ANDIJAN STATE UNIVERSITY

Tadqiqot.uz

DOI 10.26739/conf_20/05/2020_2

MATERIALS

OF THE INTERNATIONAL ONLINE DISTANCE CONFERENCE ON

MODERN INFORMATICS AND ITS TEACHING METHODS (MITM2020)

SECTION II.
INFORMATION PROCESSING METHODS AND ALGORITHMS

PART - 2

ANDIJAN
UZBEKISTAN

2020

Modern informatics and its teaching methods (MITM2020) [Electronic resource]**:** collection of materials of the international scientific-practical conference. - Andijan, May 20, 2020 Section II. Information processing methods and algorithms DOI 10.26739/conf_20/05/2020_2

The collection includes abstracts of the international scientific-practical conference "Modern informatics and its teaching methods (MITM2020)", which was held on May 20, 2020 in Andijan.

The international scientific-practical online-Internet conference includes scientific theses in the field of education, which includes the achievements of modern computer science and its teaching methods.

The conference was divided into four sections:

- modern methods of teaching informatics in high and medium education;
- modern informatics and management;
- methods and algorithms for processing information;
- the role of information and communication technologies in preschool education.

Abstracts are published in the original edition (with corrections and revisions). Authors of published materials are responsible for the accuracy of the information, quotes and information. Their opinion may not coincide with the opinion of the editorial board.

SECTION II. INFORMATION PROCESSING METHODS AND ALGORITHMS

ОБОСНОВАНИЕ ГЕОФИЛЬТРАЦИОННОЙ СХЕМАТИЗАЦИИ МАТЕМАТИЧЕСКОЙ МОДЕЛИ ДЛЯ ОПРЕДЕЛЕНИЯ РАСЧЕТНЫХ ГИДРОГЕОЛОГИЧЕСКИХ ПАРАМЕТРОВ
(на примере Дамходжинского водозабора)

Р.А.Юсупов, Ш.С.Ахралов, Жуманов Ж.Ж., Ташкентский университет информационных технологий, Узбекистан.

Аннотация. *Тадқиқотда Зарафшан дарёси водийсидаги Дамхўжа сув олиш иншоати мисолида гидрогеологик тизимларнинг тавсилотларини ҳисоблаш жараёнларини моделлаштириш учун геофильтрация схемалаш масалалари ҳақида фикр юритилади.*

Калит сўзлар*: ер ости сувини тутувчи қатлам, математик моделлаштириш, фильтрация жараёнлари, геофильтрацион схемалаштириш.*

Аннотация. *В работе рассмотриваются вопросы геофильтрационной схематизации для определения расчетных параметров гидрогеологических систем при моделирование на примере Дамходжинского водозабора долины р.Зарафшан.*

Ключевые слова: *водоносный горизонт, математическое моделирование, фильтрационные процессы, геофильтрационные схематизации.*

Annotation. *In the article is considering the questions of geofiltration schematics for the determination calculi parameters of hydrogeological systems to the modeling on example Damxodja water intake are of the Zarafshan river valley.*

Key words: *aquifers, mathematical modeling, filtration of processes, geofiltration schematics.*

Геофильтрационная схематизация проводятся на основе оценки фильтрационных параметров водоносных горизонтов по результатам опытно-фильтрационных работ (ОФР) и опирается на солидный теоретический базис физико-математические основы фильтрации, созданный трудами Ч.Тейса, Ч.Джейкоба, М.Хантуша, В.А.Мироненко, Л.С.Язвина, В.М.Шестакова и др. С точки зрения математического модели наиболее полный обзор имеющихся методик расчета изложены в работе Ф.Б.Абуталиева, И.Хабибуллаева, Р.Усманова, Ж.Х.Джуманова и др. в которой приведены теоретические зависимости, описывающие установившийся и неустановившийся характер процесса снижения уровня подземных вод при откачках из совершенных и несовершенных скважин для типов водоносных горизонтов (напорный, полунапорный, полубезнапорный, безнапорный) рассматривающая гидравлику скважины, работающей в условиях перетекания. Все эти работы,

а также многочисленные отдельные зарубежные и отечественные публикации, посвященные вопросам геофильтрационной схематизации и оценки геофильтрационных параметров пластов по данным ОФР, были проанализированы с точки зрения достоверности методик расчета научно-исследовательские работы.

В частности, что каждая гидрогеологическая задача расчленяется на геологическую и механико-математическую составляющие, а также геоинформационного обеспечения. С этой точки зрения составлены комплекс программ и апробирована на тестовых примерах, показавших определенную степень совпадения рассчитанных на компьютере параметров с заданными. Была выполнена выборочная интерпретация ОФР из водоносных горизонтов различного типа на основе моделирования, которая показала реальность рассчитанных параметров и их близость к величинам, полученным при расчета аналитическую и с применением компьютерного моделирования.

Механико-математической и расчетный блок включает в себя набор различных модели, методов и схем, позволяющих от приближенных оценок параметров перейти к достаточно обоснованным их значениям. Геоинформационных блок включает в себя территориально распределенных характеристик об объекте, географические расположение, геометрические параметры, т.е, необходимую информацию о конструкции скважин и водоносного пласта, об условиях проведения опытных работ, об исходных данных к расчету, о других признаках, характеризующих водоносный горизонт. Эти блоки представляются обязательной структурой технологии моделировании, как при аналитических расчетов вручную, так и при автоматизации их на компьютере.

Обработки результатов ОФР на основе компьютерного моделирования показала, что комплексный подход в использовании методик расчета позволяет выйти на достоверные параметры пласта при стабильных их значениях по всем пьезометрам. При этом в основу логики расчетной структуры легли теоретические представления о процессах, происходящих в системе «скважина и водоносный пласт» с учетом индивидуальности каждого из этих объектов и из взаимодействия при откачках. Они подтверждены данными большого объема опытных работ по водоносным горизонтам различного типа, следующим образом:

-гетерогенность пласта по упругости проявляется во всех типах водоносных горизонтов за исключением собственно напорных, которые развиты на больших глубинах (1500 – 2000 м);

- для водоносных горизонтов, развитых в четвертичных отложениях, которые является объектами изучения Дамходжинской водозаборе долины р.Зарафшан Самаркандской области, гетерогенность горизонтов по упругости максимально проявляется в верхней части разреза (сырдарьинско-голодностепский комплекс), где водоносные горизонты являются собственно безнапорными, и уменьшается с глубиной с переходом горизонтов в зависимости от степени изолированности в потенциально безнапорные;

Анализ гидродинамического режима подземных вод в естественных и нарушенных условиях, решение специфических задач региональной динамики подземных вод (оценка условий питания и разгрузки водоносных структур, анализ региональной взаимосвязи водоносных комплексов, первоначально коэффициент водопроводимости оценивается по приближенной зависимости для безнапорного пласта: $T=100*q$, где q– удельный дебит, л/с; T – коэффициент водопроводимости опробуемого слоя, м^2/сут. Эта величина выступает как минимальное значение параметров, которое может быть принято в качестве предварительного ориентира.

Значения параметров по результатам обработки данных ОФР на ПЭВМ

Талица 1

№ скв.	T	a	B
1ок	1400	$1,5*10^4$	-
2ок	1155	$1,2*10^4$	-
3ок	140	$2*10^5$	140
4ок	154	$2*10^5$	125
5ок	300	$1,7*10^5$	-
6ок/20	800	$0,9*10^4$	-
6ок/75	200	$2*10^5$	70,0
7ок	1200	$2*10^4$	-
8ок	300	$1,76*10^4$	250
9ок	1000	$1,3*10^4$	-
6ок/20	1000	$2*10^4$	-
6ок/75	199	-	150-
11ок	400	$4*10^5$	60

Для квазинапорных пластов числовой коэффициент может быть увеличен до 125, а для собственно напорных – до 150. Для безнапорных горизонтов при расположении скважин около реки числовий коэффициент должен быть уменьшен до 50-70.

На начальных этапах откачки (восстановления) в основе закономерностей снижения уровня лежат зависимости: $S=Q/4\pi T*W(u)$; где Q – дебит опытной откачки для начального этапа ОФР, м^2/сут; $W(u)$– интегральные функции скважин, (для напорного пласта с перетеканием через разделяющие слои), безразмерная величина; $u=r^2/a*t$, аргумент функций; м^2/сут; a*–коэффициент пьезопроводности, м^2/сут; $r=\mu/B$ – расстояние наблюдательных скважин от центральной, безразмерная величина; B – параметр дренирования, характеризующий процесс замедления проявления реальной (гравитационной) водоотдачи; t – время, сут;

На основе линейной анаморфозы определяется T– коэффициент водопроводимости пласта, м^2/сут; –временного прослеживания с использованием характеристик прямой линии, проходящей через опытные точки по зависимостям: $T=0,183*Q/I*a = R^2/2,25*T_0$, где, Q – дебит опытной откачки для начального этапа ОФР, м^2/сут; I – уклон прямой, осредняющей точки 1 этапа; T_0 – значение времени на шкале $lg\ t$ в точке пересечения её с прямой, сут. a – коэффициент пьезопроводности пласта, м^2/сут.

Метод заключается в совмещении фактических данных с эталонными при условии соблюдения параллельности осей координат, снятии с обоих кривых характеристик одной любой точки, которые служат основанием для расчета параметров T, a, B.

Для всех этапов расчет параметров можно проводить площадными метода-ми, т.е. используя данные понижений на различные моменты времени одновре-менно в нескольких наблюдательных скважинах. Для

характеризующихся к стационарному режимом, расчет пара-метров проводится по зависимостям: $T=0,366\ Q/C_z$; $B=0,89*r_0$; где C_z – уклон прямой графика S –$lg\ r$, осредняющей понижения в наблюдательных скважин; r_0 – точка пересечения графика на оси lgr; а по этапам характеризующихся нестационарным режимом, зависи-мостями: $T=0,366\ Q/C_r$; $a=0,44\ r_0^2/t_p^6$; где t_p – время от начала откачки (восстановления), на которое зафиксированы анализируемые понижения в наблюдательных скважинах, сут.

По каждому из этапов полученная *величина коэффициента водопроводимости* должна иметь примерно одинаковое значение, стабильное, как для разных этапов хода откачки (восстановления), так и для различных пьезометров опытного куста. Расчет параметров выполнен по результатам опытных откачек из скважин всех партий, которые проводили гидрогеологические исследования в области возможного влияния Дамходжинского водозаборного узла (Дамходжинской, Миankальской, Челекской, Верхнезарафшанской). Данные опытно-фильтрационных работ приведены в таблице 1. Скважины является преимущественно одиночными и опробовали отдельные интервалы водоносных горизонтов продолжительности проведены только по кустам разведочных скважин, пробуренных в процессе детальных разведок на современных Дамходжинском и Муртакском водозаборных участках.

Расчет параметров проведен по данным опытно-кустовых откачек из кустов 1ок-12ок, на территории эксплуатируемого в настоящее время Дамходжинского водозабора, перерасчет параметров по данным этих опытных откачек направлен на уточнение коэффициентов водопроводимости (T) и пьзопроводности (a^*) основного эксплуатируемого горизонта и коэффициента перетекания (B) через относительно слабопроницаемый разделяющий слой. Наблюдательные скважины пробурены по направлению параллельному урезу русла реки Карадарья на расстоянии ≈ 300 м от русла. Понижения в наблюдательных скважины на конец откачки составили до 1,5-3,0 м, т.е. срезки в наблюдательных скважинах являются достаточными для надежной оценки параметров.

Для опытного куста *3ок* и 9ок полученные значения коэффициента водопроводимости находятся в интервале от 137 м^2/сут до 299 м^2/сут, а расчеты проведенные по данным понижений в наблюдательной скважине 3а дают величину T= от 134 м^2/сут до 150 м^2/сут. Значение подтверждается расчетами по данным восстановления, для опытного куста 3ок приняты значения: T=137 м^2/сут, a=2,67*10^5 м^2/сут, B=154 м (таблица 1).

Расчеты, проведенные по данным опытных работ в скважинах опытного куста 11ок, дают более низкие значения коэффициента водопроводимости (от 411 м^2/сут до 739 м^2/сут). Наиболее стабильные значения Т получены по данным в пьезометре 11а и комбиниронным способом методом Хантуша (от 441 м^2/сут до 474 м^2/сут). Среднее значение Т=456 м^2/сут принято за достоверное, а=4,35*10^5 м^2/сут, В=155 м.

Результаты расчета параметров по данным ОФР. При расчете параметров исходили из следующих положений: -продуктивный водоносный комплекс, приуроченный к аллювиальным современным и верхне-четвертичным отложениям является безнапорным в своей верхней части (до глубины 10-20м) и субнапорным (потенциально-безнапорным) в своей нижней части (глубины 35-55м); -наличие слабопроницаемого пласта, разделяющего верхний и нижний водоносные горизонты, обуславливает в процессе опробования одного из горизонтов привлечение, путем перетекания через этот пласт, подземных вод смежного горизонта. Такой процесс наиболее характерен при опробовании нижнего водоносного горизонта; -трехслойное строение продуктивного водоносного комплекса (2 водоносных горизонта, разделенных слабопроницаемым пластом) характерно для южной части месторождения, в которой располагается и действующий водозабор.

Литература.

1. Мироненко В.А. Динамика подземных вод. –М. Изд-во МосГосГорУнивер. 2009. 529 с.

2. Бэр Я., Заелавеки Д. Физико-математические основы фильтрации воды. -М., Мир, 1971.

3. Айтметов Р., Мирюсупов Ф.М., Нагевич П.П. Проведение доразведки эксплуатируемого месторождения современной долина р.Зарафшана участке действующего водозабора Дамходжа в связи с изменением водохозяйственных условий. Отчет Дамходжинская ГГП на 2001-2003, -Ташкент, 2004, ГГП «Узбекгидрогеология».

4. Нагевич П.П. – Формирование и оценка фильтрационных свойств водоносных толщ. -Т., ГИДРОИНГЕО, 1998.

ANALYSIS OF CLUSTERING ALGORITHMS OF WEKA TOOL ON AUTOMOBILE DATASET

Muhammad Umer Ijaz, Department of Computer Science, Riphah, Lahore, Pakistan.

Abstract: *Data mining is production of interesting patterns (relevant, constructive, unpredictable and very important) or data from a large data stack. In other words, it is a combination of links, associations and whole patterns contained in massive databases however hidden or unknown. So as to perform the analysis, we'd like software system and tools. Set of tool, those permits to user analyze information for various perspectives and angles, in order to find meaningful relationships. In this paper, we study and compare the varying algorithms and technique used the group analysis that is used for Weka tool. Our comparisons of these seven clustering algorithms according to their executed time, the number of these iterations, the estimated bugs rate the log chance. Finally on the basis of the result obtained we analyze and judge the performance of the algorithms in relation to each other.*

Keyword: *Clustering, Automobile, Weka Tool*

I. INTRODUCTION:

At present, the process of extracting important information and verity from knowledge has become a lot of scientific than scientific. Even before information is collected and processed, the kind of information that will be extract data that lays the human brain, which is why human understanding remains undeniable. Various data extraction techniques were developed, each designed for a specific set of data. Consolidation is a "natural" classification method for the labeling of data objects that are not written in such a way that one-member objects are different from objects in another group. It can be considered as the most important and most important learning process that can be managed in Data Mining. Clustering is responsible for obtaining structure from an unread data group. There is some kind of similarity among these object which will be present within same clusters and at an equivalent time, dissimilarity is observed among the objects to belong different clusters. In clustering algorithms are wont to organize, model, reason and compress information [1]. Throughout the analysis, the input datasets and thus the clusters used area unit varied in range to measure the performance of cluster algorithms. In the paper we are discussing the various clustering approaches and techniques utilized in processing then within the part, we are compared and analysis of different algorithm various factor.

CLUSTER ANALYSIS IN WEKA

Data Mining is not exclusively perform by application tool and software, as the fact, there is many tools that act to counter part to expensive tools. These tools are also called Weka[2], These tools are easily used newly data science worker. The toolkit has been very widely used to remainder for extended time. Weka tool

written in java language and provide graphic user interface which is used to connected the data set and generated visualization output, such like as graph, table and more other. In the paper we used three types on clustering technique these names are Partitioning method, Density based method and Hierarchal methods.

II. REVIEW OF LITERATURE

Few of the researchers to improved clustering algorithms whereas others have enforced new ones, there are few others who have analyzed and compared the already existing clustering algorithms. [5] In their work, they need discussed the technique and have showed the performance of the algorithms with reference to the executing time and at speed.[6]. Researchers have used agglomerative technique so as to create a simple heuristic method is used for dendrogram and partition the info. They studied about the similarity based agglomerative clustering algorithms and presented its effectiveness. Compared the various clustering algorithms consistent with the subsequent component of data set ,number of these clusters, variety of data set and there for software used to clustering.

III. METHODLOGY OR MATERIALS AND METHODS.

CLUSTERING CONCEPTS

Clustering referred to unsupervised learning. A dataset divide into the partition a number to disjoint or overlapping groups.

Clustering referred to the natural clustering of the information objects in such the simplest way that the objects within the same group are similar with reference to the objects gift within the different teams.

Details of the Data Set

Dataset table is show below the attribute of number and the instance of number to the given database.

Table 1 In the given dataset table shows that instance and attributes.

Name of Dataset	*No of Attributes*	*No of Instance*
Automobile Dataset	*26*	*205*

In This Paper work to analysis of various techniques that are used those algorithms are:

- K-means Algorithm.
- Hierarchal clustering
- EM
- MDB Cluster
- Canopy
- Filtering Cluster
- Farthest First

K-Mean Algorithm:

K-means is an entire clustering algorithm. Every objects that is assign to precisely set of a one cluster. For these methods of cluster we will start to decide how much cluster we will like from our dataset. So we will call the value of k. These k values is suppose smallest integer like as 2, 3, or 4, but they may be large. So we move backed and decide that which values of k will be selected in the given values are there. [3]

Figure 2. K-mean Algorithm

Hierarchical algorithm:

In the hierarchical algorithm the integration of the hierarchy, a state of things is created. These two types are group cohesion: agglomeration (top down) & division (top down). The linear combinations, we will started one data of objects and we start with one data object and then moderately form a group when it is in the process of classification; We will started with all of dataset and dividing the data into groups.[2]

Figure 3. Hierarchical Clustering

Expectation–maximization Algorithm:

In using of em clustering technique that estimate mean and std deviation of each cluster. So as to maximize the like hood of cluster data set. In the paper em algorithm work deeply to make a cluster of given dataset. The em clustering to make cluster 11 of different categories. So number of irritation 4.

Figure 4. EM Algorithm

FARTHEST FIRST CLUSTERING:

Farthest initial could be a variant of K means places every cluster centre successively at the purpose furthest from the present cluster centre. This time should lay at intervals the info space. This greatly accelerates the clump in most of the cases since less duty assignment and adjustment is required.

Figure 5. Farthest first

Canopy:

Canopy bunch is usually used as associate degree initial step in additional rigorous bunch techniques, like K-Means bunch. By beginning with associate degree initial bunch the amount of dearer distance measurements will be considerably reduced by ignoring points outside of the initial canopies.

Figure 6. Canopy

MDB Cluster:

The Density-based bunch tool works by detection square measures wherever point's square measure targeted and wherever they're separated by areas that are empty or distributed. ... These algorithms square measure thought of unattended as a result of they are doing not need any coaching on what it means that to be a cluster.

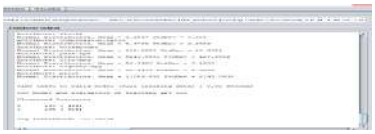


Figure 7. MDB Cluster

K-means clustering:

In K-means algorithm show the cluster to highest instance of this case. K-means algorithm makes two clusters. Cluster (1) 57% and cluster (2) 43% show in the fig 8

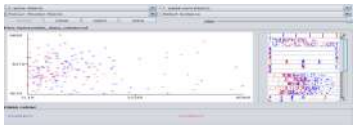


Figure 8. Visual K-mean Clustering

In k-mean clustering we find out to comparison between price and peak rpm. That may shows the cluster the highest range of rpm that shown in the fig 9.

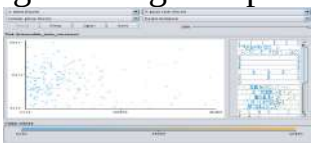


Figure 8. K-means Clustering Analysis Figure

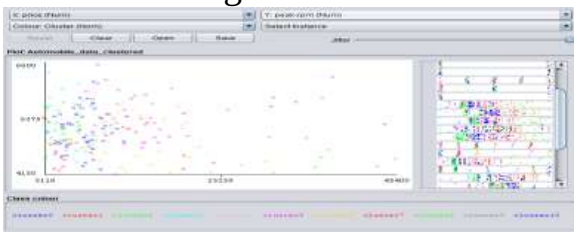
K-means clustering to highest price out of peak rpm in cluster 1. Show in the fig

K-means Cluster 1 Performance

K-means clustering that cluster 0 to group that to low price to comparative to other cluster.

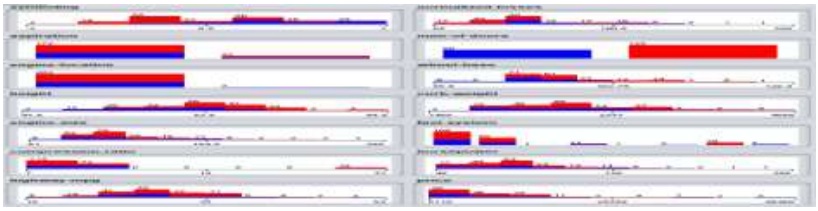
K-means Cluster 2 Performance

Expectation–maximization Algorithm

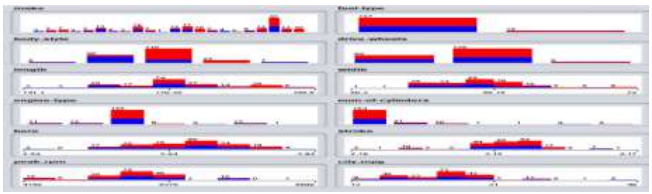


Visual E-M Clustering figure 10

Visualization Dataset:



Dataset visualization



EXPERIMENTAL SETUP

In the paper we work out to comparison of clustering algorithms using of Weka tool. Weka tools are data mining tool that consist of machine learning algorithm. Weka tool consists for preprocessing, clustering, classification, association rules, regression and visualization of data. So we use dataset Automobile dataset. In our working we comparison the seven clustering algorithm these are (K-means Algorithm, Hierarchal clustering, EM, MDB Cluster, Canopy, Filtering Cluster) on basis of the no of clusters, no off instance, Square errors, and how much time are required to build a model and log likelihood.

IV. RESULTS & DISCUSSION

The following table shows the final result of clustering algorithm is used. Table represents the analysis process of all four algorithms.

Table 1: Comparison result of different clustering algorithms

Name	***Data set name***	***No. Of clusters***	***Cluster distribution***	***No of iterations***	***Sum of squared error***	***Log likelihood***	***Time taken to build model(sec)***
K means	*Automobile*	*2*	*116 (57%) 89 (43%)*	*4*	*697.325*		*0.03*
Hierarchal clustering	*Automobile*	*2*	*202(99%) 3(1%)*				*0.025*

EM	*Automobile*	*11*	*15(7%)* *10(5%)* *38(19%)* *12(6%)* *18(9%)* *30(15%)* *14(7%)* *12(6%)* *8(4%)* *16(8%)* *32(16%)*	*2*		*-58.7033*	*10.22*
MDB Cluster	*Automobile*	*2*	*100(49%)* *105(51%)*	*4*	*697.325*	*-71.191*	*0.03*
Farthest first	*Automobile*	*2*	*76(37%)* *129(63%)*				*0.02*
Canopy	*Automobile*	*3*	*5 (2%)* *31 (15%)* *14 (7%)* *6 (3%)* *5 (2%)* *7 (3%)* *16 (8%)* *3 (1%)* *40 (20%)* *8 (4%)* *27 (13%)* *6 (3%)* *8 (4%)* *4 (2%)* *4 (2%)* *5 (2%)* *6 (3%)* *10(5%)*				*0.02*
Filtering Cluster	*Automobile*	*2*	*116(57%)* *89(43%)*	*4*	*697.325*		*0.02*

V. CONCLUSIONS

Recently data mining is the branch of computer science and information technology. There are large amount of data is spreading all over the world. In data mining branch are used in education, medical, and banking etc. To get relevant data to mining technique used. Working of data mining, you have to knowledge of algorithm [4]. We have performing seven algorithm these are K-means Algorithm, Hierarchal clustering, EM, MDB Cluster, Canopy, Filtering Cluster. All of these algorithm results are generate similar object and time that create clusters. For using of EM algorithm that take consume more time than other of k-means, hierarchical, canopy, and other. Best of working algorithm is k-means clustering. It performance of result are very efficient and take less time than other clustering algorithm to finding similarity cluster through Weka tool for automobile dataset.

REFERENCES

[1] Rajneet Kaur and Aastha Joshi "Review: Comparison of clustering in data mining" CSSE, 2013

[2] Andrew: "Hierarchical and K-means Clustering - Tutorial".

[3] DATA MINING AND KNOWLEDGE DISCOVERY HANDBOOK.

[4] Shai Ben and Reza Zadeh "Theorem for Clustering", Stanford.edu.

[5] Chen G , Tanaka T, Banerjee N, Zhang M and Kom: "comparison clustering algorithms analysis to express data", Statistics Sinica, vol 12

[6] Osama Abbas: "Compare b/w data clustering Algorithms", IAJIT, Vol. 5, 2008

ПОВЫШЕНИЯ КАЧЕСТВА ИЗОБРАЖЕНИЙ НА ОСНОВЕ ЭКВАЛИЗАЦИИ ГИСТОГРАММЫ

Равшанов Н., Суванов Ш.М., Научно-инновационный центр информационно-коммуникационных технологий, Узбекистан.

***Аннотация**: В данной статье предоставлен метод повышения контрастности изображения путем применения эквализация гистограммы к одной из частей изображения. Изображение делится на две части на основе средней интенсивности пикселей. Эквализация гистограммы применяется к одной из двух частей, в то время как другая часть остается неизменной.*

***Ключевые слова:** Эквализация гистограммы, контрастность изображения, яркость.*

***Annotation**: This paper presents a method to enhance the contrast of an image by applying histogram equalization to a part of the image. The image is partitioned into two parts based on the pixel intensities with respect to the median intensity. Histogram equalization is applied to either one of the two parts while keeping the other part unchanged as in the image.*

***Keywords:** histogram equalization, image contrast enhancement, sub-image.*

Улучшение контрастности изображения нацелено на повышение качества видимости изображения. Эквализация гистограммы (ЭГ) является одним из распространенных методов, используемых для улучшения контраста в цифровых изображениях [1]. ЭГ распределяет интенсивность пикселей по всему диапазону, используя функцию распределения. ЭГ позволяет областям с более низким контрастом принимать более высокий контраст и следовательно делает невидимые детали видимыми. В изображении ЭГ полезнее когда фон и передний план является яркими или темными оттенками цвета.

Адаптивное ЭГ (АЭГ) является одним из ЭГ для того чтобы усилить контраст изображения, для этого нужно использовать ЭГ по всему широкому диапазону [2]. АЭГ использует несколько гистограмм, каждая из них соответствует отдельной части изображения. Однако АЭГ излишне усиливает шум в относительно однородных областях изображения.

В этой статье представлен метод повышения контраста изображения путем применения ЭГ к части изображения. Для начала мы разделяем изображение на две части относительно средней интенсивности. Мы применяем ЭГ к одной части изображения, в то время как другая часть остается неизменной. Выбор части для применения ЭГ основан на положении высшей точки интенсивности в гистограмме изображения. Объединение двух частей гистограммы создает более контрастное изображение по сравнению с применением глобального ЭГ к изображению.

Метод который используется в данной статье, является одним из вариантов АЭГ. АЭГ предназначен для локального улучшения контрастности изображения в нескольких областях, но наш метод предназначен для повышения общего контраста изображения. В данной работе мы показываем наши результаты, в которых сравниваем качество изображения с нашим методом и с глобальным ЭГ.

Предложенный метод эквализация гистограммы использован для повышения контрастности изображения. Исходя из входного изображения в оттенках серого с диапазоном интенсивности [0:255], мы разбили изображение на две части относительно его средней интенсивности. Одна часть состоит из пикселей, интенсивность которых находится в пределах [0:127], называемых левым под-изображением, а другая часть состоит из пикселей, интенсивность которых находится в пределах [128:255], называемых правым под-изображением. Мы выбираем либо левым под-изображением, либо правым под-изображением, в которых задействована высшая точка интенсивности для входного изображения, чтобы применить эквализацию гистограммы.

Применяя эквализацию гистограммы к одному из под-изображения, в то время как другая остается неизменной, изменяется интенсивность пикселей как на темной стороне, так и на яркой стороне изображения. После этого объединение их в одно изображение создает более контрастное изображение, чем применение глобальной эквализации гистограммы ко всему изображению. Предлагаемый способ суммирован на рисунке 1.

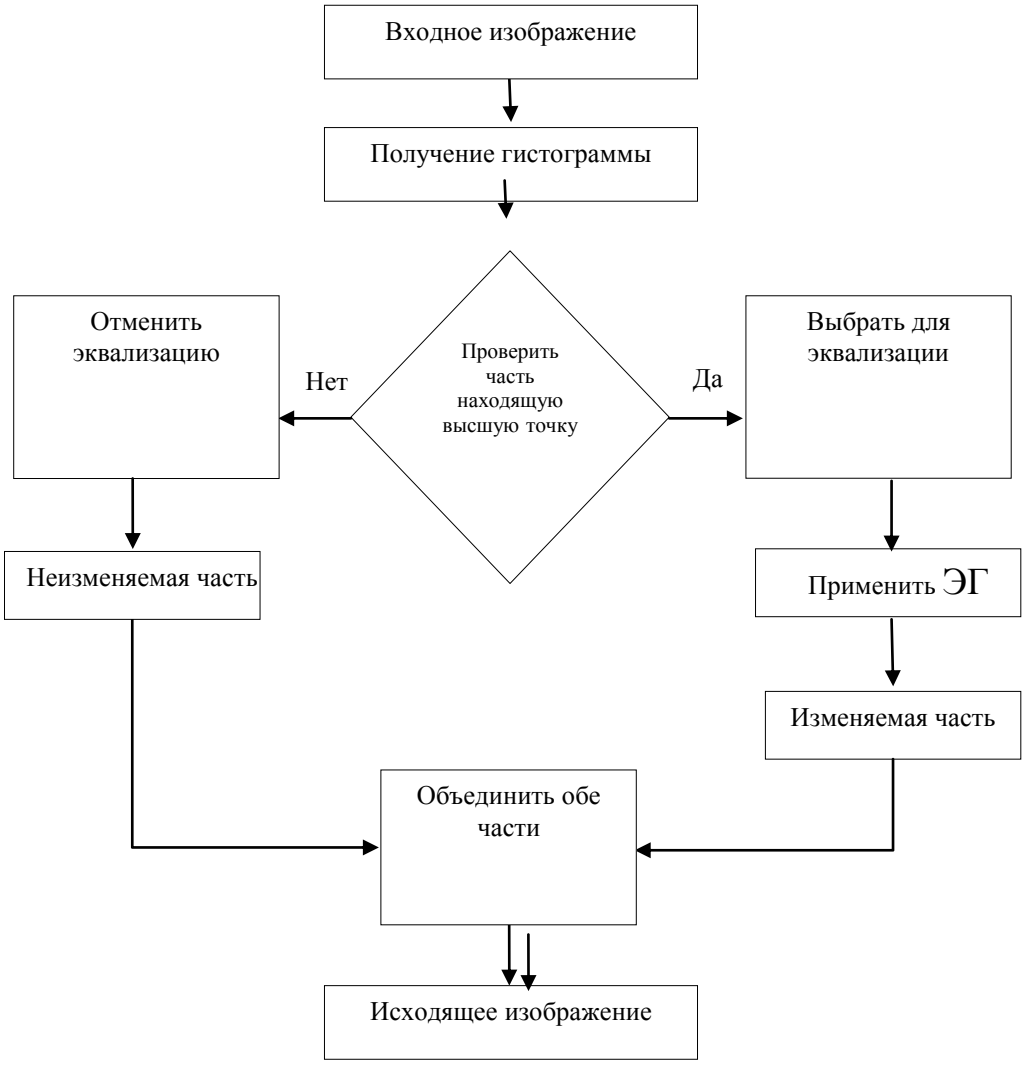


Рисунок 1 – Блок-схема для данного метода

Таким образом, повышения контрастности изображения путем эквализации к одному из частей изображения путем применения гистограммы. Изображения были разделены на две части относительно средней интенсивности пикселей.

Литература

[1] R. C. Gonzalez and R. E. Woods, Digital Image Processing 3rd Ed., Prentice Hall, 2007.

[2] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. T. H. Romeny, J. B. Zimmerman, Adaptive histogram equalization and its variations, Computer Vision, Graphics, and Image Processing, 39(3):355-368, 1987.

ELEKTROKARDIOGRAMMA SIGNALLARINI TAHLIL QILUVCHI MOBIL ILOVA DASTURINI LOYIHALASH.

Kuchkorov T.A.,[1] Jo'rayev A.I.,[1] Odinaev M.X., [1] Tojiboyeva Yo.R.[2]
[1]Muhammad al-Xorazmiy nomidagi Toshkent Axborot Texnologiyalari Universiteti, [2]Andijon davlat tibbiyot instituti, O'zbekiston.

Annotatsiya. *Ushbu maqolada yurak - qon tomir kasalliklarini o'z vaqtida tezkor ravishda tashxislash uchun elektrokardiogramma signallarini tahlil qiluvchi mobil ilova dasturini ishlab chiqish masalalari ko'rib chiqilgan.*

Kalit so'zlar: *Elektrokardiografiya, konvertor, bluetooth, server.*

Аннотация*. В этой статье обсуждается разработка мобильного приложения, которое анализирует сигналы электрокардиограммы для своевременной диагностики сердечно-сосудистых заболеваний.*

Ключевые слова: *электрокардиография, конвертер, bluetooth, сервер.*

Annotation*. This article discusses the development of a mobile application that analyzes electrocardiogram signals for the timely diagnosis of cardiovascular diseases.*

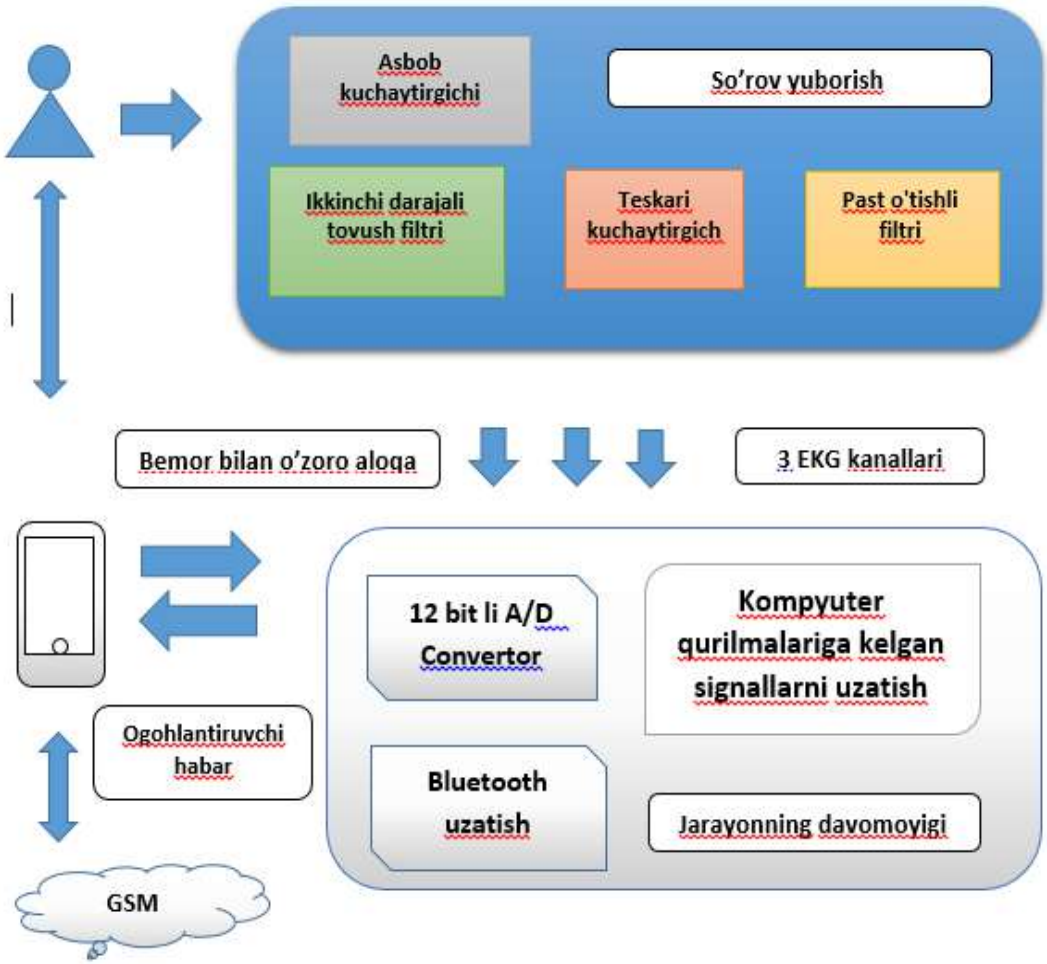
Key words*: electrocardiography, converter, bluetooth, server.*

Yurak - qon tomir kasalliklari O'zbekistonda keksa yoshdagi insonlar o'limning yetakchi sabablaridan biri hisoblanadi. Ushbu oqibatlarning tarqalishiga nafaqat tashqi ta'sirlar balki nosog'lom turmush tarzi sabab bo'lishi ham mumkin.

Hozirgi kunda barcha qurilmalar, raqamli texnologiya signallari orqali tegishli natijalarni 86% aniqlikda olishga erishilmoqda. Aynan ushbu natijalar orasida insonni holatini kuzatishning ko'plab usullari orasida elektrokardiografiya (EKG) yetakchi o'rinni egallaydi. EKG signali katta miqdordagi ma'lumotni o'z ichiga oladi va bemorning EKG signalini batafsil avtomatik tahlil qilish, noxush holatlarga to'sqinlik qiluvchi signallarni o'z vaqtida ishlab chiqarishga imkon beradi. EKG signalini tahlil qilishda birinchi navbatda shovqinni filtrlash, undagi kardiotsikllarni tanlash va yurak urish tezligini (HR) hisoblash va shu bilan yurak faoliyatini nazorat qilish kerak. Qurilma signalni qabul qilayotganda kardiyosikllarga bo'lish algoritmining kiritilishida dastlabki ishlov berish bosqichlaridan o'tgan signal uzatiladi, ya'ni shovqinlarni filtrlash va signalni siqish. Ushbu bosqichning asosiy maqsadi EKG signalini ECSda yurak qisqarishiga mos keladigan xarakterli joylarni tanib olish yo'li bilan olingan kardiyosikllarning ketma-ketligi shaklida namoyish etishdir.

Bizning ushbu loyiha ishimizda, insonlarni yurak bilan bog'liq holatlar bezovta qilganida, shifokor qabuliga borib kardiogramma (EKG) qildirishlari shart emas. Foydalanuvchilar o'zlarining mobil telefonlariga biz ishlab chiqayotgan EKG mobile ilovasini yuklaydilar va yurak urishini yozib olish uchun elektrod (липучка) orqali ko'krak qafasiga qo'yadilar va olingan ma'lumotlarni mobil telefon orqali serverga yoki o'zlaring doimiy qabulida bo'lib turadigan shifokorlariga yuboradilar. Server olingan EKG ma'lumotlarini avtomatik

taxrirlaydi va mobil telefonga javob yuboradi yoki shifokor masofadan turib qanday choralarni ko'rishlari kerakligini aytadi.



1-rasm: Ishlab chiqilayotgan tizimning ishlash strukturasi va EKG signallari, signallarni qayta ishlash va bemor bilan interfeys holatlari ko'rsatilgan. Bu uyali telefon orqali simsiz aloqa hodisalarni yozib olish va signal signallarini yaratishga imkon beradi.

2-rasm: EKG mobile ilovasining funksional tuzilmasi

Uzatiladigan tibbiy ma'lumot va favqulodda vaziyat yuzaga kelganda, shifokorga tibbiy yordam zarurati haqida xabar beradi. Belgilangan davr uchun bemorning ahvoli haqidagi maxsus hisobot, jadvallashtirish va bemorning holatini tahlil qilish shifokor tomonidan amalga oshiriladi. EKG mobile olgan ma'lumotlarini avtomatik ravishda veb serverga yuboradi, veb serverda olingan ma'lumotlar tahlil qilinib bemorning o'ziga yoki bemorning shifokoriga yuboriladi.

a b

3-rasm: a-mobil ilova ishga tushirilgandagi ko'rinishi, b-mobil ilova EKG singnallarini yozib olish jarayoni

Ushbu loyihada ishlab chiqilayotgan mobil ilova ko'proq doimiy ravishda yurak hastaliklari bilan og'rigan, yani aritmiya yoki stenokardiyasi bor, infarkt o'tgazgan bemorlar uchun ayni muddaodir. Odatda bunday bemorlar doimiy ravishda ma'lum kardiolog nazoratlarida bo'lishlarini lozim. Agarda bezovtalik holatlari yuzaga kelsa bemor shifokor qabuliga borib, soatlab navbat kutib, vaqtini ketgazishi mumkin. Bazan esa bu bezovtalik sababi yurak bilan bog'liq bo'lmay ovora bo'lishlari ham mumkin. Shunday paytlarda darxol EKG mobile ilovasini ishga tushirib, u orqali kerakli EKG ni olib, shifokorga yuborilsa, uning tavsiyasi bilan biror choralar ko'riladi. EKG ni yozish jarayoni yakunlangandan so'ng, mobil ilovada yurak pulsi hamda stress holatlari ham aks etadi. Ushbu EKG ni PDF fayl shaklida saqlash imkoniyati ham mavjud yoki to'g'ridan to'g'ri shifokorning elektron pochtasiga yuborish imkoni mavjud.

Adabiyotlar:

1. *Samuel E. de Lucena, Daniel J. B. S. Sampaio," ECG Monitoring Using Android Mobile Phone and Bluetooth", 2015 .*
2. *Sandeepa. B Diagnosis of EKG arrhythmias in wireless telecardiology. 2018.*
3. *E. K. Hodson, D. R. Thayer, and C. Frankli (1981): Adaptive Gaussian filtering and local frequency estimates using local curvature analysis, IEEE Trans. A coust. Speech Sig. Proc. 29, 854-859.*

ФИЗИК ЎЛЧАШЛАРНИ КОМПЬЮТЕРЛАШТИРИШНИНГ АФЗАЛЛИКЛАРИ

А.Артиқов, Андижон давлат университети, Ўзбекистон

Annotatsiya: *Маърузада информацион технологияларнинг ўқув жараёнларига қўлланилишининг айрим дидактик, таълимий ва дастурий асослари билан боғлиқ бўлган айрим масалаларга тўхталиб ўтилади. Кейинги йилларда физика практикумларини компьютерлаштириш борасида олиб борилаётган изланишлар ва уларнинг айрим натижалари ҳам баён қилинади.*

Калит сўзлар ва иборалар: *дидактик, дастурий, физика практикумлари, компьютерлаштириш, технологиялар, интерфейс, интерфейс асослари, радиотехника, автоматика*

Аннотация: *В докладе рассматриваются некоторые вопросы, связанные с некоторыми дидактическими, образовательными и программными основами применения информационных технологий в образовательных процессах. В последующие годы будут также объяснены исследования по компьютеризации физических практик и некоторые их результаты.*

Ключевые слова и фразы: *дидактика, Программирование, физический практикум, компьютеризация, технология, интерфейс, основы интерфейса, радиотехника, автоматизация*

Annotation: *The report discusses some issues related to some didactic, educational and programmatic foundations of the use of information technologies in educational processes. In the following years, research on computerization of physical practices and some of their results will also be explained.*

Key words and phrases: *didactics, Programming, physical practice, computerization, technology, interface, interface basics, radio engineering, automation*

Янги информацион технологияларнинг ўқув жараёнларига қўлланилиши ўқитиш самарадорлиги кескин ортиб боришига олиб келмоқда[1-4]. Информацион технологияларнинг ўқув жараёнларига қўлланилишида аввало таълим принципларига амал қилинади. Таълим методлари ва тамойилларисиз бундай технологияларни таълимда самарали қўллаш мумкин эмас.

Шу контекстда, физик практикумларни компьютерлаштириш борасидаги изланишлар кейинги йилларда тобора кўпроқ олиб борилмоқда. Практикумни бажариш давомида ўлчанаётган барча катталиклар тўғридан тўғри компьютерга махсус ўзгартиргич(преобразователь)лар орқали киритилади. Компьютерда ҳар қандай кичик ўзгаришларни ва бу ўзгаришлар билан боғлиқ бўлган қонуниятларни узлуксиз кузатиб бориш имконияти яратилади. Ўлчаш аниқлиги кескин ортади ва осонлашади. Ушбу қисқагина мавзуда интерфейс асосларини тўла баён этиш фикридан жуда йироқмиз,

лекин ўзбек тилида бу соҳада адабиётлар камлигини ҳисобга олиб, физик ўлчашларда интерфейсдан фойдаланишнинг механизмини қисқача баён этишни маъқул деб топдик.

Ҳозирги кунда замонавий ишлаб чиқаришда ҳам кескин ўзгаришлар содир бўлмоқда. Уларнинг моҳияти шундаки, ривожланиш оқибатида фундаментал фанларнинг ютуқлари жуда тез амалиётга кириб келмоқда, шу сабабли ўқитиш жараёнида олдинги давр (классик) лаборатория методикасига асосланган, университет ёки таълим тизимининг бошқа босқичилари шароитида бажаришнинг имконияти бўлмаган замонавий амалий ишларни бажарилиши зарур бўлиб қолди. Акс ҳолда талаба фундаментал фанни билса ҳам, уни амалиётга талаб даражасида умуман қўллай олмайди. Масалан, разрядланиш конденсатор сиғимига ва занжирнинг таркибига боғлиқ. Бундай ҳолатлар эса наинки радиотехникада, балки автоматикада катта роль ўйнайди, яъни занжир параметрларининг озгина ўзгариши билан заряд хусусиятлари кескин ўзгариб кетиши мумкин. Талаба бундай ҳолатлар тўғрисида амалий тушунчага эга бўлиши керак. Ўшандагина у амалиётда қўлланиладиган юқоридагига ўхшаш жараёнларни тушуна олади, зарур бўлса бошқара олади ёки ўз навбатида таълим муассасаси тўла қонли билим бера олади. Бу фикрлар ва танланган мисолдан мақсад компьютерсиз бундай лаборатория ишларини бажариш мумкин эмаслигини кўрсатишдан иборатдир. Демак, компьютердан фойдаланадиган бўлсак, олдинги шароитда умуман бажариб бўлмайдиган ишларни ҳам бажариш мумкин ва лаборатория ишларини сонини, мазмунини кескин оширишимиз имкони яратилади.

Яна шуни айтиш керакки, компьютер билан лаборатория ишлари бажарилганда кўп ҳолларда улар фундаментал тадқиқотларнинг ранг–баранг натижаларини ўз ичига қамраб олади. Масалан: конденсатор сиғимини аниқлашда талабада электродинамика фанининг зарядларнинг ҳаракат қонунига оид, майдонларни ўзгариш қонунларига оид билимларни батафсилроқ ўрганиш,тушуниш имконини ярата олади. Фақат доскада буларни тушунтириш анча қийин кечади. Ўқувчи ва талабалар жараённинг механизмини англаши, тушуниши деярли амалга ошмайди.

Шахсий компьютер(ШК)ларни ўқув жараёнига тадбиқ этиш учун интерфейс асосларини билиш зарур. Бугунги кунда дастурлаш усуллари жуда ривожланиб кетди. Физик тадқиқотларга тадбиқ қилинаётган дастурлар, дастурлаш кодлари(крокодил,MEPL,…) ўрганилаётган физик жараёнларнинг механизмларини, кинетикасини, геометриясини(структура) ва ҳ.к. ларни батафсил ўрганиш имконини бермоқда. Лекин физик тажриба ўтказиш жараёнида ўлчанаётган катталикларни тўғридан-тўғри ШК га киритиш орада кечадиган машаққатларни умуман йўқотиш имконини беради.

Ўлчаш натижалари бўйича информация чиқарилаётганда ташқи қурилмага, зарур ахборотлар процессор “аккумулятори”дан берилади. Ахборот компьютерга киритилаётганда у процессор “аккумулятори”га

ёзилади. Умумий ҳолатда жадвалда кўрсатилгандек киритиш–чиқариш буйруғи учта қисмдан иборат бўлади.

1. Операция коди.
2. Танлаб олинаётган масала коди.
3. Буйруқ коди.

Физик катталикларни электр сигналларга айлантирувчи ўзгартиргичлар автоматика системаларини ва ўлчаш тизимларини энг асосий элементлари ҳисобланади. Функционал ўлчов асбобларини яратиш автоматика, ҳисоб техникаси, радиоэлектроника, ахборот ўлчов техникаси ва метрология ютуқларига таянади. Физик катталикни узатувчи ахборот канали таркибига кетма–кет қуйидаги асбоблар киради. Электр сигналини ҳосил қилувчи ўзгартиргич, электр сигналини функционал ўзгартиргичи, масштабли ўзгартиргич ва сигнални фойдаланишга қулай ҳолатга келтирувчи ўзгартиргич. Буларнинг ҳаммаси ўзгартиргич деб аталади.

Ностабилловчи факторларни таъсирида ҳосил бўлган хатоликни одатда структуравий ёки консерватив усулларда йўқотишга ҳаракат қилинади. Бу докладда таклиф қилинаётган усулни эса, дастурий усул дейилади. Ўзгартиргичлар анологли ёки анолог сонли бўлиши мумкин. Бизда анолог сонли ўзгартиргичлар қўлланилган.

Физик катталикларни ўзгартиргичлари қуйидаги йўналишларда юксалиб бормоқда. Сезгирлик даражаси ва тезлиги ортяпти. Ташқи шароитга чидамлилиги ортиб бормоқда.

Албатта, юқорида айтилганларнинг барчаси ҳар бир физик параметрни сезгир элементини чизиқли диапазони учун ўринлидир. Амалда қўлланилган сезгир элементларни назарий тахлил қилиш анча қийин масала ва махсус изланишларни талаб қилади. Лекин тажрибада аниқланган хатоликлар нисбатан жуда кичик, яъни функционал системанинг аниқлиги механик аниқликка қараганда анча юқори эканлиги аниқлан[1,5-7]. Масалан, Юнг модулини аниқлаш стендида эгилиш ёйини линейка билан ўлчашга нисбатан компьютерда ўлчаш жуда катта аниқликни беради.

Адабиётлар

1. *Юров В. «Ассемблер» С.П. 2001.*
2. *Раҳмонқулова С. «IBM PC шахсий компьютерида ишлаш». Т. 1998 й.*
3. *Бегимқулов У. «Физика дарсларида информацион технологиялардан фойдаланиш»,// «Физика, математика ва информатика журнали». 2002 й.*
4. *Йўлдошев Б. Ж., Рўзимов Ш.М. «Мультимедиа тизимидан фойдаланган ҳолда физик ҳодиса ва жараёнларни моделлаштириш. // Физика, математика ва информатика». 2002 й. №4(6), 54–с.*
5. *Соучек Б. Мини ЭВМ в системах обработки информации. М., 1976, 518*
6. *Сальтиков А.И, Семашко Г.Л Программирование для всех. Москва «Наука» 1986г. 176 стр.*
7. *Бухтияров А.М, Маликов Ю.П, Фролов Г.Д Практикум по программированию на фортране. М. 1979.*

ШНОР СХЕМАСИГА АСОСЛАНГАН ЭЛЕКТРОН РАҚАМЛИ ИМЗО АЛГОРИТМИ ТАДҚИҚИ

Худойназаров У.У., Мухаммад ал-Хоразмий номидаги Тошкент ахборот технологиялари университети Фарғона филиали, Ўзбекистон.

Аннотация. *Бугунги кунда электрон рақамли имзо алгоритмлари маълумотлар аутентификациясида, электрон тижоратда, криптовалюталар соҳасида ва сайлов тизимида кенг фойдаланилмоқда. Ушбу тезисда Шнор схемасига асосланган электрон рақамли имзо алгоритмининг моҳияти баён этилади.*

Таянч сўзлар: *Криптография, электрон рақамли имзо, дискрет логарифмлаш, криптовалюталар, элекьрон хужжат, очиқ калит, ёпиқ калит.*

Аннотация: *Сегодня алгоритмы электронной цифровой подписи широко используются в аутентификации данных, электронной коммерции, криптовалютах и избирательной системе. Этот тезис описывает сущность алгоритма электронной цифровой подписи на основе схемы Шнора.*

Ключевые слова: *Криптография, электронная цифровая подпись, дискретный логарифм, криптовалюты, электронный документ, открытый ключ, закрытый ключ.*

Annotation. *Today, electronic digital signature algorithms are widely used in data authentication, e-commerce, cryptocurrencies, and the electoral system. This thesis describes the essence of the electronic digital signature algorithm based on the Schnor scheme.*

Key words: *Cryptography, electronic digital signature, discrete logarithm, cryptocurrencies, electronic document, public key, private key.*

Электрон рақамли имзо қоғозли ҳужжат алмашинувидаги анъанавий шахсий имзо хусусиятларидан фарқли бўлиб, иккилик саноқ тизими хусусиятлари билан белгиланадиган хотира регистрлари битларига боғлиқ. Хотира битларининг маълум бир кетма-кетлигидан иборат бўлган электрон имзони кўчириб бирор жойга қўйиш ёки ўзгартириш компьютерлар асосидаги алоқа тизимларида мураккаблик туғдирмайди[2]

Бугунги юқори даражада ривожланган бутун дунё цивилизациясида ҳужжатлар, жумладан махфий ҳужжатларнинг ҳам, электрон кўринишда ишлатилиши ва алоқа тизимларида узатилиши кенг қўлланилиб борилаётганлиги электрон ҳужжатлар ва электрон имзоларнинг ҳақиқийлигини аниқлаш масалалари ечимларининг муҳимлигини келтириб чиқармоқда.[2]

ЭРИ ахборот-коммуникация тармоғида электрон ҳужжат алмашинуви жараёнида қуйидаги учта масалани ечиш имконини беради:

- электрон ҳужжат манбасининг ҳақиқийлигини аниқлаш;
- электрон ҳужжат яхлитлигини (ўзгармаганлигини) текшириш;

- электрон ҳужжатга рақамли имзо қўйган субъектни муаллифликдан бош тортмаслигини таъминлаш.

Электрон рақамли имзонинг умумий тан олинган схемаси учта жараённи ўз ичига олади:

– электрон рақамли имзо калитларини генерациялаш;
– электрон рақамли имзони шакллантириш;
– электрон рақамли имзони ҳақиқийлигини тасдиқлаш

ЭРИни шакллантириш жараёни

Имзо қўйиш муаллиф томонидан, фақат унга маълум бўлган шахсий калит билан амалга оширилади. Имзонинг ҳақиқийлигини текшириш эса исталган шахс томонидан, имзо муаллифининг очиқ калити билан амалга оширилиши мумкин.[3] Электрон рақамли имзо алгоритмларини шакллантириш ҳам ҳозирги кунда дастурий, аппарат ва аппарат дастурий воситалардан фойдаланилмоқда.

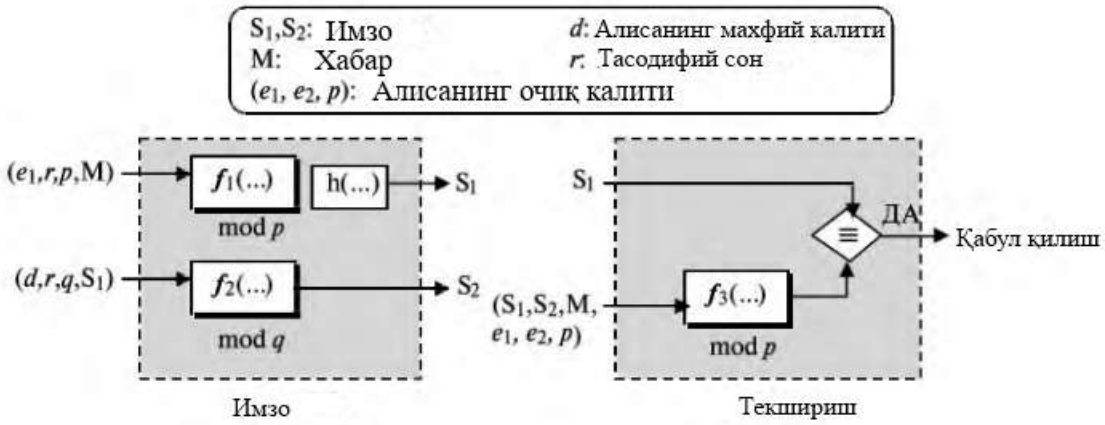
Рақамли имзонинг самарадорлиги унинг асосини ташкил этувчи криптографик алгоритмнинг тезкорлиги, бардошлилиги ва ҳисоблаш машиналарига мослашувчанлигига боғлиқ.

ЭРИни текшириш жараёни

Ушбу тезисда электрон ҳужжатларни алмашиниш жараёнларидаги ахборот хавфсизлигини таъминлаш муаммосини ҳал этишда муҳим ўрин эгаллаган криптографик восита Шнор схемасига асосланган электрон рақамли имзо алгоритмининг моҳияти баён этилади.

Эл-Гамалнинг рақамли имзо схемаси билан боғлиқ муаммо шундаки, Zp дискрет логарифмини қийинлаштириши учун p жуда катта бўлиши керак. P-узунлиги камида 1024 бит бўлиши тавсия этилади. Сиз 2048 бит ҳажмида

имзо қўйишингиз мумкин. Имзо ҳажмини камайтириш учун Шнорр Эл-Гамал схемасига асосланган, аммо қисқартирилган имзо ҳажмига эга янги схемани таклиф қилди. Қуйидаги расмда Шнорнинг рақамли имзо схемаси ҳақида умумий тушунча берилган.[1]



Шнорр рақамли имзо схемасининг умумий ғояси

Имзоланиш жараёнида иккита функция иккита имзо ҳосил қилади; текшириш жараёнида битта функциянинг чиқиши текшириш учун биринчи имзо билан таққосланади. Юқоридаги расмдаҳар бир функциянинг киришлари кўрсатилган. Энг муҳими шундаки, паллада p ва q иккита модул ишлатилади.1 ва 3 функциялари p қўлланилади; 2 фунцсияси q ни қўллайди.

Калитларни ҳосил қилиш.

Хабарга имзо қўйишдан олдин, Элис калитларни яратиши ва ҳамма учун очиқ калитларни эълон қилиши керак.

1. Алиса узунлиги 1024 битга тенг бўлган p туб сонни танлайди.
2. Алиса криптографик хеш функцияси томонидан яратилган дайжест ҳажмига тенг келадиган яна бир қ ни танлайди. q туб сони $p-1$га бўлиниш керак. Яъни $(p-1)=0 mod q$ шарт қаноатланиши керак.
3. Алиса яна e_1сонни танлайди. Бунинг учун Алиса Zp тўпламга тегишли e_0примитиве элементни танлайди ва $e_1 = e_1^{(p-1)/q}$ модр ни ҳисоблайди.
4. Алиса махфий калити сифатида d бутун сонни танлайди.
5. Алиса $e_2 = e_2^d$ модр ни ҳисоблайди.
6. Алисанинг очиқ калити- (e_1, e_2, p, q), унинг махфий калити-(d).

Имзолаш жараёни

1. Алиса тасодифий r рақамини танлайди. Эътибор беринг, очиқ ва ёпиқ калитлардан кўплаб хабарларни имзолаш учун фойдаланиш мумкин. Аммо Алиса ҳар сафар янги хабар юборганида r-ни ўзгартириши керак. Шуни ҳам унутмангки, r қиймати 1 ва q орасида бўлиши керак.
2. Алиса биринчи имзони ҳисоблайди $S_1 = h(M \mid e_1^r \bmod p)$. Хабар M олдин $e_1^r \bmod p$ қийматига қўшилади, сўнгра дайжест ҳосил қилиш учун ҳаш функцияси қўлланилади. Шуни эсда тутиш лозимки, хеш функцияси хабарга бевосита тааллуқли эмас, лекин у М ва $e_1^r \bmod p$ қийматларнинг йиғиндисидан олинади.

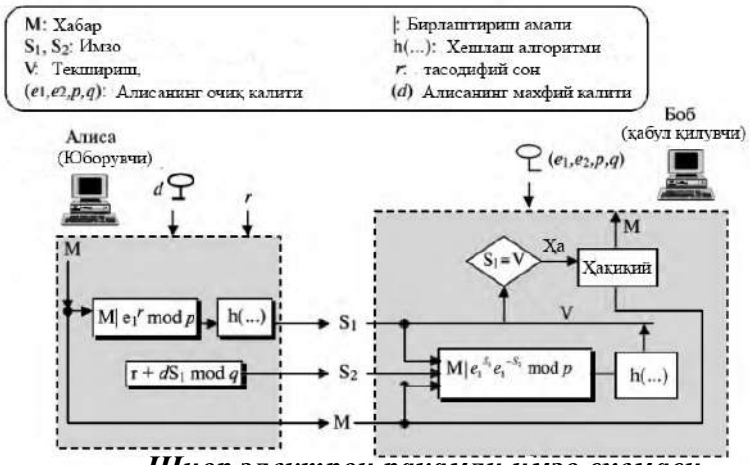
3. Алиса иккинчи имзони ҳисоблвйди $S_2 = r + d\,x\,S_1\ mod\ q.$
4. Алиса $M., S_1$ ва S_2 ни узатади.

Имзони текшириш жараёни.

Қабул қилувчи, Мисол учун Боб $M., S_1$ ва S_2 маълумотларни қабул қилади.

1. Боб $V = h\,(M \mid e_1^{S_1} e_2^{-S_1}\ mod\ p)$ ни ҳисоблайди.
2. Агар S_2 модул p бўйича V га мос бўлса, хабар қабул қилинади; акс ҳолда рад этилади.

Имзо қўйиш ва имзони текшириш жараёнларининг схемаси қуйидаги расмда ифодаланган.[1]



Шнор электрон рақамли имзо схемаси

Шнор схемасига асосланган электрон рақамли имзо алгоритмлари қуйидаги афзалликларга эга.

- Хабарда имзо яратиш учун зарур бўлган ҳисоб-китобларнинг боғлиқлигини минималлаштиради.

- Фаол рақибга қаршилик кўрсатиш учун аутентификация протоколи нол очилган далил бўлиши этарли. Ҳалигача ким Шнорр схемаси учун нол очиш хусусиятини исботлай олмаган.

- Бугунги кунда Шнор схемаларига асосланган алгоритмлар криптовалюталарнинг самарадорлигини оширишда фойдаланилмоқда[4]

- Чизиқлилик: Шнорр имзолари диққатга сазовор хусусиятга эга: бир нечта томонлар биргаликда очиқ калитлари йиғиндиси учун яроқли имзо яратишлари мумкин. Бу мултиимзолар ва бошқа ақлли шартномалар каби самарадорлик ва махфийликни оширадиган турли хил юқори даражадаги дизайнлар учун қурилиш блоки бўлиб хизмат қилиши мумкин.

Агар жорий этилаётган криптовалюталар тизимида Шнор схемасига асаланган рақамли кўп имзоли алгоритмларидан фойдаланилса, самарали натижаларга эришилади.

Адабиётлар

1. *Криптография и безопасность сетей: Учебное пособие / Фороузан Б.А.; пер. с англ. под ред. А.Н. Берлина. — М.: Интернет-Университет Информационных Технологий: БИНОМ. Лаборатория знаний, 2010. — 784 с.: ил., табл. — (Основы информационных технологий)*

2. *Акбаров Д., Хасанов П., Хасанов Х., Ахмедова О. Криптографиянинг математик асослари. Ўқув қўлланма.– Тошкент, 2010 – 210 б.*

Акбаров Д.Е. "Ахборот хавфсизлигини таъминлашнинг криптографик усуллари ва уларнинг қўлланилиши". Тошкент 2009.

БИОМЕДИЦИНА СИГНАЛЛАРИГА РАҚАМЛИ ИШЛОВ БЕРИШДА ТРЕНДЛАРНИ ОЛИБ ТАШЛАШ УСУЛЛАРИ

Зайнидинов Х.Н., Махмуджанов С.У., Муҳаммад ал-Хоразмий номидаги Тошкент ахборот технологиялар университети, Ўзбекистон.

***Annotation**: In this study, the results of studies in the field of spectral processing of relatively low-frequency biomedical signals (in the range of 0.01 - 0.25 Hz) to determine their local characteristics in the time and frequency domains are considered. Often, phenomena such as trends (long-term tendencies of the signal being studied) manifest themselves when viewing the gastrogram and enterograms, which is especially evident in sick patients.*

***Key words:** biomedical signals, trend, electrocardiography (ECG), electroencephalography (EEG), Fure, gastrogram*

***Аннотация**: В данной работе рассматриваются результаты исследований в области спектральной обработки относительно низкочастотных биомедицинских сигналов (в диапазоне 0,01 - 0,25 Гц) для определения их локальных свойств во временной и частотной областях. Часто такие явления, как тренды (долговременные тенденции изучаемого сигнала), проявляются при просмотре гастрограммы и энтерограмм, что особенно заметно у больных пациентов.*

***Ключевые слова:** биомедицинские сигналы, тренд, электрокардиография (ЭКГ), электроэнцефалография (ЭЭГ), Фуре, гастрограмма*

***Аннотация**: Ушбу тадқиқот ишида нисбатан паст частотали биотиббий сигналларни спектрал ишлов бериш (0,01 - 0,25 Hz диапазонида) вақт ва частота соҳаларида уларнинг маҳаллий хусусиятларини аниқлаш йўналишидаги тадқиқотлар натижалари кўриб чиқилган. Кўпинча гастрограм ва энтерограмларни кўришда трендлар (ўрганилаётган сигналнинг узоқ муддатли тенденсиялари) каби ҳодисалар ўзини намоён қилади, бу айниқса касал беморларда яққол намоён бўлади.*

***Калит сўзлар:** биомедицина сигналлари, тренд, электрокардиография (ECG), электроэнсефалография (EEG), Фуре, гастрограмма*

Биомедицина сигналларни ўрганиш учун замонавий компьютер усуллари ва воситаларини тадқиқ қилиш, айниқса электрокардиография (ECG) ва электроэнсефалография (EEG) соҳаларида кенг ривожланмоқда. Тиббиёт ва ахборот технологияларида фундаментал ва амалий характерга эга муҳим муаммолар мавжуд, уларни ҳал қилмасдан тиббиёт соҳасига кириб бориш ва билимларни эгаллаш имконсиз ҳисобланади.

Сўнгги йилларда турли мамлакатларда, шу жумладан Ўзбекистонда ҳам турли хилдаги биомедицина сигналларни қайта ишлаш ва тадқиқ қилиш учун янги ихтисослашган ҳамда компьютерлаштирилган усул ва воситалар жадал ривожланмоқда. Шу жумладан маьлумотни таҳлил қилишнинг янги усуллари

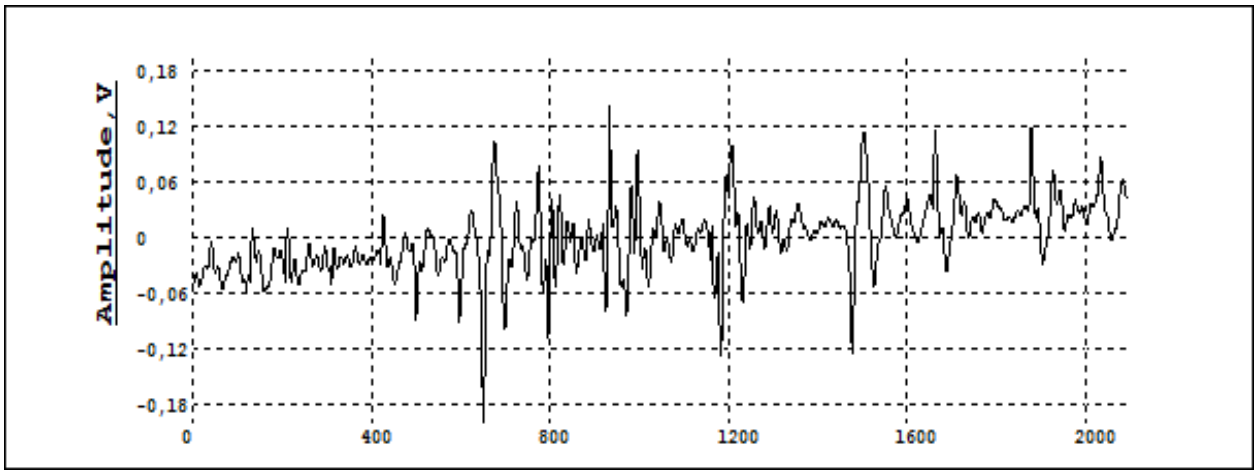
ишлаб чиқилмоқда ва клиник амалиётга кенг жорий этилмоқда. Тиббиётда компьютерлар узоқ вақтлар давомида қўлланилиб келмоқда ва кўплаб замонавий диагностика усуллари ахборот технологияларига (АТ) асосланган. Ультратовуш ёки компьютер томографияси каби текширув усуллари ҳам компьютер технологияларисиз умуман аҳамиятга эга эмас. Ҳозирда компьютерлардан фаол фойдаланилмайдиган тиббиёт соҳасини топиш деярли имконсиз. Аммо фақатгина тиббиётда компьютерлардан фойдаланиш диагностика билан чекланмайди. Улар тобора кўпроқ турли хил касалликларни даволашда ҳам қўлланилмоқда ва нафақат даволашда шу билан бирга, мақбул даволаш режасини тайёрлашдан тортиб, процедуралар ўтказиш давомида турли хил тиббий жиҳозларни бошқаришгача.

Тиббиётдаги ахборот тизимларининг ҳозирги ривожланиш босқичида компьютер технологияларига асосланган электрокардиография (ECG) ва электроансефалографияларда (EEG) энг ривожланган усуллари ва воситалари ўрин олган. Шу билан бирга, сўнгги йилларда дунёнинг кўплаб мамлакатларида тиббиётнинг гастроентерология соҳаси учун компьютер диагностик тизимини яратишда катта ютуқларга эришилмоқда. Қорин бўшлиғи органларининг ҳолатини текширишнинг инвазив усуллари, яъни беморларда оғриқли ҳисларни келтириб чиқарадиган (масалан, балонография, гастроскопия, pH ўлчагич) танага сенсорларни киритиш билан боғлиқ бўлган усуллар инвазив бўлмаган усулларга йўл очмоқда. Сўнггилари кўпинча электрогастрографиянинг (EGG) ютуқларига асосланган, яъни тананинг сиртқи қатламидан олинган биопотенсиалларни қайд этиш ва динамикасини таҳлил қилиш ҳамда ошқозон-ичак тракти (ОИТ) мушакларининг фаолияти тўғрисида маълумот бериш усулларига асосланади. EGG ҳамда тегишли жиҳозларнинг хусусиятларини яхшилаш учун математик ва дастурий ахборот тизимларини янада такомиллаштириш зарур. EGG соҳаси билан боғлиқ бўлган методлар ошқозон-ичак тракти касалликларини даволашда терапевтик ҳамда жарроҳлик усуллар ёрдамида қўлланилади. Уларнинг афзаллиги шундаки, инсон органларининг ҳолатини миқдорий жиҳатдан баҳолаш имкониятини беради. Бу, айниқса, аҳоли касалликларининг прафилактик олдини олиш, шунингдек ошқозон-ичак трактининг аниқланган бузилишларини ўз вақтида даволаш нуқтаи назаридан жуда муҳимдир. Ошқозон-ичак трактидаги жараёнлар ностационарлиги ва бошқа органларга нисбатан паст частотали тебраниш характери билан фарқланади.

Гастро ва энтерограмм вақт-частота таҳлилида кенг қўлланилган дискрет қисқавақтли Фуре ўзгартиришларига (ҚФЎ) асосланган гастро сигналининг маҳаллий стационар математик модели ҳисобланади. Бу, хусусан, маълум бир вақт ичида қувватнинг спектрал зичлиги ва сигналларнинг бошқа хусусиятларининг ўзгаришини ўрганишга имкон беради.

Ушбу тадқиқот ишида нисбатан паст частотали биотиббий сигналларни спектрал ишлов бериш (0,01 - 0,25 Hz диапазонида) вақт ва частота

соҳаларида уларнинг маҳаллий хусусиятларини аниқлаш йўналишидаги тадқиқотлар натижалари кўриб чиқилган. Кўпинча гастрограм ва энтерограмларни кўришда трендлар (ўрганилаётган сигналнинг узоқ муддатли тенденсиялари) каби ҳодисалар ўзини намоён қилади, бу айниқса касал беморларда яққол намоён бўлади. Трендни ўз ичига олган электрогастрограмма 1 - расмда келтирилган.



1-расм Трендни ўз ичига олган гастрограмма

Чизиқли тенденцияни ўз ичига олган $x^*(t)$ сигнални шаклнинг математик модели билан тавсифлаш мумкин:

$$x_j^*(t) = x_j(t) + \gamma t \qquad (1)$$

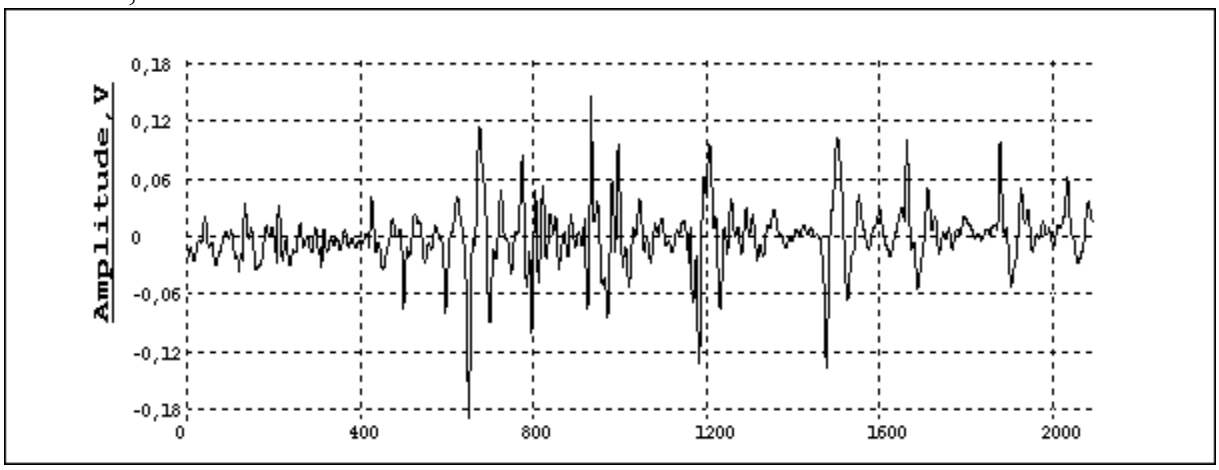
бу ерда $x_j(t)$ - сигналнинг j-инчи намунасининг амплитудаси, у1 - (0; 0) ва $(t; x_j(t))$ орқали ўтадиган тўғри чизиқнинг қиялигининг тангенси. Юқори даражадаги трендлар, шунингдек, алоҳида ёзув сегментларида чизиқли моделлар тўплами билан тавсифланиши мумкин.

Трендларни бартараф қилиш учун жуда самарали, эмпирик усул қўлланилади. i-сегментда сигналнинг ўртача қиймати ва гастроентерограмманинг (i + k)-сегментидаги ўртача қиймат ҳисобланади. Шундан сўнг, бошланғич қисмдаги ўртача қийматга тенг бўлган аралашиш бартараф этилади кейин эса ёзувнинг эгилиш бурчаги аниқланади. Гастроентерограмма олдиндан белгиланган бурчак остида айланади. Гастроентерограмманинг $[i_0; i]$ сегментидаги трендини бартараф этиш учун дастур моделини қуйидагича ёзиш мумкин:

$$\gamma^{(i)} = \frac{\sum_j \left(x^{*(i)}_j - x^{*(i+k)}_{j+k}\right)}{(i - i_0)^2}; j = \overline{i_0, i} \qquad (2)$$

$$x_j^{(i)} = x^{*(i)}_j - j \cdot \gamma^{(i)}$$

2-расмда (2) га мувофиқ олиб ташланган тренд, гастрограмма тақдим этилган;



2-расм Тренд олиб ташланган гастрограмма

Ҳар қандай биомедицина сигналга олдиндан қайта ишлов берилади. Гастроентерологик сигналларни олдиндан қайта ишлаш деганда умумий бошланғич сигналларни рақамли филтрлаш (одатда 0,01 - 0,25 Hz частота диапазонида) ОИТ аналог сигнал каналларини динамик ажратиш мақсадида, шунингдек трендларни олиб ташлаш жараёни тушунилади.

PHP WEB DASTURLASH TILIDA MATNLI FAYLLAR BILAN ISHLASH IMKONIYATLARI

Jabborova M., Alimqulov N. Andijon davlat universiteti, O'zbekiston.

Annotation*: This article mainly discusses the possibilities of working with text files in PHP web programming language.*

Keywords*: PHP, web, file writing, file reading, text files.*

Аннотация*:В этой статье в основном рассматриваются возможности работы с текстовыми файлами на языке веб-программирования PHP.*

Ключевые слова*: PHP, веб, запись файлов, чтение файлов, текстовые файлы.*

Annotatsiya*:Ushbu maqolada asosan PHP veb dasturlash tilida matnli fayllar bilan ishlash imkoniyatlari ko'rib chiqiladi.*

Kalit so'zlar*:PHP, web, faylga yozish, fayldan o'qish, matnli fayllar.*

Fayllar bilan ishlash, ya'ni fayllarni yaratish, tekshirish, o'qish yoki yozish – bu ixtiyoriy dasturlash tilining asosiy instrumenti hisoblanadi. RNR tili ham bu imkoniyatlardan holi emas. Bu tilda yuqorida keltirib o'tilgan imkoniyatlar va boshqa qo'shimcha imkoniyatlarga ega bo'lgan ko'plab funksiyalar mavjud. Bularga bir qancha misollar ko'rib chiqamiz. Bundan tashqari ruxsat tiplari xam mavjud. Ruxsat tipi parametri quyidagi qiymatlardan birontasini qabul qilishi mumkin.

1-jadval. *Faylning ruxsat tiplari.*

Ruxsat tipi	Nomlanishi
r	O'qish rejimi — faylni faqat o'qish uchun ochadi; ko'rsatkich faylga uning boshidan boshlab o'rnatiladi.
r+	O'qish rejimi — faylni faqat o'qish va yozish uchun ochadi; ko'rsatkich faylga uning boshidan boshlab o'rnatiladi.
w	Yozish rejimi — fayl faqat yozish uchun ochiladi; ko'rsatkich faylni boshidan boshlab o'rnatiladi. Agar fayl mavjud bo'dsa, fayl nol uzunlikkacha qisqartiriladi, ya'ni tarkibidagilarning barchasi o'chiriladi. Agar fayl mavjud bo'lmasa, uni ochish vaqtida faylni yaratadi.
w+	Yozish rejimi — fayl o'qish va yozish uchun ochiladi; ko'rsatkich faylni boshidan boshlab o'rnatiladi. Agar fayl mavjud bo'dsa, tarkibidagilarning barchasi o'chiriladi. Agar fayl mavjud bo'lmasa, uni ochishga xarakat qiladi va natijada faylni yaratadi.
a	Qo'shish rejimi — fayl faqat qo'shish (yozish) uchun ochiladi; ko'rsatkich fayl tarkibining oxiriga o'rnatiladi. Agar fayl mavjud bo'dmasa, uni ochishga xarakat qiladi va natijada faylni yaratadi.
a+	Qo'shish rejimi — fayl faqat qo'shish (yozish) va o'qish uchun ochiladi; ko'rsatkich fayl tarkibining oxiriga o'rnatiladi. Agar fayl mavjud bo'dmasa, uni ochishga xarakat qiladi va natijada faylni yaratadi.

x	Faqatgina yozish uchun fayl yaratadi va ochadi; ko‘rsatkichni faylning boshiga o‘rnatadi. Agar fayl mavjud bo‘dsa, u xolda fopen() false ni chiqaradi va ogoxlantirish beradi. Agar fayl mavjud bo‘lmasa, u xolda uni yaratishga urinib ko‘radi. Ruxsatning bu tipi PHP 4.3.2 versiyasidan boshlab qo‘llaniladi va faqatgina lokal fayllar bilan ishlaydi.
x+	Faqatgina yozish va o‘qish uchun fayl yaratadi va ochadi; ko‘rsatkichni faylning boshiga o‘rnatadi. Agar fayl mavjud bo‘dsa, u xolda fopen() false ni chiqaradi va ogoxlantirish beradi. Agar fayl mavjud bo‘lmasa, u xolda uni yaratishga urinib ko‘radi. Ruxsatning bu tipi PHP 4.3.2 versiyasidan boshlab qo‘llaniladi va faqatgina lokal fayllar bilan ishlaydi.

Faylga yozish uchun w (write) ya’ni yozmoq tipidan foydalanamiz. Bu funksiyani ishlatish xavfli xisoblanadi. Sababi zaminit qilib yozadi ya’ni fayl ichidagi barcha ma’lumotlarni foydalauvchidan so‘ramasdan o‘chirib yuboradi. Eski ma’lumot ustiga yangisini yozadi. Yozish uchun faylni quyidagicha ochamiz.

```
<?php
$f=fopen("ruyhat.txt","w");
$x=rand(5,100);
fputs($f,$x);
fputs($f,"\n");
fclose($f);
?>
```

Fayl oxiriga qo‘shish uchun a (append) ya’ni qo’shmoq tipidan foydalanamiz. Bu funksiyani ishlatish xam xafvli xisoblanadi. Sababi u foydalanuvchidan so‘ramasdan ma’lumotlarni faylni oxiriga o‘zi qo‘shib ketadi.

```
<?php
$f=fopen("ruyhat.txt","a");
$x=rand(5,100);
fputs($f,$x);
fputs($f,"\n");
fclose($f);
?>
```

Faylni o‘qish uchun r (read) ya’ni o’qimoq tipidan foydalanamiz. Bu tip xam ancha xavfli xisoblanadi. Sababi faylda qancha ma’lumot borligini oldindan bilish qiyin. $x=fgets($f); echo $x; echo "<br>"; ushbu satlar bir marta yozilsa, fayldagi ma’lumotlarning birinchi qatori o‘qiladi. Ikki marta yozilsa, ikkinchi qatori o‘qiladi.

```
<?php
$f=fopen("ruyhat.txt","r");
$x=fgets($f);
echo $x;
echo "<br>";
fclose($f);
?>
```

Bu ko‘rinishda esa fayl ichidagi ma’lumotlar oxirigacha o‘qiladi. Ekranga barcha ma’lumotlar chop etiladi.

```
<?php
$f=fopen("ruyhat.txt","r");
while (!feof($f)){
$x=fgets($f);
echo $x;
echo "<br>";}
fclose($f);
?>
```

Adabiyotlar.

1. Aripov M. A-75 WEB texnologiyalari: o’quv qo’llanma / M.Aripov, S.Dottoyev, M.Fayziyeva; O’zbekiston Respublikasi Oliy va o’rta maxsus ta’lim vazirligi. – Toshkent: O’zbekiston faylasuflari milliy jamiyati nashriyoti, 2013. – 280 b.

2. O.Jakbarov, T.Jo’rayev. INTERNETDA DASTURLASH muammoli ma’ruzalar matni. Namangan – 2004.

AYRIM OPTIK JARAYONLARNI O‘RGANISHDA DASTURIY MODELLASHTIRISHLARDAN FOYDALANISH TAJRIBASIDAN

A.Artikov, J.G'ulomov. Andijon davlat universiteti, O'zbekiston.

Annotation*: This article covers methods of teaching some optical effects (phenomena) through computer modeling, action programs. To explain the optical effects below, an initial version of the software package called "optical effects" was developed using the Visual Basic programming language, which described some of its aspects.*

Keywords*: computer, modeling, programs, optical effects, complex, Visual Basic, Crocodile, MatLab, physical processes, Kerr effect*

Аннотация*: В данной статье освещены методы обучения некоторым оптическим эффектам(явлениям) посредством компьютерного моделирования, программы действий. Для объяснения оптических эффектов ниже был разработан первоначальный вариант программного комплекса под названием "оптические эффекты" с использованием языка программирования Visual Basic, в котором были описаны некоторые его аспекты.*

Ключевые слова: *компьютер, моделирование, программы, оптические эффекты, комплексные, Visual Basic, Крокодил, MatLab, физические процессы, эффект Керра*

Annotatsiya*: Mazkur maqolada ayrim optik effektlar(hodisalar)ni kompyuterda modellashtirish orqali o‘rgatish usullari, harakatdagi dasturlar asosida yoritilgan. Optik effektlarni tushuntirish uchun quyida Visual Basic dasturlash tili yordamida "Optik effektlar" nomli dasturiy kompleksning dastlabki varianti ishlab chiqildi va uning ayrim jihatlari bayon qilindi.*

Tayanch so'zlar: *kompyuter, modellashtirish, dasturlar, optik effektlar, kompleks, Visual Basic, Krokodil, MatLab, fizik jarayonlar, Kerr effekti*

Keyingi yillarda kompyuterlardan barcha sohalardagi kabi ta'lim sohasida ham keng qamrovli foydalanib kelinmoqda. O'quvchi va talabalarni o'z sohalari bo'yicha chuqur bilimga ega bo'lishi uchun turli animatsiyalar, dasturlar va multimediyalar samarali vositalar bo‘lib xizmat qilib kelmoqda[1-5]. Fizik jarayonlarni modellashtirish, laboratoriya ishlarini kompyuterlashtirish va standart dasturlardan, Krokodil, MatLab kabi dasturlashning eng zamonaviy usullaridan foydalanish esa murakkab fizik jarayonlarni har tamonlama,chuqur va mukammal o‘rganish imkoniyatlarini yaratmoqda. Jumladan o‘zining murakkabligi bilan ajralib turadigan optik effektlarni tushuntirish uchun quyida Visual Basic dasturlash tili yordamida "Optik effektlar" nomli dasturiy kompleksning dastlabki varianti ishlab chiqildi va uning ayrim jihatlari bayon qilindi. Bu dastur o'z ichiga bir qancha effektlarni oladi. Dastur ishga tushurilishi bilan, monitorning o‘ng

tamonidagi darchada o‘rganilayotgan jarayon(effekt)larning qisqa nazariyasi bayon qilinadi. Chapda esa, jarayonning multimedia modeli ishga tushadi.

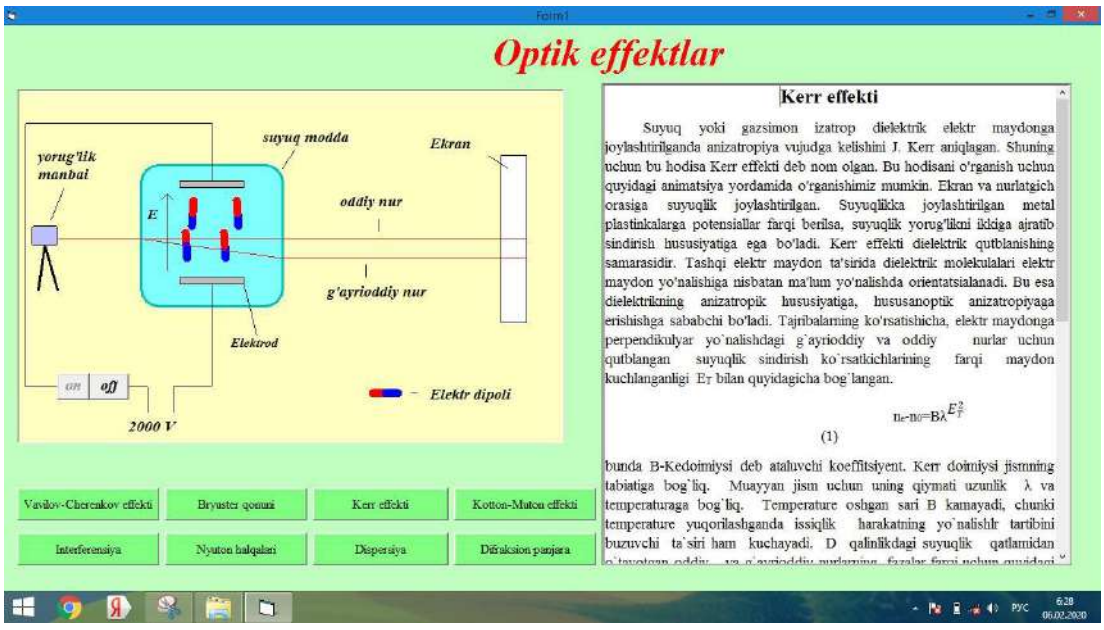
1.KERR EFFEKTI

Suyuq yoki gazsimon izatrop dielektrik elektr maydonga joylashtirilganda anizatropiya vujudga kelishini J. Kerr aniqlagan. Shuning uchun bu hodisa Kerr effekti deb nom olgan. Bu hodisani quyidagi animatsiya yordamida o’rganishimiz mumkin. Ekran va nurlatgich orasiga suyuqlik(nitrobenzol,uning eritmalari kabi) joylashtirilgan. Suyuqlikka joylashtirilgan yassi metal plastinkalarga potensiallar farqi berilsa, suyuqlik yorug’likni ikkiga ajratib sindirish hususiyatiga ega bo’ladi. Kerr effekti aslida dielektrik qutblanishining samarasidir. Tashqi elektr maydon ta’sirida dielektrik molekulalari elektr maydon yo’nalishiga nisbatan ma’lum yo’nalishda orientatsialanadi.Bu jarayonni dastlab Klauzius va Massoti o’rganganlar. Bu esa dielektrikning anizatropik hususiyatiga, hususan optik anizatropiyaga erishishga sababchi bo’ladi. Tajribalarning ko’rsatishicha, elektr maydonga perpendikulyar yo`nalishdagi g`ayrioddiy va oddiy nurlar uchun, qutblangan suyuqlik sindirish ko`rsatkichlarining farqi maydon kuchlanganligi $E_T$ bilan quyidagicha bog`langan.

$$n_e-n_0=B\lambda E_T^2 \qquad (1)$$

bunda B—Kerr doimiysi deb ataluvchi koeffitsiyent. Kerr doimiysi jismning tabiatiga bog`liq. Muayyan jism uchun uning qiymati uzunlik $\lambda$ va temperaturaga bog`liq. Temperatura oshgan sari B kamayadi, chunki temperatura yuqorilashganda issiqlik harakatning yo`nalishlar tartibini buzuvchi ta`siri ham kuchayadi. Korrelyatsiya faktori deb ataluvchi juda muhim kattalik ham shu qonuniyatga bo‘ynsunadi.Bu kattalikni Klauzius-Masoti tenglamasidan keltirib chiqariladi[6]. Qandaydir d qalinlikdagi suyuqlik qatlamidan o`tayotgan oddiy va g`ayrioddiy nurlarning fazalar farqi uchun quyidagi ifodani yozish mumkin

$$\Delta\varphi = \frac{2\pi}{\lambda} d(n_e - n_0) = \frac{2\pi}{\lambda} d\ B\lambda E_T^2 = 2\pi d\ BE_T^2 \qquad (2)$$



1-rasm: Dastur tarkibidagi Kerr effekti

Adabiyotlar

1. *Begimqulov U. «Fizika darslarida informatsion texnologiyalardan foydalanish»,// «Fizika, matematika va informatika jurnali». 2002*

2. *Yo'ldoshev B. J., Ro'zimov Sh.M. «Multimedia tizimidan foydalangan holda fizik hodisa va jarayonlarni modellashtirish. // Fizika, matematika va informatika». 2002 y. №4(6), 54–s.*

3. *Соучек Б. Мини ЭВМ в системах обработки информации. М.1976,518 с*

4. *Сальтиков А.И, Семашко Г.Л Программирование для всех. Москва «Наука» 1986г. 176 стр.*

5. *Бухтияров А.М, Маликов Ю.П, Фролов Г.Д Практикум по программированию на фортране. М. 1979*

Артиков А. Корреляция фактори ва суюқликларда содир бўлувчи жараёнларни ўрганиш истиқболлари// АДУ илмий хабарномаси. №4. 2016.12-14 с

CHIZIQLI DASTURLASH MASALASINI AMALIY DASTURLAR PAKETI YORDAMIDA YECHISH

Ismanova K.D., Isomadinov U.M., Jabbarov A.M.
Namangan muhandislik-texnologiya instituti, O'zbekiston

Abstract*. This article solves linear programming problems that represent mathematical models of economic processes. Computer technologies, in a package of practical mathematical programs, were used to search for objective functions and their optimal solutions and to analyze the results.*

Key words. *Linear programming, mathematical modeling, optimization, objective function, boundary conditions, mathematical software package, maximizing and minimizing problems.*

Аннотация. *В данной статье peshаются задачи линейного программирования, представляющие математические модели экономических процессов. Компьютерные технологии, в частности пакет практических математических программ, использовались для поиска целевых функций и их оптимальных peshений и анализа результатов.*

Ключевые слова. *Линейное программирование, математическое моделирование, оптимизация, целевая функция, граничные условия, пакет математических программ, задачи максимизации и минимизации.*

Annotatsiya. *Ushbu maqolada iqtisodiy jarayonlarning matematik modellarini ifodalovchi chiziqli dasturlash masalasi yechildi. Bunda maqsad funksiyalari va ularni maqbul yechimlarini topishda kompg'yuter texnologiyalaridan xususan, amaliy matematik dasturlar paketidan foydalanildi va natijalar tahlil etildi.*

Kalit so'zlar. *Chiziqli dasturlash, matematik model, optimallashtirish, maqsad funksiyasi, chegaraviy shartlar, matematik dasturlar paketi, maksimallashtirish va minimallashtirish masalalari.*

Tadbirkorlik faoliyati bilan shug'ullanuvchi korxona va tashkilotlarda iqtisodiy jihatdan turli xulosalar chiqarish ularning to'g'ri va asosli ekanligini isbotlab bera olishda miqdoriy usullarni va zamonaviy statistik kompyuter dasturlarini qo'llash muhim ahamiyat kasb etadi. Noaniqlik sharoitida optimal qaror qabul qilishda iqtisodiy usullarning ahamiyati tobora ortib bormoqda.

Faraz qilaylik, optimal dasturlash masalasining chegaraviy shartlari chiziqli tenglamalar va tengsizliklar sistemasidan iborat bo'lsin, yahni optimal dasturlash masalasi quyidagi ko'rinishda berilsin.

Faraz qilaylik, n xil mahsulot ishlab chiqarish uchun m xil xom ashyodan foydalanilsin.

Quyidagi belgilashlarni kiritaylik:

b_i- i xil xom ashyoning mavjud miqdori;

A_j (j=1..n)-ishlab chiqariladigan mahsulot turi;

a_{ij}- j-xil mahsulot birligiga sarflanadigan i xil xom ashyo miqdori;

c_j- j xil mahsulot birligidan olinadigan daromad;

Ishlab chiqarilishi kerak bo'lgan mahsulot miqdorlarini x_j bilan belgilasak masalaning matematik modeli quyidagicha bo'ladi:

$$x_j \geq 0,\ j = \overline{1,n} \qquad b_i \geq 0,\ i = \overline{1,m}$$

j

$$\sum_{i=1}^{m} a_{ij} x_j \leq A_j, \qquad F_{max} = \sum_{i=1}^{m} C_i x_i \qquad (1)$$

Mazkur chiziqli dasturlash masalasini yechishning bir necha xil usullari mavjud. Biroq zamonaviy kompyuter texnologiyalari yordamida bu masala juda oson hal etiladi.

Buning uchun chiziqli dasturlash masalasiga keluvchi quyidagi masalani qaraymiz.

Fabrika ikki xil A va B tikuv mahsuloti ishlab chiqaradi. Bu mahsulotlarni ishlab chiqarishda uch xil N1, N2, N3 turdagi materiallarni ishlatadi. Omborda 1-materialdan 15 m, 2-materialdan 16 m, 3-materialdan 18 m zahira mavjud. A mahsulotni ishlab chiqarish uchun N1-dan 2m, N2-dan 1m, Nz-dan 3m ishlatadi. B- mahsulotni ishlab chiqarish uchun N1-dan 3m, N2-dan 4m, N3-dan 0m ishlatadi. A- mahsulotning bir birligidan keladigan foyda 10 ming sumni, B - mahsulotdan keladigan foyda 5 ming so'mni tashkil qiladi.

Ishlab chiqarishning shunday rejasini tuzish kerakki, fabrika maksimal foyda olsin.

Yaqin kungacha foydalanuvchi o'zining matematik masalasini yechish uchun nafaqat matematikani bilishi balki kompg'yuterda ishlashni, kamida bitta dasturlash tilini bilishi va murakkab hisoblash usullarini o'zlashtirgan bo'lishi kerak edi. Hozirda esa dasturlashni bila olmaydiganlar uchun tayyor ilmiy dasturlar majmualari, elektron qo'llanmalar va tipik hisob- kitoblarni bajarishga mo'ljallangan dasturiy vositalar bo'lgan - *amaliy dasturlar paketlari* (ADP) mavjud.

Mashina matematikasini ADP yordamida o'rganish foydalanuvchida matematikaning o'zini o'rganish illyuziyasini yaratadi. Ammo shuni aytish joizki, mazkur paketlarda yaratilgan har qanday chiroyli menyu foydalanuvchini oddiy matematik tushunchalardan va usullardan ozod qila olmaydi. Xususan, agar foydalanuvchi matritsa nimaligini bilmasa, u holda matritsa algebrasi dasturiy paketi unga hech qanday yordam bera olmaydi, yoki foydalanuvchi noaniq bo'lmagan integralni sonli usullar yordamida hisoblashga uringanda, u haqiqatdan ancha yiroq bo'lgan javobni olishi yoki javobni umuman ololmasligi ham mumkin. Ixtiyoriy keng imkoniyatlarga ega paket universal yondashishga bogliq. Matematik paketlarni ishlatishda mutaxassis undan ongli foydalanib, chegirmalar qilishi mumkin: paketni uning muammosiga rostlashi, dasturni modifikatsiyalash, yangilash, hisoblash vaqtini tejash va h.k.

Shu bois yuqoridagi chiziqli dasturlash masalasini Mathcad yordamida yechish jarayonini qaraymiz.

Mathcadda chiziqli dasturlash masalasi yechishda maximize va minimize funksiyalaridan foydalanish mumkin. Bu funksiyalar umumiy holda quyidagi

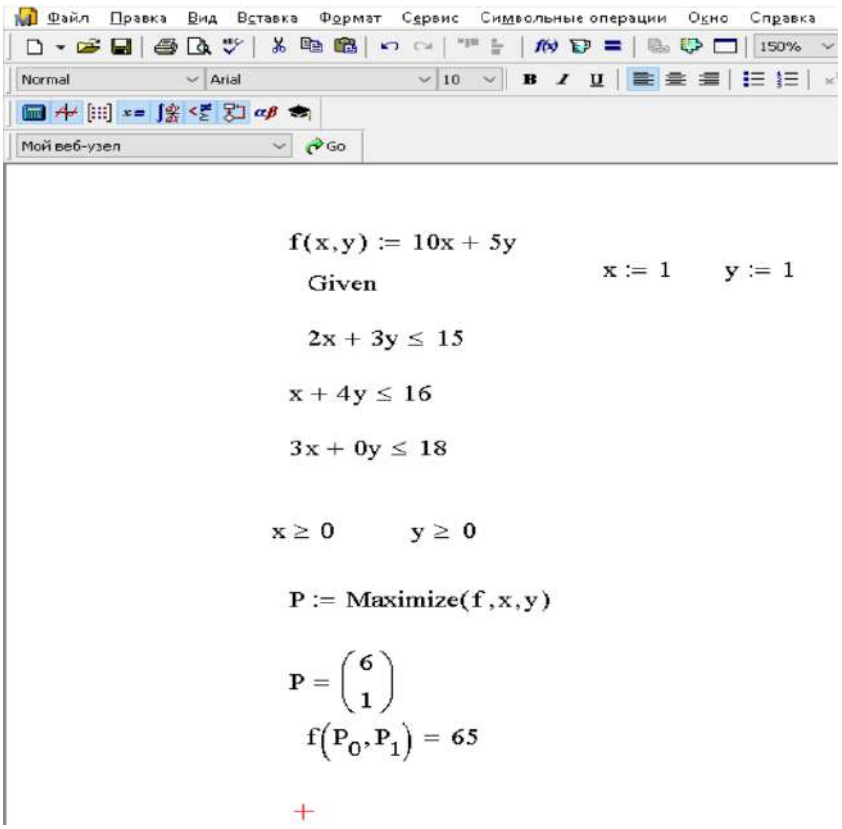
ko’rinishda yoziladi:

Maximize (F,<o’zgaruvchilar ro’yxati>)

Minimize (F,<o’zgaruvchilar ro’yxati>)

Mathcadda chiziqli dasturlash masalasini yechish quyidagi bosqichlarda bajariladi (1-rasm):

1. Mathcadni ishga tushirgandan so’ng, maqsad funksiyasi yoziladi, f(x,u)=<funksiya ko’rinishi> va o’zgaruvchilarning boshlang’ich qiymati kiritiladi.



MathCadda yechish jarayoni.

2. Given kalit so’zi yoziladi.
3. Tengsizliklar tizimi va cheklanishlar kiritiladi.
4. Biror o’zgaruvchiga maximize yoki minimize funksiyasi yuboriladi.
5. Shu o’zgaruvchi yozilib tenglik kiritiladi. Natija vektor ko’rinishida hosil bo’ladi.
6. Maqsad funksiyasi qiymatini hisoblash uchun, masalan $f(P_0,P_1)$ yozilib tenglik belgisi kiritiladi.

Xulosa qilib aytganda, bugungi kunda iqtisodiy masalalarni yechishda turli xil zamonaviy dasturlash texnologiyalari |qo’llaniladi. Ularning barchasi turli xil usulda talab etilgan natijalarga erishishga imkon berishi mumkin. Biroq mutaxassis qo’yilgan masalaning mazmuniga va natijalarning qay shaklda olinishi talablariga qarab ulardan birini tanlaydi. Yuqorida olingan natijalar tavsiya qilingan usulning samarali ekanligini va ularni barcha turdagi chiziqli dasturlash masalalariga tadbiq etilishi mumkinligini ko’rsatadi.

Adabiyotlar:

1. *M. Olimov, K. Ismanova, P. Karimov, SH. Ismoilov “Amaliy matematik dasturlar paketi”, Toshkent-2015.*
2. *Ракитин, В.И. Руководство по ВМ и приложения MathCad. М.: ФМ, 2005.264* с.

ANALYSIS OF ONLINE CUSTOMERS PURCHASE INTENTION USING MACHINE LEARNING ALGORITHMS

Tayba Asgher. Department of Computer Science, Riphah International University Lahore, Pakistan.

Abstract. *Online shopping is a huge business nowadays because of the rapid development of buying a product online is raised. Internet technologies play a very important role in online shopping and pay bills online by transection. Data mining is that process which can be used to extract useful data from a large amount of material or data. Now, we use data mining and in which we use machine learning algorithms to predict the appropriate transaction or purchase intention of customers or consumers. To analyzed customer predicting intention, we analyzed the customers' history-related buy product then get predicted values [1]. In this paper, we use empirical data to analyzed the purchase intention of shoppers and developed a better predicting model to check the purchase intention of customers. We have analyzed many different classification algorithms in which some are naive bayes, j48, and logistics regression to predict the customer visit to the webpage, customers buy a product or not. Our dataset online shopper's intention makes this study unique.*

Keywords: *Data Mining, Classification, Online customers intention, E-commerce*

INTRODUCTION

In the modern world, the internet plays a very superficial role in the life of humans begins. It builds a limited amount of transactions all over the world [2]. The Internet does not change only the source of business, also changed the living style of people. In the data mining community, customer attitude is complex [3]. In 2016 about 310 million people bought 136 billion goods from Amazon which is the leading online platform for product purchasing [4].

In Asia, 60% of internet users have to buy online by the year 2018 [5]. Alibaba is one of the most leading online shopping platforms, their active customer increased by 43.3% from the last or previous year [6]. Convenience, varieties of products, friendly return policy, customer reviews, etc are few factors that have been work behind online shopping [1]. For marketing, it is important to analyze the customer attitude or behavior and online purchase intension. Online customers are different from traditional customers and also the way to buy products, pay bills also different. That's why it helps us to improve the customer demands and experience, customers give us feedback in return which helps to enhance the sales.

In this paper, the analysis of data became an emerging field of research in computation and data mining. It has become very challenging and more difficult to predict the customer purchase intention because there was no direct interaction

between the customer and seller. It can be possible only in the past data of customer history.

Data mining techniques have been exploited to help enterprises for knowledge discovery and decision making by analyzing past data [7]. It is also very challenging to predict individual customer intention in real-time. Here, we analyzed a few data mining techniques, it also involved some classification algorithms and clustering.

literature review

Our aim to get knowledge about the different or some basic machine learning algorithms of classification and the categories they purchase of consumers in an online platform.

Consumers are listed on the base of their style, act, and attitude on the online market website by Moe, Wendy W [8]. Customers activities in the online purchase website turn to their intention that beliefs on that intention were classified and list clickstream of the customers. To put looks of the customer action was cluster and support to the k means clustering to look the customers. A customer who visits the page and buys the product directly are categories" Direct buying", "knowledge building", "search", "shallow". In shallow, those customers are listed who visit two pages and leave the website.

Create customer profiles based on transaction history and viewing the page of customers where the user act in real-time can be happened and enhanced the customer likelihood by using two clustering schemes [9]. Another customer clickstream data gathered for analysis [8].

The E-consumers are listed on the server based on their log data. Many actions are gathered to analyzes the appropriate marks which showed the high probability of building a purchase [9].

A content-based algorithm was used to analyze the data and prediction of the intention of the customer, it can use the approach of deep learning. So we use a different technique to predicted that a customer buys a product or not for this we use a recommendation system which finds the intention of new consumer and creates a recommendation of goods to the user accordingly [10].

methodology

Now we organized our study to use machine learning algorithms of classification and clustering. At the start, we need to prepare or preprocessing our dataset so it is reliable for use. Move forward, we check is there any missing values in our attribute. Then we apply some algorithms and get accurate results in real-time. Our study represents in figure 1.

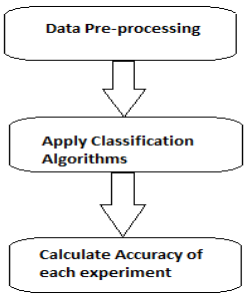


Figure:1 Study framework

Classification Algorithms

It is a supervised ML model that learns the data by using predicts the test set foe where the model didn't have knowledge of actual class and labeled train set. There are many classification models available a few of them we discussed here.

a) Naïve Bayes: Naïve Bayes is a classification method based on Bayes theorem, in simple words, it estimates the availability of appropriate action in a class and it cannot relate to the availability of any other action. This model easy to create and use a large amount of data set. It finds the dependency between the actions and in training data. In test data, actions, and training set to build on the frequency and the dependency on the base of predicted results. Naïve Bayes is a fast classifier that's why it could be used to make predictions in real-time. By using this we predict the probability of multiple classes of targeted variables and many more.

b) J48 Decision Tree: It performs classification to build a tree-like structure [11]. By using this algorithm, we split out data into small subsets in every step and also write rules of breakdown. That's why we get an appropriate decision tree in which each node gives the results to classify the data. It contains two nodes one is a decision node and the other is a leaf node. The intermediate node called classifier and the other called results. J48 implementation developed by the weka project team.

c) Logistic Regression: It is a machine learning algorithm for classification, it used for prediction analysis and relay on the concept of probability. For example, it can be extended model classes to finds an image having a glass, jug, etc. every object detected in the image would be assigned a probability in 0 and 1 and some addition. This is used in various fields of ML and other medical fields.

Clustering Algorithm

Clustering is a method of data mining that can be used to cluster the data of related groups.im other words we say that it processes the partitioning the data into a similar class group, the data of one class is much similar to the data of other clusters.no we can say that data object partitions into subclasses called cluster. The quality depends upon the methods that we used in clustering.

Simple K-Mean:

Here we use a simple K means clustering algorithm to get the quality result of our data set. It is one of the simplest and popular unsupervised algorithms of

machine learning. If it shows the k number of centroids, and allocate data point to the nearest cluster, kept the centroids as tiny as possible.

dataset and experiment

Dataset:

The dataset is present vectors belonging to 12,330 sessions. The dataset was formed so that each session would belong to a different user in 1 year to avoid any custom to a specific campaign, special day, customer profile, or period [12] [13]. Figure 2 shows the attributes.

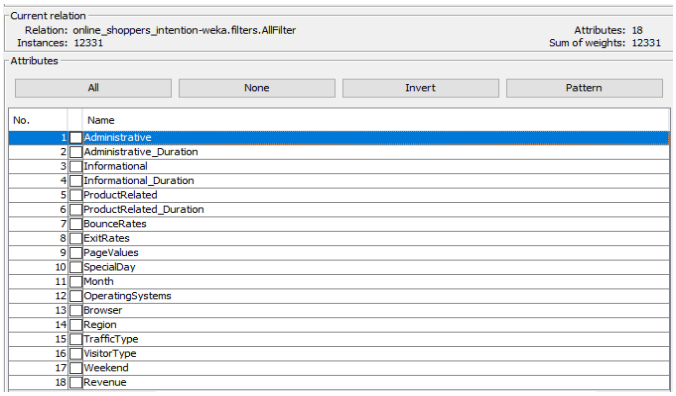


Figure:2 Attributes of the dataset

Experiment:

Now we want to perform some selected classification algorithms on our dataset. then we compare their performance and find the accuracy results. We also state that why we need pre-processing in our dataset. we fill the all missing value in our dataset then we apply classification and clustering on this dataset and get results. And also we show their visualization. We use a full training dataset for these results. For each, we calculate the accuracy rate that we show.

Logistic Regression:

Table 1 shows the summary result for this, and figure 3 shows the visualization prediction result for this algorithm.

Table: 1 Accuracy Result of Logistic Regression

=== Summary ===		
Correctly Classified Instances	10903	88.4194 %
Incorrectly Classified Instances	1428	11.5806 %
Kappa statistic	0.4463	
Mean absolute error	0.1697	
Root mean squared error	0.2918	
Relative absolute error	64.8666 %	
Root relative squared error	80.6867 %	
Total Number of Instances	12331	
=== Detailed Accuracy By Class ===		

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.976	0.619	0.896	0.976	0.934	0.479	0.894	0.976	FALSE
	0.381	0.024	0.747	0.381	0.504	0.479	0.894	0.642	TRUE
Weighted Avg.	0.884	0.527	0.873	0.884	0.868	0.479	0.894	0.925	

=== Confusion Matrix ===

a	b	<-- classified as
10177	246	a = FALSE
1182	726	b = TRUE

Figure 3 shows the appropriate result and classified results for instances and incorrectly classified instances. This is an actual summary of our working using an online shopper's intention dataset. Revenue in our dataset represents the class or act as a class in all working. It gives the prediction of the intention of our customers.

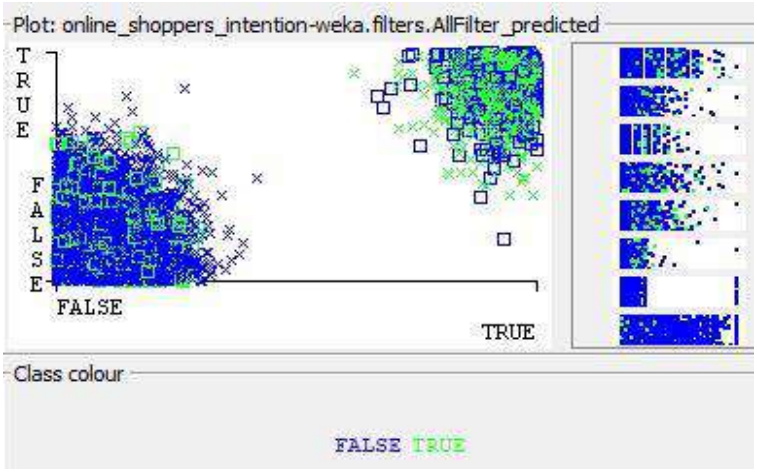


Figure: 3 Visualization Result

Naïve Bayes:

Table 2 shows the summary result for this, and figure 4 shows the visualization for this algorithm.

Table:2 Accuracy Result of Naïve Bayes

=== Summary ===		
Correctly Classified Instances	9992	81.0315 %
Incorrectly Classified Instances	2339	18.9685 %
Kappa statistic	0.4148	
Mean absolute error	0.2338	
Root mean squared error	0.3959	
Relative absolute error	89.3759 %	

```
Root relative squared error            109.4714 %
Total Number of Instances              12331

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC      ROC Area
PRC Area  Class
0.834    0.320    0.934      0.834    0.881      0.432    0.838     0.961
FALSE
0.680    0.166    0.429      0.680    0.526      0.432    0.838     0.494
TRUE
Weighted Avg.    0.810    0.296    0.856      0.810    0.826      0.432
0.838     0.889

=== Confusion Matrix ===

a    b   <-- classified as
8695 1728 |    a = FALSE
611 1297 |    b = TRUE
```

Here we use special day attribute for the customer to check the purchase of customer and then we predict using their history.

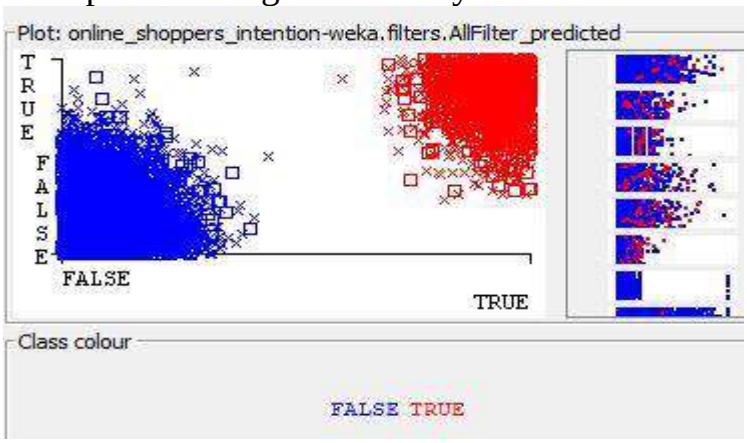


Figure: 4 Predict Revenue Visualization Result

J48 Decision Tree:

Table 3 shows the summary result for this, and figure 5 shows the visualization for this algorithm.

Table: 3 Accuracy Result

```
=== Summary ===

Correctly Classified Instances       11025               89.4088 %
Incorrectly Classified Instances      1306               10.5912 %
Kappa statistic                          0.5669
Mean absolute error                      0.1418
```

```
Root mean squared error              0.3013
Relative absolute error             54.2035 %
Root relative squared error          83.3196 %
Total Number of Instances           12331

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC
ROC Area  PRC Area  Class
                 0.952    0.422    0.925     0.952    0.938     0.570    0.787
0.915     FALSE
                 0.578    0.048    0.688     0.578    0.628     0.570    0.787
0.524     TRUE
Weighted Avg.    0.894    0.364    0.888     0.894    0.890     0.570
0.787     0.854

=== Confusion Matrix ===

    a    b   <-- classified as
 9922  501 |    a = FALSE
  805 1103 |    b = TRUE
```

Revenue prediction shows the predicted visualization for the dataset. It gives the prediction of the intention of our customers.

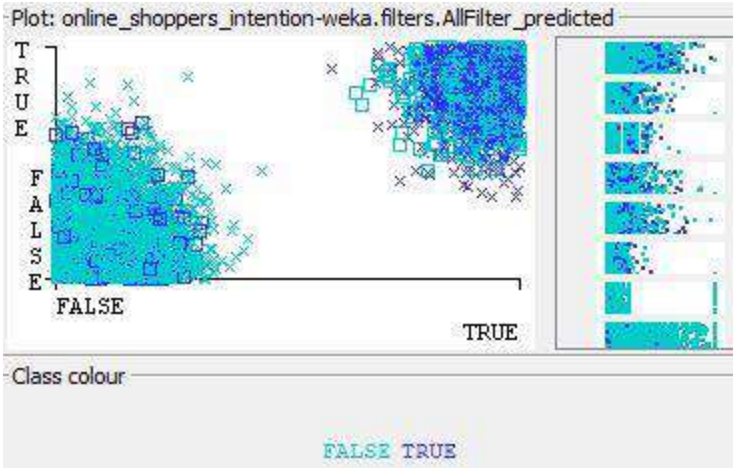


Figure: 5 visualization

Simple K-Mean:

Table 4 shows the clustering result for this, and figure 6 shows the visualization for this algorithm.

Table: 4 clustering centroids

```
Final cluster centroids:
                                        Cluster#
Attribute                  Full Data            0            1
                           (12331.0)     (3339.0)     (8992.0)
=================================================================
```

```
====================
Administrative                2.3161          2.5409          2.2327
Administrative_Duration       80.8044         91.1604         76.9589
Informational                 0.5056          0.605           0.4688
Informational_Duration        34.4642         41.8296         31.7292
ProductRelated                31.7321         31.4274         31.8452
ProductRelated_Duration       1194.5831       1194.7932       1194.5051
BounceRates                   0.0232          0.0165          0.0257
ExitRates                     0.043           0.0316          0.0472
PageValues                    5.8888          6.0752          5.8196
SpecialDay                    0.0614          0.0029          0.0831
Month                         May             Mar             May
OperatingSystems              2.1238          2.1153          2.127
Browser                       2.3569          2.2363          2.4017
Region                        3.1471          2.681           3.3202
TrafficType                   4.0693          3.7194          4.1992
VisitorType                   Returning_Visitor Returning_Visitor Returning_Visitor
Weekend                       FALSE           TRUE            FALSE
Revenue                       FALSE           FALSE           FALSE
```

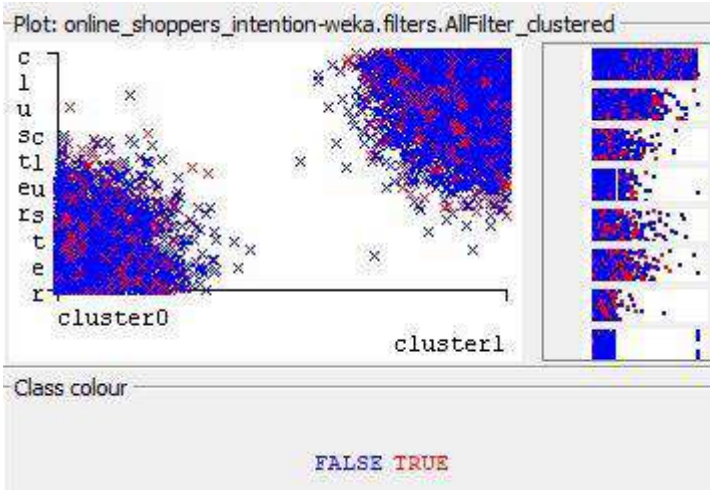


Figure: 6 visualization

results

We run many classification algorithms on our dataset and analyze the enhancement of every approach. Accuracy tells us how many percentages of test cases have been classified correctly. For correct positive results, precision gives the ratio of all positive test cases.

The accuracy of different experiments shows in table 5.

Table: 5 Results

Algorithms	Accuracy(%)
Naïve Bayes	89.40%
J48 Tree	81.03%
Logistics Regression	88.5%

We get the results in which we get the most reliable algorithm is Naïve Bayes its accuracy is more efficient than the other two algorithms.

CONCLUSION AND FUTURE WORK

In this study, we get the enhancement of different supervised machine learning algorithms on empirical data on our dataset. Our goal for this work was to identify an appropriate model that can predict the buying intention of a customer visits the web-pages of online shop more accurately. We apply some different classification and one clustering algorithm by using the Weka tool. For our experiment, we found Naïve Bayes is more suitable than other algorithms. That's why we say that Naïve Bayes for predicting the purchase intention of online consumers can make the e-commerce business more beneficial, profitable, and comfortable for all online customers and sellers.

If we want to increase the accuracy of our dataset, then we can use Ensemble Methods to enhance the accuracy rate in the future.

REFERENCES

[1] Faisal Bin Ashraf (2019) Analysis of Different Predicting Model for Online Shoppersâ€™ Purchase Intention from Empirical Data, 22nd International Conference on Computer and Information Technology: Faisal Bin Ashraf.

[2] S.Meher Taj, A.Kumaravel (February 2020) Intentions Of Online Shoppers Prediction By Fuzzy Petri Nets Construction, Volume 9, edn., International Journal Of Scientific & Technology Research: S.Meher Taj.

[3] Erevelles, S., Fukawa, N., & Swayne, L (2016) Big Data consumer analytics and the transformation of marketing, edn., Journal of Business Research, 69(2), 897-904.:

[4] (2017) Statistics and facts about global e-commerce, edn., Retrieved from https://www.statista.com/topics/871/online-shopping.

[5] (2017) Statistics and facts about e-commerce in India, edn., https://www.sta tista.com/topics/2454/e-commerce-in-india,

[6] (2018) Number of active buyers on Alibaba in Bangladesh from 2015 to 2017, edn., https://www.statista.com/statistics/898967/bangladesh-e-commerceexport-distribution

[7] Rygielski, Chris, Jyun-Cheng Wang, and David C. Yen. "Data mining techniques for customer relationship management." Technology in Society 24.4 (2002): 483-502.

[8] Yomi Kastro Inveon (n.d.) Information Technologies Consultancy and Trade, edn., 34335 Istanbul, Turkey:

[9] Mobasher, Bamshad, et al. (n.d.) Discovery and evaluation of aggregate usage proï¬les for web personalization, edn., Data mining and knowledge discovery 6.1 (2002): 61-82

[10] Shekasta, Michael, et al (2019) New Item Consumption Prediction Using Deep Learning, edn., arXiv preprint arXiv:1905.01686.

[11] Llora, Xavier, and Josep M. Garrell (2001) Evolution of decision trees, edn., Forth Catalan Conference on Artii¬ cial Intelligence (CCIA2001).

[12] C. Okan Sakar (n.d.) Department of Computer Engineering, Faculty of Engineering and Natural Sciences, Bahcesehir University, edn., 34349 Besiktas, Istanbul, Turkey.

[13] Yomi Kastro Inveon (n.d.) Information Technologies Consultancy and Trade, edn., 34335 Istanbul, Turkey.

SANOQ SISTEMALARI BILAN ISHLOVCHI ZAMONAVIY DASTURIY VOSITALAR

Sodikjanov A. A. Namangan Muhandislik Qurilish Instituti, O'zbekiston
Xamidxanov J.N., 41-maktab, Namangan viloyati XTB, O'zkekiston.

Annotation*. In this state there is a review of the system of calculations and types of types, as well as the program for the implementation of the operation. The functions of the program working with number systems were studied, examples were solved and the goal was achieved.*

Key words: digital computer technology, technical base, mathematical base, number systems, Android platform, Base Calculator software.

Аннотация*. В этой статье дается обзор систем счисления и их типов, а также выбирается программа для выполнения операций над ними. Были изучены функции программы, работающей с системами счисления, решены образцовые примеры и достигнута поставленная цель.*

Ключевые слова: цифровая компьютерная технология, техническая база, математическая база, системы счисления, платформа Android, программное обеспечение Base Calculator.

Annotatsiya. Ushbu maqolada sanoq sistemalari va ularning turlari haqida tushunchalar berildi va ular ustida amallar bajaruvchi dastur tanlab olindi. Dasturning sanoq sistemalari bilan ishlovchi funksiyalarini o'rganib chiqildi, na'munaviy misollar yechildi va ko'zlangan maqsadga erishildi.

Kalit so'zlar: *raqamli kompyuter texnologiyalari, texnik asos, matematik asos, sanoq sistemalari, android platformasi, Base Calculator dasturi.*

Hozirgi kunda axborot texnologiyalari hayotimizning ajralmas bir bo'lagiga aylanib ulgurdi. Axborot texnologiyalaridan faqat o'quv jarayonida emas, uzluksiz ta'lim tizimida faoliyat yuritayotgan o'qituvchilarni ilmiy-texnik va maxsus axborot bilan ta'minlaydigan axborot ishida, ta'lim tizimini boshqarishda shu bilan birga davlat boshqaruv tizimlarida, davlat xizmatlarini amalga oshirishda ham foydalanish mumkin. Lekin biz foydalanilayotgan kompyuter va unga asoslangan barcha texnikaning asosiy qismi raqamli kompyuterlar tashkil etishini ham unutmasligimiz kerak.

Ushbu maqolada biz raqamli kompyuterlarning matematik asosini tashkil etuvchi sanoq sistemalari va ular ustida ammallar bajaruvchi dasur haqida o'z fikr-mulohaza va tavsiyalarimizni berishga harakat qilamiz. Har bir Oliy va o'rta maxsus ta'lim muassasalarida tahsil olayotgan talabalarga Informatika va axborot texnologiyalari fanlarining kirish qismi sifatida sanoq sistemalari va ular ustida amallar bajarish mavzulari o'rgatiladi. Shunday ekan, sanoq sistemalarini bilish, o'rganish va o'rgatish, hamda u bilan bog'liq boshqa jarayon va vositalar, jumladan turli sanoq sistemalariga o'tkazish algoritmlarini o'rganish va o'rgatish uchun maxsus dasturlar yaratish ancha muhimdir. Sanoq sistemasi – sonlarni

o‘qish va arifmetik amallarni bajarish uchun qulay ko‘rinishda yozish usuli hisoblanadi.

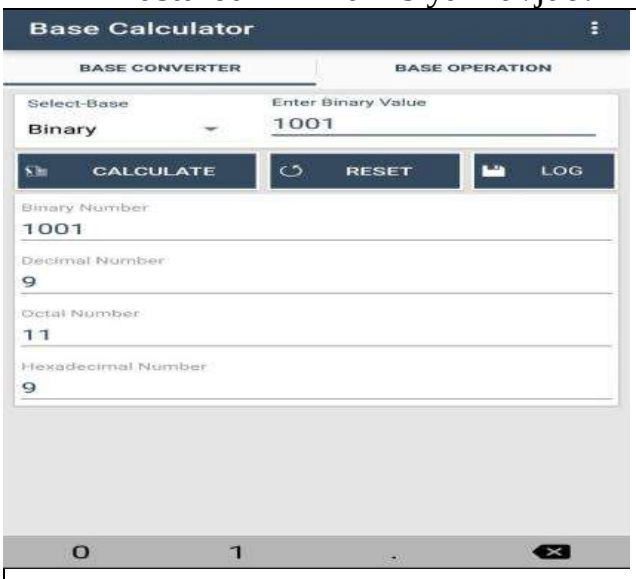
Elektron hisoblash mashinalarida 2 lik, 8 lik, 10lik va 16 lik sanoq sistemalari keng foydalaniladi. Shu bois ushbu sanoq sistemalarida barcha ishlar amalga oshiriladi. Odatda sonlarni qaysi sanoq sistemasiga mansubligi ularning quyi indeksida ko‘rsatilgan son bilan farqlanadi. Masalan, 798_{10} – o‘nli sonoq sistemasidagi 798 soni, $A87_{16}$ – o‘n oltili sonoq sistemasidagi A87 soni va h.k. bildiradi.

Ikkilik sanoq sistemasining asosiy kamchiligi – sonlardagi xona(razryad)larning juda tez ortib ketishidir. O‘nlik sanoq sistemasidan ikkilikka va teskari o‘tkazishlarni kompyuterning o‘zi bajaradi. Lekin kompyuterning imkoniyatlaridan oqilona foydalanish uchun uning tilini tushunish zarur bo‘ladi. Shular sababli sakkizlik va o‘n oltilik sanoq sistemalari ishlab chiqilgan.

Sanoq sistemalari va ular ustida amallar bajaruvchi dastur sifati biz Base Calculator dasturini tanladik. Ushbu dastur Android platformasi uchun juda qulay dastur hisoblanib,hozirga qadar 10 mingdan ortiq foydalanuvchiga ega. Foydalanuvchilar dasturni Play Marketdan yuklab olishlari mumkin.

Dasturning afzalliklari: dastur offlayn tarzda ishlaydi, dastur hajmi xotiradan uncha katta bo’lmagan joyni 12 megabaytni egallaydi holos.

Dasturda 2 xil funksiya mavjud:

	
1)Bir sanoq sistemasiga tegishli bo’lgan sonni boshqa sanoq sistemalariga avtomatik ravishda o’tkazish.	2) 2 xil sanoq sistemalari ustida amallarni bajarish: qo’shish,ayirish, ko’paytirish va bo’lish.

Shunday qilib, mazkur dasturiy vosita informarika va axborot texnologiyalari yo'nalishida ta'lim olayotgan talabalarda sanoq tizimlari ustida amallar bajarish va olingan natijalarni tahlil etishda, ishlab chiqilgan natijalarni taqqoslash va muayyan qarorlar qabul qilish ko'nikmalarini hosil qilishda ahamiyatli hisoblanadi.

СОДДА МИСОЛЛАР УЧУН МАЪЛУМОТЛАР БАЗАСИНИ ЯРАТИШ (PYTHON)

Икромов Х.Х. Андижон давлат университети, Ўзбекистон.

Аннотация: *Тезисда Python дастурлаш тилидан фойдаланувчилар учун маълумотлар базасини яратишнинг қулай усуллари кўрсатиб ўтилган. Шунингдек, маълумотлар базаси ва дастур лойиҳаси интеграциясидаги келиб чиқиши мумкин бўлган хатоликларни бартараф қилиш йўли берилган.*

Калит сўзлар: *Ядро, маълумотлар базаси, функция, ўзгарувчи, консол.*

Аннотация: *В тезисах приведены удобные методы создания база данных в языке Python для пользователей. А также, методы устронения погрешности возникающих при интеграции база данных с проектом программирования.*

Ключевые слова: *Ядро, база данных, функция, перменный, консол.*

Abstract: *Theses provide convenient methods for creating a database in Python for users. And also, methods for eliminating errors arising from the integration of the database with the programming project.*

Keywords: *core, database, function, variable, console.*

“Ядро” кодини ёзишда файллар билан бирга ишлаш инобатга олинмаган ва маълумотлар базаси дастурнинг ишлаш жараёнида рўйҳатда жойлашган ҳамда уни сақланиши учун керакли функция рўйҳатдан қаторни ажратади ёки аксинча. Бу бошқа дастурлаш тиллари билан маълумотлар базасини интеграция қилиш жараёнида қулайлик туғдиради.

Маълумотлар базаси билан бирга ишловчи функция ёзилиши натижасида маълумотлар базасини яратилади. Функцияга init ёки parse номини берилади.

init функцияси дастур кодидаги устунлар ва қаторлар сони ҳақидаги маълумотга эга бўлган рўйҳатни қайтариш қуйидаги формула ёрдамида ҳисоблайди:

Маълумот ёзилган катакларнинг умумий сони / устунлар сони.

Parse функцияси қаторни қабул қилади, аммо маълумотлар базасидаги рўйҳатни қайтаради.

```
def init():
    dblist = [1]
    #$
    return dblist
def parse(string):
    dblist = []
    var = ''
    x = 0
```

```
    y = len(string)
    #-
    if(string[y - 1] == '!'):
        while(x != y):
            while(string[x] != '!'):
                #-              var += string[x]
                x += 1
            dblist += [var]
            var = ''
            x += 1        #-
        var = dblist[0]
        dblist[0] = int(var)
        #$
        return dblist
    else:
        return '$err'
```

dblist ўзгарувчиси бу – мавжуд маълумотлар базаси, маълумотлар базасини ўзгартириш функциялари рўйҳатни параметр сингари қабул қилади, ўзгартиради ва қайтаради.

parse функцияси хатолик келтириб чиқариши мумкин, хатолик келтириб чиқарувчини ҳам четлаб ўтиш мумкин.

Қуйида маълумотлар базасидан ёзиш ва ўқиш функциялари ёзилган:

```
def write(dblist, pos, data):
    leN = len(dblist)
    leN -= 1
    #-    if (pos > leN):
        while(leN != pos):          #-
            dblist.append(' ')
            leN += 1
        #-        dblist[pos] = data
    else:        #-
        dblist[pos] = data     #$
    return dblist
def read(dblist, pos):     leN = len(dblist)
    if (pos > leN):        #$
        return '$err'
    else:        #$
        return dblist[pos]
```

Маълумотлар жойлашишини нотўғри кўрсатиш ҳам read функцияси ишлашида хатоликни келтириб чиқаради.

Энг содда устун яратиш ва ўчириш функциялари:

```
def mkcol(dblist):     dblist[0] += 1
   #$     return dblist
def delcol(dblist):     dblist[0] -= 1     #$
   return dblist
```

Ва ниҳоят, ушбу қаторда сақлаш функцияси:

```
def save(string, dblist):
   x = 0
   while(x != len(dblist)):
   #-         string += str(dblist[x])
      string += '!'         #-
x += 1
   return string
```

Хатолик келтириб чиқарувчиларни ҳам унутмаслик керак. Функция хатолик қайтарган тақдирда у консолда кўринади.

```
import core as db
# "ERR_PARSE_END-OF-DB"
# "ERR_READ_POS"
def init():
   return db.init()
def parse(string):
   ret = db.parse(string)
   if ret == '$err':
      print("ERR_PARSE_END-OF-DB")
   else:
      return ret
def save(string, dblist):     return save(string, dblist)
def mkcol(dblist):     return db.mkcol(dblist)
def delcol(dblist):     return db.delcol(dblist)
def write(dblist, pos, data):     return db.write(dblist, pos, data)
def read(dblist, pos):     ret = db.read(dblist, pos)
   if ret == '$err':         print('ERR_READ_POS')
   else:
      return ret
```

Ушбу маълумотлар базасини мураккаб бўлмаган видеоўйинларнинг персонаж ва объектлари ҳақидаги маълумотларни сақлашда қўллаш мумкин. Теглар python3даги маълумотлар базаси учун ёзилган.

Фойдаланилган адабиётлар:

1. *И. А. Хахаев – “Практикум по алгоритмизации и программированию на Python", ўқув қўлланма, “Альт Линукс” нашриёти, 2010.*

М. Лутц – “Изучаем Python”, ўқув қўлланма, “O'Reilly” нашриёти, 2011.

COMPUTER MODELING AND INFORMATION PROCESSING IN HETEROGENEOUS ENVIRONMENTS

Zhuzbayev S.S., Eurasian National University named after L.N. Gumileva, Nur-Sultan, Kazakhstan.
Adilova A.K., Eurasian National University named after L.N. Gumileva, Doctoral student, specialty Information Systems, Nur-Sultan, Kazakhstan.

Annotation: *The necessity of analyzing wave processes in a deformed body and predicting the laws of its behavior has led to the improvement of mathematical methods for solving problems. In connection with the development of information technology and accessibility to the computing power of computers, numerical solution methods have developed: the finite difference method, the splitting method and the method of spatial characteristics, etc. The article is devoted to computer modeling and information processing in heterogeneous environments.*

Keywords: *computer modelling, mathematical model, equations, bi-characteristic method.*

Аннотация: *Необходимость анализа волновых процессов в деформируемом теле и прогнозирования закономерностей его поведения привели к совершенствованию математических методов решения задач. В связи с развитием информационных технологий и доступности к вычислительным мощностям компьютеров, получили новое развитие численные методы решения: метод конечных разностей, метод расщеплении и метод пространственных характеристик и т.д. Статья посвящена компьютерному моделированию и обработка информации в неоднородных средах.*

Ключевые слова: *компьютерное моделирование, математическая модель, уравнения, метод бихарактеристик.*

The solution of modern problems of managing complex systems and objects, the large-scale implementation of information technologies, the need to increase the reliability and quality of the source information underlying the formation of effective managerial influences urgently require further development of methods and algorithms for processing experimental data, analysis and modeling of processes characterizing properties. Recently, algorithms based on the use of artificial neural networks have been gaining increasing use. Such algorithms can be very effective precisely in the analysis of heterogeneous data.

The validity and reliability of scientific provisions, conclusions and recommendations are confirmed by the correct use of methods of mathematical statistics, the theory of random processes, the apparatus of the theory of artificial neural networks, the results of simulation modeling and practical use of the developed algorithms and methods for processing heterogeneous information.

Of great interest today are the transformations that occur in solids as a result of external deformation. This interest is reflected in theoretical, and mostly in

mechanical fields. The existence of static loads and their influence during the construction or design process can be calculated. While dynamic loads can bring significant financial costs if they are not taken into account in the design. So, to study dynamic loads, it is necessary to examine the body in a spatio-temporal projection in order to analyze wave processes in this body. Therefore, the study of unsteady processes in solid mechanics is now becoming increasingly important. Since the process of wave propagation and analysis of deformation in the body cannot be carried out in real conditions, computer mathematical modeling of the wave process becomes relevant. The main task of computer modeling is the construction and manipulation of abstract mathematical or graphical representations of economic, engineering, industrial, social and other types of situations and natural phenomena modeled using a computer system [1]. The study of computer models is more convenient and is applied in all spheres of human activity: in economics, engineering, manufacturing, etc.

Computer modeling is a method for solving the problem of analysis or synthesis of a complex system based on the use of its computer model. The essence of computer modeling is to obtain quantitative and qualitative results of the model. Qualitative conclusions obtained as a result of analysis and synthesis make it possible to obtain new information about a complex object or system, such as integrity, stability, dynamics, etc. Quantitative conclusions help predict the future of a system or, conversely, explain past events or values of variables, reflecting the essence of the system. The study of computer models is more convenient and is used in cases where the actual setting of the experiment is impossible or is unpredictable.

If we denote the variables affecting the process as x_1, x_2, ..., x_n, and values after modeling as y_1, y_2, ..., y_k then you can symbolically indicate the relationship between them using the function:

$$y_k = F_j (x_1, x_2, ..., x_n); (j = 1, 2, ..., k),$$

whereF_j - is the action that must be taken with respect to xn in order to get y_k.

In a rectangular coordinate system is consideredperfectly elastic body D is a flat section, which is a rectangle $/x_1/ \le L_1$, $0 \le x_2 \le L_2$.

At the initial time, the body is at rest:

$$v_i = \sigma_{ij} = 0, (i, j = 1,2) \quad (1)$$

In addition, the left side boundary of the body is subjected to local dynamic perturbation, at:

$$x_2 = 0, | L_1 - x_1 | \le R: \sigma_{22} = f(t), \sigma_{12} = 0,$$
$$x_2 = 0, | L_1 - x_1 | > R: \sigma_{22} = 0, \sigma_{12} = 0 \quad (2)$$

The right lateral and upper border of the body are free of stress, with:

$$x_2 = L_2 , |x_1| \le L_1 : \sigma_{22} = 0, \sigma_{12} = 0,$$
$$0 \le x_2 \le L_2, x_1 = -L_1 : \sigma_{11} = 0, \sigma_{12} = 0 \quad (3)$$

Thelowerboundaryisrigidlyfixed,with:

$$x_1 = L_1, 0 \le x_2 \le L_2: v_1 = 0, v_2 = 0 \quad (4)$$

It is necessary to study the stress state of the elastic body D for t>0.

The construction of a mathematical model. To solve the problem with given initial and boundary conditions, we use a system of equations consisting of equations of motion and the generalized Hooke law:

$$\begin{cases} \sigma_{11,1} + \sigma_{12,2} = \rho \dfrac{\partial^2 u_1}{\partial t^2} \\ \sigma_{21,1} + \sigma_{22,2} = \rho \dfrac{\partial^2 u_2}{\partial t^2} \end{cases} \tag{5}$$

$$\begin{cases} \sigma_{11} = \lambda\left(u_{1,1} + u_{2,2}\right) + 2\mu u_{1,1} \\ \sigma_{22} = \lambda\left(u_{1,1} + u_{2,2}\right) + 2\mu u_{2,2} \\ \sigma_{12} = \mu\left(u_{1,2} + u_{2,1}\right) \end{cases} \tag{6}$$

where ρ- is the density, λ- μ are the Lamé constants, δ_{ij} is the Kronecker delta.

For convenience, independent dimensionless variables and unknown quantities are introduced[2]:

$$\bar{t} = \frac{tc_1}{b}; \quad \bar{x}_i = \frac{x_i}{b}; \quad \bar{v}_i = \frac{1}{c_1}\frac{\partial u_i}{\partial t}; \quad \bar{\sigma}_{ij} = \frac{\sigma_{ij}}{\rho c_1^2};$$

$$\gamma_{12} = \frac{c_2}{c_1}; \quad \gamma_{11} = 1 - 2\gamma_{12}^2; \quad (i,j = 1,2),$$

где b -characteristic length, $c_1 = \sqrt{\dfrac{\lambda + 2\mu}{\rho}}; c_2 = \sqrt{\dfrac{\mu}{\rho}}$ - speed.

In the future, the line over dimensionless parameters is omitted.

After combining the dimensionless quantities, the equations of motion (5) and the time-differentiated relations of the generalized Hooke's law (6), they take the form:

$$\begin{cases} \dot{v}_1 = \sigma_{11,1} + \sigma_{12,2} \\ \dot{v}_2 = \sigma_{21,1} + \sigma_{22,2} \\ \dot{\sigma}_{11} = v_{1,1} + \gamma_{11} v_{2,2} \\ \dot{\sigma}_{11} = \gamma_{11} v_{1,1} + v_{2,2} \\ \dot{\sigma}_{12} = \gamma_{12}^2 \left(v_{1,2} + v_{2,1}\right) \end{cases} \tag{7}$$

The indices after the decimal point mean the partial derivatives with respect to the Cartesian coordinates, the dot above the partial derivatives with respect to time.

Computer simulation today has become one of the most powerful information technologies. The study of computer mathematical modeling opens up great opportunities for understanding the connection between informatics and mathematics, physics and other sciences.

REFRENCES:

1. *Tarabrin G.T. The use of the bicharacteristics method for solving non-stationary problems of the dynamics of anisotropic massifs. //M., Structural Mechanics and Structural Analysis, 1981, No. 4, p. 38 - 43.*
2. *Tarabrin G.T. Difference schemes of wave problems of the theory of elasticity: monograph /G.T.Tarabrin; Volgograd State Technical University. - Volgograd: RPK "Polytechnic", 2000. - 148 p.*
3. *Sabitova D.S., Zhuzbayev S.S., Juzbayeva B.G. Development of an information analysis system for analyzing wave processes in a homogeneous medium // International Journal of Mechanical Engineering and Technology. – 2018. – Vol. 9(12). – P. 499-508.*
4. *Zhuzbayev S., Sabitova D., Sarsenov B. Computer mathematical modeling of wave processes // 4th International conference on computer and technology applications (ICCTA). – Istanbul, 2018. – P. 1-5.*
5. *Zhuzbayev S.S., Adilova A.K. Introduction to mathematical modeling. International Conference: The Scientific Thought of the Information Age. Poland, 2018.*
6. *Zhuzbayev S.S., Sabitova D.S, Adilova A.K. The use of information-analytical systems for the analysis of wave processes. Scientific conference of mathematicians of Kazakhstan "Actual problems of mathematics", Turkestan, 2018, p.159-162.*
7. *Sirota A. A. Analysis and computer modeling of information processes and systems / E. K. Algazinov, A. A. Sirota; Under the total. ed. prof. Doctor of Technical Sciences E.K. Algazinov. - M.: DIALOGUE-MEPhI, 2009. - 416 p.*

BLACKFIN МАХСУС ПРОЦЕССОРЛАР ЯДРОСИНИНГ АСОСИЙ ХУСУСИЯТЛАРИ

Ибрагимов С.С., Андижон машинасозлик институти, Ўзбекистон.

Annotation: *This article reviews functional descriptions of Analog Devices' Blackfin dual-core processor architecture. The architecture and components of the CPU core are described. Also included are internal memories of the Blackfin family of CPUs, first-class L1 memories in the core, and comparing tables processors.*

Keywords: *Analog Devices, Blackfin family of processors, dual-core processors, processor architecture, processor core, memory architecture, memory card, L1 memory.*

Аннотация*: В этой статье рассматриваются функциональные описания архитектуры двухъядерного процессора Blackfin компании Analog Devices. Описания архитектура и компоненты ядра процессора. Также включены внутренняя память о семействе процессоров Blackfin, первоклассная память L1 в ядре и диаграммы сравнения процессоров.*

Ключевые слова: *Analog Devices, семейство процессоров Blackfin, двухъядерные процессоры, архитектура процессора, ядро процессора, архитектура памяти, карта памяти, память L1.*

Аннотация*: Ушбу мақолада Analog Devices компаниясининг Blackfin икки ядроли процессорларни ядро архитектурасининг функционал тавсифлари кўриб чиқилган. Процессор ядросининг архитектураси ва ташкил этувчилари тасвирланган. Шунингдек Blackfin процессорлари оиласининг ядрода жойлашган биринчи даражали L1 хотиралари ва процессорларнинг хусусиятларини таққослаш жадваллари келтирилган.*

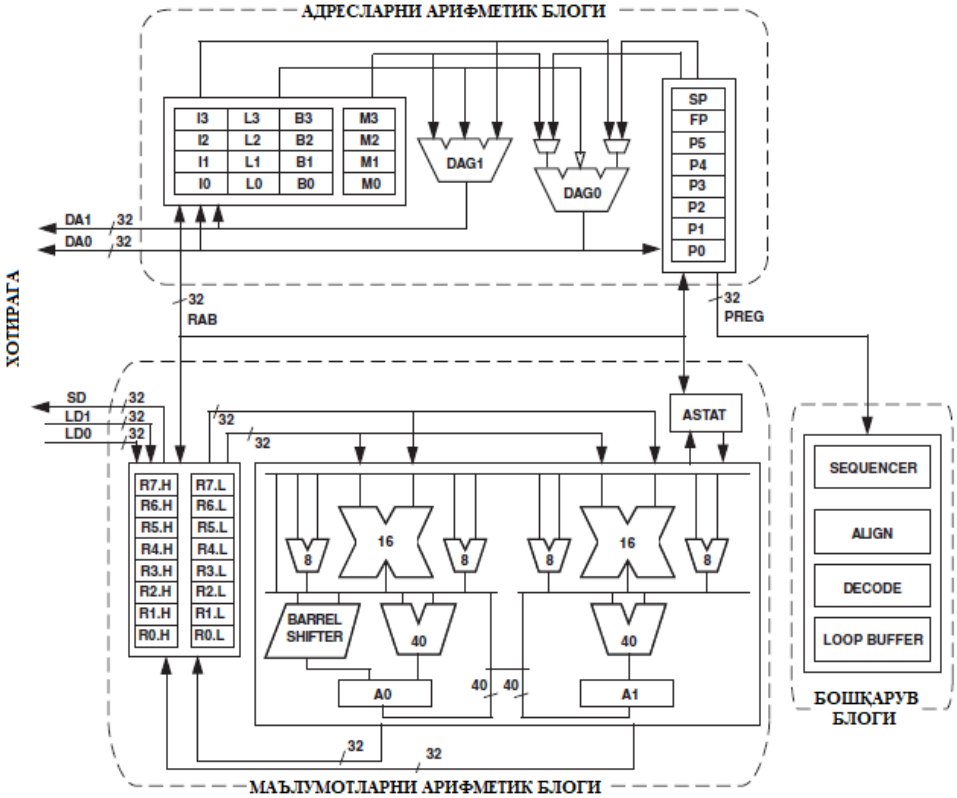
Калит сўзлар: *Analog Devices, Blackfin процессорлари оиласи, икки ядроли процессорлар, процессорлар архитектураси, процессор ядроси, L1 хотира.*

Бугунги кунда Blackfin процессорлари аудио, видео ва алоқа тизимларида маълумотларни қайта ишлаш ва энергия сарфи талабларига жавоб берадиган 16-32 разрядли микропросессорнинг янги авлоди ҳисобланади. Blackfin Процессорлари Intel корпорацияси билан биргаликда ишлаб чиқилган MSA (Micro Signal Architecture) архитектураси асосида RISC процессорларига хос бўлган 32-разрядли буйруқлар тўплами ва иккита блокли 16-разрядли MAC кўпайтириш-қўшиш сигналларни қайта ишлаш функциясини универсал микроконтроллерларда мавжуд фойдаланиш учун қулай хусусиятлари билан бирлаштиради.

Ушбу мақолада Analog Devices компаниясининг Blackfin икки ядроли процессорларни ядро архитектурасининг функционал тавсифларини кўриб чиқамиз.

Икки ядроли Blackfin процессорлари маълумотлар узатишнинг юқори даражада ўтишини таъминлаш учун комплекс бошқариш ва сигналларни қайта ишлаш вазифаларини талаб қиладиган ўлчаш, тиббий тизимлар ва маиший электрониканинг кенг доираси учун оптималлаштирилган. Blackfin процессорлари битта интеграциялашган пакетда периферик қурилмалари ва хотирага бой тўпламни бирлаштирган ҳолда сигналларни қайта ишлашга мўлжалланган дастурлар учун афзалроқ платформа ҳисобланади.

Процессор ядросининг архитектураси. Процессор юқори самарали инфратузилма, катта ҳажмдаги ички хотира, шу жумладан USB 2.0 HS OTG, иккита 10/100 Ethernet MAC контроллерлари ва RSI (Removable Storage Interface) интерфейси каби такомиллаштирилган периферик модулларини ўз ичига олган кенгайтирилган алоқа интерфейсига эга. Процессорлар 500 МГц гача ишлашни, шунингдек, кам статик қувват сарфини таклиф қилади. Кам қувватли ва паст кучланишли дизайн методологияси билан ишлаб чиқарилган бўлиб, улар юқори даражадаги қувватни бошқариш ва ишлашни таъминлайди.



1-Расм. ADSP-BF60x Процессорларининг ядроси архитектураси

Blackfin ADSP-BF60x процессорларининг ҳар бир ядросида иккита 16-битли кўпайтиргич, иккита 40-битли аккумулятор, иккита 40-битли арифметик мантиқий қурилмалар (АМҚ), тўртта 8-битли видео АМҚ ва битта 40 битли сургич мавжуд. Ҳисоблаш бирликлари регистр файлидан 8-, 16- ёки 32-битли маълумотларга ишлов беради (1-расм).

Хисоблаш регистрлари файлида саккизта 32 битли регистрлар мавжуд. 16-битли операндларнинг маълумотлари билан ҳисоблаш ишларини амалга ошираётганда, регистр файллари 16 та 16-битли регистрлар каби ишлайди.

Ҳисоблаш амаллари учун барча операндлар кўп тармоқли регистрлар файлидан ва доимий майдонлардан олинади. Ҳар бир MAC битта тактда 16 та 16 битгача кўпайтириши ва натижаларни 40 битгача жамлай олиши мумкин. Бунда ишорали ва ишорасиз форматлар қўлланилиши мумкин.

Қуйидаги 1-жадвалда Blackfin икки ядроли махсус процессорларни бошқа бир ядроли процессорлар билан бир неча хусусиятлари таққослаб келтирилган.

1-жадвал.

Blackfin процессорларининг хусусиятларини таққослаш жадвали

Процессор тури	Ядро сони	Такт частотаси МГц	ишлаш самарадорлиги (MMACS)	L1 ички хотираси
ADSP-BF504	1	400	800	68К байт
ADSP-BF512	1	400	800	116К байт
ADSP-BF527	1	600	1200	132К байт
ADSP-BF531	1	400	800	52К байт
ADSP-BF533	1	600	1200	148К байт
ADSP-BF538	1	533	1066	148К байт
ADSP-BF542	1	600	1200	132К байт
ADSP-BF561	2	600	2400	192К байт
ADSP-BF606	2	400	1600	0.5M byte
ADSP-BF607	2	500	2000	0.5M byte
ADSP-BF608	2	500	2000	0.5M byte
ADSP-BF609	2	500	2000	0.5M byte

Хулоса: ADSP-BF60x процессорлар оиласи CrossCore® Embedded Studio платформасида C/C++ дастурлаш тилларидан фойдаланиб дастурлаштирилади, шунингдек юқори тезликда ишлайдиган электрон эмуляторлари, EZ-Kit баҳолаш платалари ва махсус EZ-Extender® кенгайтириш платаларида қўлланилади. Икки ядроли процессорлар реал вақт режимида ишлаши керак бўлган рақамли ишлов бериш мосламаларида параллел дастурлаш имконини беради. Бу эса сигналларни қайта ишлаш алгоритмларини юқори тезликда бажарилишини таъминлайди.

Адабиётлар

1. Вальпа О.Д. Разработка устройств на основе цифровых сигнальных процессоров фирмы Analog Devices с использованием Visual DSP++. – М.: Горячая линия-Телеком, 2007. – 270 с.
2. DSP-BF60x Blackfin® Processor Hardware Reference, Analog Devices, Inc. 2013.

Blackfin Dual Core Embedded Processor ADSP-BF606/ADSP-BF607/ADSP-BF608/ADSP-BF609, Analog Devices, Inc. All rights reserved. Technical Support, 2014.

DASTURIY TA`MINOTNI ISHLAB CHIQISH JARAYONIDA TEKSHIRISH VA TESTLASHNING O`RNI.

Zulunov R.M. TATU Farg'ona filiali, O'zbekiston.

Annotatsiya: *Maqolada dasturiy ta'minotni ishlab chiqishda testlash va tekshirishning o'rni, bosqichlari va shu jarayondagi dolzarb muammolarni yechish yo'llari qo'rib chiqilgan.*

Kalit so'zlar: *dastur, testlash, verifikatsiya, validatsiya.*

Аннотация: *В статье рассматриваются роль и этапы тестирования и верификации в разработке программного обеспечения, а также способы решения текущих проблем в процессе.*

Ключевые слова: *программа, тестирование, верификация, валидация.*

Abstract: *The article discusses the role and stages of testing and verification in software development, and ways to address current issues in the process..*

Keywords: *program, testing, verification, validation.*

Dastur tuzish quyidagi bosqichlardan amalga oshiriladi:

-Vazifani qo'yilishi. Bu bosqichda dasturchi buyurtmachi yordamida echilishi kerak bo'lgan vazifani qo'yadi. Texnik topshiriq tuziladi. Bunda dasturning asosiy xarakteristikalari, muddatlar va ma`sul shaxslar aniqlanadi.

-Algoritmni tuzish. Dasturchi vazifani tahlil qilib kerakli bo'lgan algoritmni tanlaydi. Tanlangan algoritm to'liq tahlil qilinadi va uning blok-sxemasi chiziladi.

-Dasturlashtirish bosqichi. Dastur yaratish tili tanlanadi. Dastur qabul qilingan algoritmda tuziladi.

-Dasturni tuzatish bosqichi.

-Dasturni testdan o'tkazish bosqichi.

Odatda dasturni testdan o'tkazishda bosqichlarga bo'lib o'rganiladi. Bunda har bir modulni tekshirishdan tortib, to butun tizimni yakuniy tekshirishlar kabi bosqichlarni oladi. Agar bunda biron bir ishonchli ketma - ketlikka yondashilmasa, ishonchli ta`minlovchi dastur olish juda qiyindir. Testlash strategiyasi ikkita usuldan birortasiga asosan bajariladi: odatiy quyidan - yuqoriga qarab testlash, yoki zamonaviy yuqoridan - pastga qarab testlash.

Quyidan - yuqoriga qarab testlash.

Bu usul keng tarkalgan usul bo'lib, unda eng kuyu pog'onadagi boshlangich yozilgan modullar tekshiriladi. So'ngra yuqori qatlamdagi elementlar dasturlanadi va testlanadi. Bu jarayon to yozilgan dastur butunlay yakunlanmaguncha davom etadi. Quyidan - yuqoriga qarab testlash usuli xozirgi vaqtda yuqoridan - pastga qarab testlovchi va dasturchilvr tomonidan qo'llanmayapti. Bu usulda interfeys va algoritmdagi ko'pgina xatolar aniqlanmay qolib ketmokda. Bu esa dasturni qayta-qayta o'zgartirishdan so'ng buzishga olib keladi. Ikkinchi kamchiligi esa: har xil pog'onadagi elementlarni testdan o'tkazishda yangidan yangi testlovchi moslamalarni, drayverlarni va testlovchi ma`lumotlarni talab qilmoqda. Bu esa o'z-o'zidan dasturlashda katta hajmda mehnat talab qiladi.

Yuqoridan pastga jarayoni.

Bu testlash usuli yuqoridan pastga qarab dasturlashni, yuqoridan pastga qarab kodlashni qo'shimcha bosqichi hisoblanadi. Bu usulda oldin asosiy dastur yoziladi va so'ngra past pog'onadagi loyihalanmagan elementlar urin bosuvchi dasturlar bilan almashtiriladi. Bunday skeletli dastur chaqiriluvchi dastur va har qanday malumotlar yo'qligida ham o'z ishini davom ettiradi. Bu tekshirish natijasida bazi hollarda bemani bo'lgan xatolar ham aniqlanadi. Keyingi qadam modul qo'shilishidan iborat bo'lib, unda bu modullar kiruvchi modullarni ko'paytiruvchi bo'lishi ham mumkin, - bu esa kiritish moduli, bazi bir yordamchi modul (oxirgini dasturlash tugash daqiqasiga qadar) bo'lishi mumkin. Bu tekshirishdan so'ng sinash oddiy bir sodda kiruvchi ma`lumotlar bilan o'tkazish mumkin.

Dasturni testlash jarayonida muxim shartlardan biri shuki dasturlashga qancha vaqt sarf bo'lgan bo'lsa, testlashga ham shuncha vaqt ajratish kerak. Chunki dasturlashda rejalashtirish ishlariga intilish (tendentciya) mavjud bo'lib, shuning uchun barcha loyxalarni muddatlarini buzgan xolda bajarilishini tushinish mumkin. Bu esa ishlab chiqishga, kodlashga, sozlashga va dasturni testlashga aloxida vaqtlar ajratishni va biron bir jarayonni bajarish vaqtini biror xaftaga surilsa, kolgan bosqichlarni bajarish mudatlari ham surilishini kuz oliga keltirishi kerak.

Test o'tkazish bosqichlari:

1. Normal sharoitda tekshirish.
2. Ekstremal sharoitda tekshirish.
3. Favqulot holatlarda tekshirish.

Normal sharoitda tekshirish.

Dasturni ishlash sharoitidan chiqqan xolda olingan kiymatlar asosida tekshirish. Dastur normal sharoitda to'g'ri qiymatlar chiqarishini ko'rsatish kerak.

Ekstremal sharoitda tekshirish.

Bu tekshirishda dasturdagi o'zgaruvchilarning chegaraviy kiymatlari asosida test utkaziladi. Masalan: Eng kichik kiymatdan maksimal kiymatga ega (sonli o'zgaruvchi uchun) hamma belgilarni chop qilishi (belgilik o'zgaruvchilari uchun). Bu tekshirishda o'zgaruvchilarni nol kiymati ham bo'ladi.

Favqulot holatda tekshirish.

Bu tekshirish o'zgaruvchining kiymatlar soxasidan tashqarida etgan kiymatlar asosida tekshirish qiladi.

Rejalashtirish.

Testdan o'tkazish jarayoni vaqtida dastur yozish bilan bir bo'lganligi hisobga olib testdan o'tkazishni ham rejalashtiriladi.

Testlash rejasiga odatda quyidagi tiplar kiradi.

1. Dasturni algoritm sxemasi bilan solishtirish.
2. Displey ekranida dasturni vizualg' nazorat qilish yoki dasturiy blankadagi originalni dastur raspechatkasini urganish, solishtirishni vizual urganish.
3. Dasturni mashina tiliga olib ko'rsatish. Bu bosqichda sintaktik xatolar aniqlanadi. Dastur listingidagi sintaktik xatolarni fortran, paskalp tilidagi ko'ramalar deagnostik xabar beradi. Xatolar xaqidagi xabarlar har xil EVM larda

va tizim versiyalarida (fortran, paskalp, beysik) shakliga qarab har xil bo'lishi mumkin.

4. Dasturni kompanovkalash va tashki aloqalarni taxrirlash.

Tashki aloqalarni taxrirlash jarayonida tashki aloqalar taxririy dasturi yoki vazifalarini tartiblovchi shunday sintaktik xatolarni aniqlaydiki, bu xatolarda kichik dasturlarda yozilgan prametrlar kiymati bir biriga mos kelmaydi, aslida yo'q bo'lgan standart dasturlarni chiqaradi. Masalan SIN o'rniga SIH va shunga o'xshash xatolar.

5. Dasturni bajarilishi.

Translyator va tashki aloqalar taxriri tomonidan barcha xatolar bartaraf etilgandan so'ng keyingi - EHM da dasturni mashina tilida bajarish bosqichiga utiladi:dastur operativ xotiraga yuklatiladi va boshlangich ma`lumotlar kiritilgandan so'ng boshlanadi. Xato borligi xaqida belgi paydo bo'lishi sozlash o'tkazishna sabab bo'ladi: belgi yo'qligi dasturda xato yo'qligini bildiradi.

Testlash rejasi boshlangich ma`lumotlarning ruxsat etilgan kiymatlarini xulosasini to'g'riligini ham tekshirishni uz ichiga jamlaydi. Dasturni testlashning oxirgi bosqichi o'zgartirishlarsiz ruxsat etilan chegarasidan tashqarida joylashgan kiymatlaridan iborat bo'lgan malumotlardan foydalangan xolda bajariladi. Dasturlarni favqulotda va qutilmagan xollarda tekshirishda bir necha maslaxatlar beriladi. Bu maslaxatlar tekshirish vaqtini tayorlashga yordam beradi.

Dasturni to'g'riligi.

Xamma dasturlar mantiqan olib karaganda ayrim bir ma`lumotlar ta`sir kursatuvchi xududlarni aniq ko'rsatishi, ya`ni dastur ish bajarish kobiliyatiga ega bo'lishi, malumotlarni kursatilgan chegaralarda turganligini aniqlash uchun operatorlarni kiritish imkoniga ega bo'lishi kerak.

Adabiyotlar:

1. Mayers, Glenford J., Corey Sandler and Tom Badgett. The art of Software Testing. John Wiley & Sons, 2011. Third Edition.
2. K.Naik and P.Tripathy. "Software testing and quality assurance", Wiley 2008.
3. Соммервилл, Иан. Инженерия программного обеспечения, 6-е издание.: Пер. с англ. – М.: Издательский дом "Вильямс", 2002. – 624 с.
4. С. Канер, Д. Фолк, Е. Нгуен. Тестирование программного обеспечения. — К.: Диасофт, 2000. — 544 с.

Г. Майерс. Искусство тестирования программ. — М.: «Финансы и статистика», 1982. — 176 с.

ВЕЙВЛЕТЫ И ИХ ИСПОЛЬЗОВАНИЯ В МЕДИЦИНЕ.

Тожибоева Ё.Р., Хакимова З.К..
Андижанский государственный медицинский институт, Ўзбекистон

Аннотация: Ушбу мақолада ностоционар рақамли сигналларни қайта ишлашда вейвлет функциялар ёрдамида тахлил қилиш усулининг қулайлиги ва ушбу масалаларнинг ечилиши келтирилган. Масалаларнинг медицина, ҳамда турбулент майдонларни ҳусусиятларини ва бошқа соҳаларда қулай имкониятлар мавжудлиги кўрсатилган.

Калит сўзлар: Вейвлет, MPEG-4, JPEG-2000, Analog Devices, NASAэ

Аннотация: В данной статье приведено анализ обработки нестационарных цифровых сигналов с помощью вейвлет функции. Приведена решения примерных задач в области медицины и турбулентных полей.

Кльючевые слова: Вейвлет, MPEG-4, JPEG-2000, Analog Devices, NASA.

Abstract: *In this issue the ways of analyzing fitting imtu shapes of outpatient number signals with the help of wavelet function was led. The decisions of exemplary tasks were ed in the field of medicine and turbulent spheres*

Keywords: wavelet, MPEG-4, JPEG-2000, Analog Devices, NASA.

В последнее десятилетие в мире возникло и оформилось новое научное направление, связанное с так называемым вейвлет - Слово “wavelet”, являющееся переводом французского “ondelette”, означает небольшие волны, следующие друг за другом. Можно без преувеличения сказать, что вейвлеты произвели революцию в области теории и практики обработки нестационарных сигналов. В настоящее время вейвлеты широко применяются для распознавания образов; при обработке и синтезе различных сигналов, например речевых, медицинских; для изучения свойств турбулентных полей и во многих других случаях.

Особо большое развитие получила практика применения вейвлетов для решения задач сжатия и обработки изображений, являющихся нестационарными по своей природе. В этой области применение вейвлет - позволило достичь одновременного снижения сложности и повышения эффективности кодеров. Вейвлет – преобразование легло в основу международных стандартов MPEG-4, JPEG-2000 стандарта на сжатие отпечатков пальцев ФБР, видеокодеков фирмы Analog Devices.

Огромный интерес к изучению теории и практики вейвлет вызвал лавинообразный поток издающейся литературы. В США и других развитых странах ежегодно издаются десятки книг, учебных пособий, тематических выпусков журналов, посвященных данной тематике. Теория и практика вейвлет - находится на стыке различных наук: математики, физики, медицна и т.д. Первое упоминание о вейвлетах появилось в литературе по цифровой

обработке и анализу сейсмических сигналов (работы А.Гроссмана и Ж.Морлета). Так как интерес авторов заключался в анализе сигналов, набор базисных функций был избыточным. Далее, математик И.Мейер показал существование вейвлетов, образующих ортонормальный базис. Дискретизация вейвлет - была описана в статье И.Добоши, которая перекинула мост между математиками и специалистами в области обработки сигналов. Дебоши разработала семейство вейвлет – имеющих максимальную гладкость для данной длины фильтра. Популярность вейвлетов увеличилась после введения С.Маллатом концепции кратно масштабного анализа. Он же, первым применил вейвлеты для кодирования изображений. И И.Добеши, и С.Маллат показали, что практическое выполнение вейвлет - осуществляется посредством двух полосного банка фильтров анализа – известного ранее в теории суб полосного кодирования. Эта теория может быть описана в терминах вейвлетов. Главное различие между этими двумя направлениями заключается в критериях построения фильтров.

Некоторые идеи теории вейвлетов частично были разработаны уже очень давно. Например, А.Хаар опубликовал в 1910 году полную ортонормальную систему базисных функций с локальной областью определения. Эти функции называются теперь вейвлетами Хаара. В настоящее время исследования в области вейвлетов ведутся по многим направлениям. Несмотря на то, что теория вейвлет - уже в основном разработана, точного определения, что же такое "вейвлет", какие функции можно назвать вейвлетами, насколько известно, не существует. Обычно под вейвлетами понимаются функции, сдвиги и растяжения которых образуют базис многих важных пространств. Эти функции являются компактными как во временной, так и в частотной области.

Вейвлеты непосредственно связаны с кратно масштабным анализом сигналов. Вейвлеты могут быть ортогональными, полу ортогональными, биортогональными. Эти функции могут быть симметричными, асимметричными и несимметричными. Различают вейвлеты с компактной областью определения и не имеющие таковой. Некоторые функции имеют аналитическое выражение, другие – быстрый алгоритм вычисления связанного с ними вейвлет. Вейвлеты различаются также степенью гладкости. Для практики желательно было бы иметь ортогональные симметричные (асимметричные) вейвлеты. К сожалению, доказана теорема о том, что такими вейвлетами являются лишь вейвлеты Хаара. Функции Хаара не обладают достаточной гладкостью и не подходят для большинства приложений, поэтому для кодирования изображений обычно используют биортогональные вейвлеты.

В настоящее время многие исследователи понимают под вейвлетами более широкий класс функций. Это и вейвлет - локальные тригонометрические базисы (вейвлеты Малвара), и мультивейвлеты, и так называемые вейвлеты второго поколения, не являющиеся сдвигами и растяжениями одной функции. В самом деле, сходные идеи появлялись на

протяжении последних десятилетий: субполосное кодирование, успешно применяемое при кодировании речи, пирамидальные схемы кодирования изображений, преобразование и функции Габора (вейвлеты Габора). С развитием теории вейвлетов произошло как бы объединение, взаимопроникновение, взаимообогащение этих идей, что привело к качественно новому результату. Так как с точки зрения практики наиболее интересными представляются быстрые алгоритмы вычисления вейвлет.

Широко распространенные приложения мультимедиа (графика, аудио, видео) с каждым днем предъявляют все более высокие требования к аппаратной базе компьютера. Ни наращивание тактовой частоты процессора, ни увеличение объема жесткого диска, ни улучшение пропускной способности каналов передачи данных не в состоянии спасти положение. Единственным путем решения этой проблемы является разработка эффективных алгоритмов видео компрессии. Задача написания новых программ видео сжатия чрезвычайно актуальна для создателей цифровых систем видеонаблюдения - ведь именно в этой области постоянно приходится обрабатывать и хранить большие объемы видеоданных.

Под видео компрессией обычно понимается сокращение объема памяти, необходимой для хранения цифровых видеоданных и передачи их по каналам связи. Цель видео компрессии - более компактное представление изображений.

Для того чтобы вейвлеты хорошо аппроксимировали исходный сигнал, они подвергаются масштабированию (сжатию или растяжению) и сдвигу (смещению). Результат вейвлет-преобразования - обычный массив числовых коэффициентов. Такая форма представления информации об изображении очень удобна, поскольку числовые данные легко обрабатывать. После этого наступает очень важный этап - пороговое преобразование. Нужно отбросить коэффициенты, значение которых близко к нулю. Следует помнить, что при этом происходит необратимая потеря информации, ведь отброшенные коэффициенты участвуют в формировании изображения. Поэтому выбранное пороговое значение коэффициентов сильно влияет на качество изображения - задание слишком высокого порога повлечет за собой падение качества. Итак, видео компрессия происходит в два этапа - на первом осуществляется сжатие с потерей информации (вейвлет-преобразование), на втором – обычная архивация данных. Для восстановления изображения необходимо повторить все действия в обратном порядке. Сначала восстанавливаются значения коэффициентов, а затем по ним, применяя обратное вейвлет-преобразование, получают изображение (сигнал).

В качестве практического применения вейвлет – преобразования рассмотрены современные подходы к сжатию изображений. Вейвлет - анализ нашел широкое применение во множестве приложений - в медицине, в биологии, в нефтегазовой отрасли, в телекоммуникациях. ФБР активно использует вейвлеты для оптимизации алгоритмов хранения

дактилоскопических баз данных, а NASA разрабатывает технологию применения вейвлет-анализа к задачам освоения космического пространства.

Между тем, во многих областях можно ожидать существенно лучших результатов за счет использования вейвлетов. Перечислим некоторые из них. Задачи, связанные с предсказанием. Это - предсказание потологии оргонизма, землетрясений, прогноз погоды.

В заключение можно отметить, что вейвлеты и сопутствующие им идеи внесли неоценимый вклад в теорию и практику кодирования изображений и, будут оставаться основным направлением исследований в этой области в ближайшем будущем.

Литературы

1. *Воробьев В.И., Грибунин В.Г. «Теория и практика вейвлет преобразования» ВУС, 1999. С.1 -204.*
2. *И.М. Дремин, О.В. Иванов, В.А. Нечитайло. Вейвлеты и их использование. – Успехи физических наук, 2001.*
3. *А. Петров. Вейвлеты и их приложения – Рыбинск, РГАТА 2007.*

SARS CLASSIFICATION IN RAPIDMINER, WEKA AND ORANGE

Bilawal Ali Jaan, Muhammad Saqib. Riphah international University Lahore, Pakistan

Abstract. *In this paper compare dissimilar type of classification algorithm working with several totally free available data mining and data discovery programs like WEKA, Rapid miner and also Orange. The Accuracy of classification algorithm such as KNN, Logistic Regression and also Naïve Bayes algorithm are compared utilizing each of 3 components. SARS 2003 Outbreak Dataset (1) data-set is employed for analyzing the Classification algorithm so as to resolve the individuals without SARS disease. Key Phrases: SARS, Classification.*

Keyword: *classification, Rapidminer, Weka, Orange, KNN, Naïve Bayes, Logistic Regression, SARS.*

1. INTRODUCTION

Due to the fact the info remains really rising, it will become problematic for a person to physically assess the exact info to tactical decision-making. Ergo people need to have assistance of data mining to mine exciting data out of the information that is available. Data mining is the method of finding interesting information from considerable sizes of information saved in data bases, data warehouses, or any different data repositories. Some of many major issues in data mining would be that the Classification that will contain finding regulations which partition specified info into predefined categories. From the find realm at which trillions of info can be properly used, the application timing of present calculations could eventually become time-consuming. Thus, there's a demand for automatic applications which may assist individuals in altering those big data right into Info.

Now we have lots of free basis data-mining applications and tool are all offered for usage such as for instance the RapidMiner (2), Waikato Environment Knowledge Analysis (WEKA) (3), Orange (4), and NLTK etc. All these courses and bids furnish a variety of algorithms and methods which support in improved diagnosis of information. These programs assist in audience investigation, information conception, reversion analysis, Conclusion bushes, Prognostic analytics, Text mining.

We ran a comparison analysis amongst classification algorithm including as KNN, Logistic Regression and also Naïve Bayes algorithm with WEKA, RapidMiner and also Orange instrument. The precision pace; that mirrors the proportion of correctly labeled examples, can be useful for estimating the operation of the classification algorithm.

2. METHODOLOGY OF LITERATURE REVIEW

In Rational Records such as (Science Direct, Springer Link, PubMed, and Google Scholar) were searched between the periods April 30, 2020, to May 3, 2020. Find a Appropriate title and keywords and used. After a search in detail, we

have a number of articles found multiple keywords were selected, based on their significance to target studies in SARS Classified,' Science on SARS. A mixture of keywords, thrown in for a detailed search for each specific data. Scientific articles on practical subjects, only for humans (in human subjects), and published in English, were considered for a systematic review.

3. CLASSIFICATION ALGORITHMS

Info Classification Algorithm can be really a method of choosing the hypothesis by a pair of choices which best suits a pair of observations. Info Classification Procedure comprises two measures:

I) Assembling the Classifier product: This the classifier is assembled by finding out exactly the practice group along with also their associated class tags. ii) Utilizing Classifier for Classification: During this measure, the classifier can be useful for classification. This evaluation data can be utilized to gauge the validity of classification regulations.

We've researched the next Classification Algorithm within our paper:

- KNN
- Naïve Bayes
- Logistic Regression

3.1 KNN

Is a type of instance-based learning, or lazy learning, where the function is only approximated locally and all computation is deferred until function evaluation. Both for classification and regression, a useful technique can be to assign weights to the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones.

3.2 Logistic Regression

In statistics, the logistic model is used to model the probability of a certain class or event existing such as pass/fail, win/lose, alive/dead or healthy/sick. This can be extended to model several classes of events such as determining whether an image contains a cat, dog, lion, etc.

3.3 Naive Bayes

Is a household of uncomplicated probabilistic classifiers in line with employing Bayes' theorem with solid independence assumptions amongst your qualities.

4. TOOLS GUIDE

4.1 RapidMiner

It has always been the study of machine learning and data analysis processes. Its open-source, free trade is fully implemented in Java. RapidMiner is a data science software developed by the company of the similar name that delivers an integrated situation for deep learning, data preparation, text mining, machine learning, and predictive analytics installation data in several ways. It contains 100 great learning strategies such as differentiation, tolerance and bonding responsibilities.

4.2 WEKA

Weka is attempted and tried open-source AI programming that can be gotten to through a graphical UI, standard terminal applications, or a Java API. It is generally utilized for educating, explore, and mechanical applications, contains a plenty of implicit apparatuses for standard AI undertakings, and also gives straightforward access to notable tool compartments, for example, scikit-learn, R, and Deeplearning4j.

4.3 Orange

It has always been a learning process for learning and using databases (written with Python). It is a transparent system for applications to get visual data and analysis, and it can likewise serve as a Python library. Orange is an open-source data visualization, ML and data mining toolkit. It features a visual programming front-end for interactive data visualization and explorative data analysis. Orange can be a very efficient software system for data mining and machine learning and statistics. The elements can be accessed by widgets and all include direct data visualization, stand-alone selection and preparation, robust investigations of understanding calculations and forecasts. After searching the data selection and the external read-out process to be deleted from the Data and Structure after the entire page has been read and you get to explore the phylogenetic tree.

5. EXPERIMENT

5.1 Dataset

We have downloaded sars_2003_complete_dataset_clean from the Kaggle [4]. We have total numbers of a row are 2538. In this data collection, we have No Certified Death Records and number recovered record. The data was collected from Kaggle. Certified death records or No is a class label used to divide into groups (patient death of SARS or Recovery). This data collection has 37 country SARS records and March 2003 to July 2003 SARS patient records. The dataset has the following attributes:

1. Country
2. Date
3. Cumulative no of cases
4. Number of deaths
5. Nu
6. mber recovered

5.2 Classification Algorithm using Rapid miner:

In Rapid miner we get highest accuracy in Naive Bayes and lowest accuracy in Logistic Regression as shown in fig 1. It’s a crucial point from the RapidMiner we get different accuracy from the different classification algorithms like Naïve Bayes, Logistic Regression. These algorithms are applied on the same dataset (sars_2003_complete_dataset_clean).

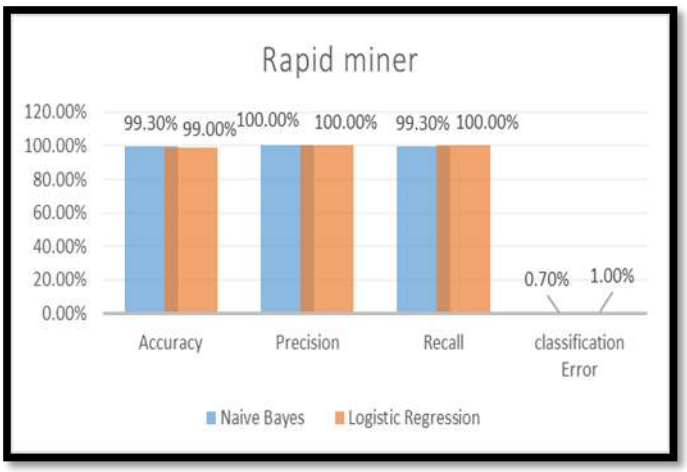


Figure 1: RapidMiner finding

5.3 Classification Algorithm using Orange

In Orange we get highest accuracy in Logistic Regression 93.2% and lowest accuracy in Naive Bayes 89.60% as shown in fig 2. It's a vital fact from the Orange we get different accuracy from the different classification algorithms like Naïve Bayes, Naive Bayes and Logistic Regression. These algorithms are applied on the same dataset (sars_2003_complete_dataset_clean). Same as theses algorithms are applied through the tool weka for the to measure the accuracy of this dataset from the different machine learning tools.

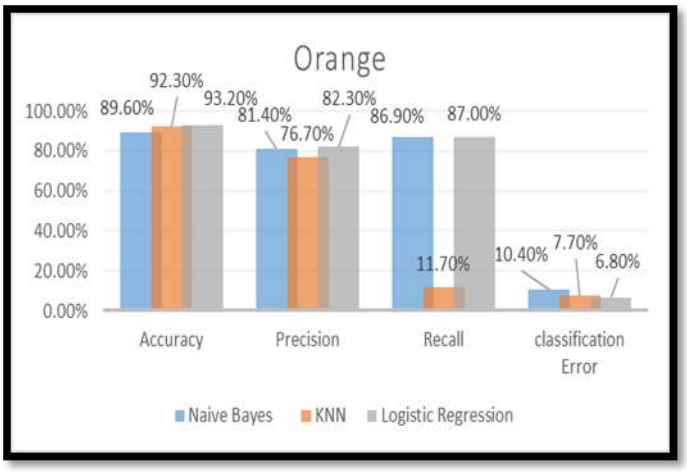


Figure 2: Orange Finding

5.4 Classification Algorithm using Weka:

Weka is the most common and the crucial tool in the machine learning but there are some drawbacks in weka like weka not support the big dataset files. In Weka we get highest accuracy in KNN 74.70% and lowest accuracy in Naive Bayes 69.50% as shown in fig 3.

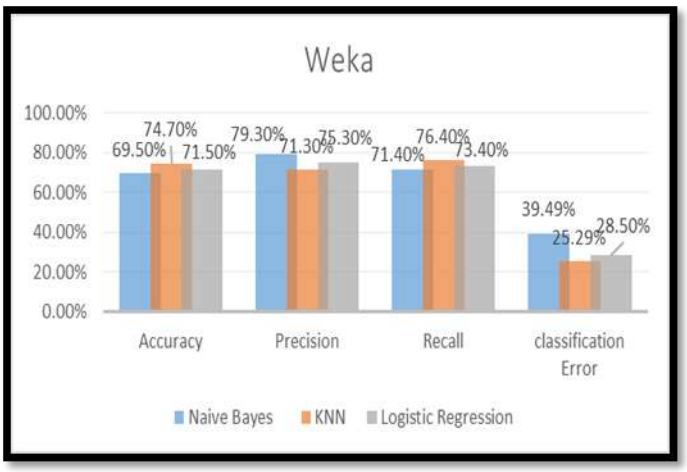


Figure 3: Weka Finding

6. CONCLUSIONS

In this paper SARS, patient data collection used and get data from the sars_2003_complete _dataset clean Kaggle Data Set. It has 2538 rows with 5 attributes and 1 class variable. There was a concert of accuracy classified models result attached, which is as clear as the tracks in Figure 4.

Figure 4: Final Accuracy

From the above in Fig it can be deduced that the WEKA tool measures the low accuracy of the Naive Bayes, but with the same process in the Rapid Miner application system it measures better accuracy when associated with WEKA. When calculating KNN the Rapid Miner tool showed lower accuracy and the orange application showed better accuracy compared to the previous one. Ultimately usually the Rapid Miner application predicts advanced accuracy forms for the whole three partitioning process. Above Fig [4] also proves that the accuracy of Rapid Miner and Naïve Bayes is well correlated with Logistic Regression.

REFERENCES

[1] https://www.kaggle.com/imdevskp/sars-outbreak-2003-complete-dataset
[2] https://rapidminer.com/
[3] https://sourceforge.net/projects/weka/
[4] https://orange.biolab.si/

SHAXS IMZOLARI TO'PLAMI ORASIDAN IMZOLARNI TAHLIL QILISH ASOSIDA SHAXSNI ANGLASH ALGORITMI

Bekmuratov D.Q., Fayziyev V., Ahrorov M.Sh., ***Muhammad al-Xorazmiy nomidagi TATU Samarqand filiali, O'zbekiston.***

Annotatsiya: *Ushbu maqolada etalon tanlovdagi shaxslarning yangi imzolari, etalon tanlovdagi har bir shaxsning imzolari bilan solishtirilgan va yangi imzolar etalon tanlovdagi qaysi shaxsning imzolariga ko'proq o'xshashligiga qarab, qaysi shaxsga tegishli ekanligini aniqlaydigan algoritmini ishlab chiqishga qaratilgan.*

Аннотация: *В этой статье разработан алгоритм, который сопоставляет новые подписи лиц введенных эталон, сравнивает с подписью лиц ранее введенные в эталлон и исходя из этого выявляет какому лицу из них пренадлежит этот подпись.*

Annotation: *This article has developed an algorithm that compares new signatures of persons entered into the standard, compares with the signatures of persons previously entered into the standard and, on the basis of this, identifies to which of them this signature belongs.*

Kalit so'zlar: *etalon tanlov, imzo, obyekt, belgi, shaxsning imzosi, sinflash, avtomatik sinflash, yangi imzo, ajratuvchi tizimlar, imzolar nabori, belgilar alfaviti, namunaviy imzo, hal qiluvchi qoida.*

Ключевые слова*: Эталонная выборка, подпись, объект, символ, подпись человека, классификация, автоматическая классификация, новая подпись, системы разделения, набор подписей, алфавит символов, образец подписи, решающее правило.*

Keywords*: Reference sampling, signature, object, symbol, human signature, classification, automatic classification, new signature, separating classes, new object, alphabet of characters, signature sample, decision rule.*

Kirish. Imzolarni avtomatik qayta ishlash va ularni aniqlash yo'nalishidagi masalalar dunyo miqiyosidagi dolzarb muammolardan biri bo'lib, bu yo'nalishda ko'plab ishlar amalga oshirilmoqda hamda usullar va algoritmlar yaratilmoqda. Bugungi kunda biz raqamlar dunyosida yashaymiz. Shunday imzoni anglash algoritmlari va qurilmalari mavjudki, bu algoritmlar va qurilmalar insonning bevosita jarayonda qatnashishini talab etmasdan, uning imzosi orqali shaxsini aniqlab olish mumkin. Buning uchun inson hujjatga imzo qo'yayotganda imzosiga boshqa shaxslar tomonidan soxtalashtirib bo'lmaydigan qandaydir belgini qo'yishi va ushbu imzoni turli algoritmlar va qurilmalar orqali kompyuterga elektron variantda uzatishi mumkin bo'lsin. Yuqorida keltirilganlarni hisobga olsak, o'zimizning an'anaviy qo'lda imzo qo'yish usuliga qaytish samarali hisoblanadi, agarda faqat imzolarni yozish(imzoni qo'yish)ni emas, balki uni yozish(imzoni qo'yish) dinamikasini ham hisobga oluvchi maxsus algoritmlar va qurilmalar yaratilsa albatta.

Masalaning qo'yilishi. Shaxs imzolari to'plami berilganda, imzolarni ketma-ket tekshirish jarayonida sinflarga ajratish va hosil bo'lgan sinflar sonini minimallashtirish algoritmlarini hamda dasturiy ta'minotlarini yaratish masalalari qo'yilgan bo'lib, xususan, kompyuterda shaxs imzolari qo'yiladi(chiziladi) yoki skaner orqali kiritiladi va o'xshash imzolar alohida sinflarga birlashtiriladi. O'xshash imzolarni o'z ichiga oluvchi sinflar hosil bo'ladi va hosil bo'lgan sinflar to'plami optimallashtiriladi. Yangi imzolar optimallashtirilgan sinflar to'plamidagi har bir sinfda joylashgan imzolar bilan solishtiriladi va qaysi sinfdagi imzolarga ko'proq o'xshasa, o'sha sinfga qarashli bo'ladi. Demak, imzo tasvirini tahlil qilish orqali shaxsni anglashning avtomatlashgan tizimini yaratish xozirgi kunning eng dolzarb muammosidir.

Ishning maqsadi. Sinflar soni oldindan ma'lum bo'lganda imzolarni ketma-ket tekshirish jarayonida imzolar to'plamini $K_1, K_2, ..., K_l$ sinflarga ajratish va hosil qilingan sinflar sonini minimallashtirish algoritmlari va dasturiy ta'minotlarini yaratish maqsadi qo'yilgan. Bu maqsadga erishish uchun quyidagi masalalarni yechish talab etiladi:

-imzolarni sinflarga ajratuvchi usullar va algoritmlarni tahlil qilish;

-imzolarni kompyuterda chizish algoritmini qurish;

-imzolarni kompyuterda kodlashtirish algoritmini ishlab chiqish;

-imzolarni ketma-ket tekshirish jarayonida ularni sinflarga ajratuvchi algoritmni ishlab chiqish;

-yaratilgan algoritmlar asosida dasturiy ta'minotni ishlab chiqish.

Masalani yechish metodi. Ishda binar kodlashtirish, kompuyterli ko'rish, timsollarni sinflarga ajratish va matematik statistika usullaridan foydalanilgan.

Masalaning yechilishi: Shaxs imzolari to'plamini imzo(yadro)larni ixtiyoriy tanlash asosida sinflash algoritmi: Ushbu ishda $S = S_1, S_2, ..., S_m$ imzolar to'plami etalon tanlov, ya'ni T_{nml}(n-belgilar, m-imzolar va l-sinflar soni) shaklda berilgan holda qaralgan..

$S = S_1, S_2, ..., S_m$ imzolar to'plami T_{nml}(n-belgilar, m-imzolar va l-sinflar soni) etalon tanlov shaklda berilmaganda, ya'ni bizga faqat $S = S_1, S_2, ..., S_m$ imzolar to'plami berilgan va ular oldindan sinflarga ajratilmagan(soha mutaxassisi tomonidan sinflarga oldindan taqsimlanmagan) holda $S = S_1, S_2, ..., S_m$ imzolar to'plami o'zini-o'zi o'rganish jarayonida $K_1, K_2, ..., K_l$ sinflarga ajratish masalasi qaraladi. Yuqorida keltirilganlardan kelib chiqib, masalaning qo'yilishi va ishning maqsadini quyidagicha shakllantiramiz.

Bizga $S = S_1, S_2, ..., S_m$ imzolar to'plami(tanlovi) berilgan bo'lsin. S tanlovdagi $S_1, S_2, ..., S_m$ imzolarni o'qituvchisiz va o'zini-o'zi o'rganish jarayonida $K_1, K_2, ..., K_l$ sinflarga ajratish talab etiladi. Imzolarni $K_1, K_2, ..., K_l$ sinflarga ajratishda sinflar soni l oldindan ma'lum bo'lishi yoki ma'lum bo'lmasligi mumkin. Agar sinflar soni l oldindan ma'lum bo'lmasa, u holda l, imzolar to'plamini sinflash jarayonida hosil bo'ladi.

Masalaning qo'yilishi va maqsadidan kelib chiqadigan bo'lsak, ushbu ishda quyidagi masalalar yechiladi:

-shaxs imzolari to'plamini skaner orqali kompuyterga kiritish algoritmini ishlab chiqish;

-shaxs imzolarini kompyuterda chizish algoritmini ishlab chiqish;

-shaxs imzolarini kompyuterda kodlashtirish algoritmini ishlab chiqish;

Masalani yechish algoritmi quyidagi qadamlardan iborat:

1. Boshlang`ich ma'lumotlar: shaxs imzolari soni, imzolar tushiriladigan retseptor to'rining o'lchami $a*b$ kiritiladi.

2. Imzolar kiritiladi:

2.1. Shaxs nomi va imzolari skaner yordamida.

2.2. Shaxs nomi va imzolari sichqoncha yordamida.

3. $j=1.$ Imzolar kodlashtiriladi:

$$S_j=\begin{cases}1, agar\ imzoning\ qismi\ a*b\ to'rning\ a_\lambda * b_\gamma\ (\lambda,\gamma=\overline{1,t})\ katakchasiga\ tushsa\\ o, aks\ holda\end{cases}$$

4. S_j imzoning $a*b$ retseptor to'ri bir o'lchamli $S_j=\alpha_{j1},\alpha_{j2},..,\alpha_{jn},(n=a*b)$ ko'rinishga keltiriladi.

5. $j=j+1.$ Agar $j\le m$ bo'lsa, u holda algoritm 3-qadamga o'tadi, aks holda 6-qadamga o'tadi.

6. Kodlashtirilgan $S_1,S_1,...,S_m$ imzolar ketma-ket T_{nm} tanlovga joylashtiriladi, bu erda $n=a*b$ - imzo belgilari soni, m - imzolar soni.

7. T_{nm} tanlovdagi $S_1,S_1,...,S_m$ imzolar orasidagi masofalar $d(S_i,S_j)=\sum_{k=1}^{m}\left|\alpha_{ik}-\alpha_{jk}\right|$ yordamida aniqlanadi. Natijada d_{ij} $(i=1,n;\ j=1,n)$ masofalar matrisasi hosil bo'ladi

$$d_{ij}=\begin{pmatrix}d_{11} & d_{12} & . & . & . & d_{1n}\\ d_{21} & d_{22} & . & . & . & d_{2n}\\ . & . & . & . & . & .\\ d_{n1} & d_{n2} & . & . & . & d_{nn}\end{pmatrix},\quad i=1,n;\ j=1,n.$$

8. d_{ij} matrisadan eng katta masofa $d_{\max}(S_i,S_j)=\max_{ij} d_{ij}$ topiladi. $d_{\max}(S_i,S_j)$ qiymatga mos keluvchi S_i va S_j imzolar boshlang'ich yadrolar sifatida tanlanadi.

9. Yadrolar $Z_1=S_i$ va $Z_2=S_j$ belgilanadi.

10. T_{nm} tanlovdagi $S_1,S_1,...,S_m$ shaxs imzolaridan $Z_1=S_i$ va $Z_2=S_j$ yadrolardan tashqari $Z_3,Z_4,...,Z_k\ (k<m)$ yadrolar ixtiyoriy ravishda tanlanadi.

11. $i=1.$ $S_1,S_1,...,S_m$ shaxs imzolardan tanlangan $Z_1,Z_2,...,Z_k\ (k<m)$ yadrolar orasidan Z_i yadro olinadi.

12. Z_i yadroga $d^i_{\min}$ asosida eng yaqin Z_j yadro orasidagi masofaning yarmi $\bar{d}_{ij}=\frac{1}{2}d_{ij}(Z_i,Z_j)$ topiladi.

13 $S_1, S_1, ..., S_m$ imzolar orasidan Z_i yadroga nisbatan hosil qilinadigan sinfga qarashliligi

$$F(S_i) = \begin{cases} S_i \in Z_i & agar \quad d_{ij} \le \bar{d}_{ij} \\ S_i \notin Z_i & boshqa\ hollarda \end{cases}$$

aniqlanadi.

14. $i = i + 1$. Agar $i \le k$ bo'lsa, u holda algoritm 11-qadamga o'tadi, aks holda 13-qadamga o'tadi.

15. Natijada $S_1, S_1, ..., S_m$ imzolar orasidan ixtiyoriy tanlangan $Z_1, Z_2, ..., Z_k (k < m)$ yadrolarga nisbatan sinflar

$$K_1 : Z_1 = \{S_1^1, S_2^1, ..., S_{m_1}^1\}$$
$$K_2 : Z_2 = \{S_1^2, S_2^2, ..., S_{m_2}^2\}$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$K_k : Z_k = \{S_1^k, S_2^k, ..., S_{m_k}^k\}$$

hosil bo'ladi. Bu erda $(m_1 + m_2 + ... + m_k) = m$.

Imzolarni qayta ishlash va ularni sinflash bilan bog'liq yaratilgan algoritmlar turli masalalarda qo'llanilishi mumkin. Xususan, kriminalistik tekshirishda, korxona, banklar va turli tashkilotlarda shaxs imzolari bazasini yaratishda uni yangilashda, imzo shaxsini ma'lumotlar bazasidan izlab topish kabi masalalarda qo'llanilishi mumkin. Bundan ko'rinib turibdiki, bajarilgan ishlar xayotdagi muhim vazifalarni tezkor hal etish imkonini beradi.

Adabiyotlar:

1. Q.A.Bekmuratov, O.A.Mamaraufov, D.Q.Bekmuratov. (2015) Timsollarni anglovchi tizimlar. Uslubiyuv qo'llanma. Firdavs Nasimov xususiy korxonasi bosmaxonasi. Samarqand. 390 bet.

2. Бекмуратов К.А., Бекмуратов Д.К. (2015). Последовательный выбор признаков, обладающих требуемой разделяющей силой. XI - Международная научно-практическая конференция. "Научные перспективы XXI века. Достижения и перспективы нового столетия". Ежемесячный научный журнал №4(11) / 2015, часть 4. Россия, г. Новосибирск, 22-23.05, 9-13 с. ISSN 34567-1769. 18-19 май 2015 г. Новосибирск. 5 стр.

3. Бекмуратов Д.К., Умаров Ф.Ш. (2016). Процедуры формирования пространства признаков k-го типа. "International Scientific and Practical Conference "WORLD SCIENCE" №7(11),vol.1, July 2016 Proceedings of the II International Scientific and Practical Conference "Topical researches of the World Science. June 29-30, 2016, Dubai UAE".(ISSN 2413-1032; ISSN 2414-6404 online).

4. Бекмуратов Д.К. (2017). Разработка алгоритм формирование системы опорных множеств признаков, обеспечивающих качество и надежность распознавания. «Проблемы вычеслительной и прикладной математики».".

Научный журнал г, часть 5, №5, 2017г.ТУИТ, Ташкент. (ISSN: 2181-8460). 74-79ст.

5. Д.Қ.Бекмуратов. (2015). Объектларни англашда ўргатувчи танловдан эталон объектларни ҳосил қилишнинг алгоритм ва дастурий таъминоти. Республика илмий-техник конференцияси. Маърузалар тўплами. Самарканд, 11-12-сентябрь, 2015 йил. ISBN 978-9943-4545-3-8. Самарқанд. 3 бет.

AYLANA TENGLAMASI YORDAMIDA AYLANA SHAKLIDAGI CUSTOM VIEW KOMPONENTASINI YARATISH

Jo'rayev M.D., Zahiriddin Muhmmad Bobur nomidagi Andijon Davlat Universiteti, O'zbekiston

Annotation: *When creating Android apps, you sometimes have to create a non-standard View component in addition to the standard View component. in this case the Custom View component is created. This article explores the creation of a Custom View component that displays images in a circle using the circle equation.*

Keywords: *view, custom view, circle, class, constructor, method, field, lifecycle, inheritance, attribute, layout.*

Аннотация: *При создании приложений Android вам иногда приходится создавать нестандартный компонент View в дополнение к стандартному компоненту View. в этом случае создается компонент Custom View. В этой статье рассматривается, как создать компонент пользовательского вида, который отображает изображения в круге с использованием уравнения круга.*

Ключевые слова: *вид, компонент пользовательского вида, круг, класс, конструктор, метод, поле, жизненный цикл, наследование, атрибут, макет.*

Annotatsiya: *Android ilovalarni yaratishda ba'zan standart View - komponentlardan tashqari nostandart ko'rinishdagi View - komponentasini yaratishga to'g'ri keladi. ushbu holatda Custom View - komponentasini yaratib olinadi. ushbu maqolada aylana tenglamasi yordamida rasmlarga aylana shaklida ko'rsatadigan Custom View - komponentasini yaratish tadqiq qilindi.*

Kalit so'zlar: *view, custom view, aylana, sinf, konstruktor, metod, maydon, hayotiy sikl, voris olish, atribut, layout.*

Andorid ilovalarini ishlab chiqish jarayonida standart View – komponentasidan tashqari dasturchi o'zi View – komponentasini qo'lda chizib, o'zinig uslubi va dizayniga asoslanga View – komponentasini yaratishga to'g'ri keladi. Shunin uchun dasturchi o'zining Custom View – komponentasini yaratadi.

Ko'pgina hollarda Custom View tarkibiy qismlarini ishlab chiqish oddiy View komponentlari bilan ishlashga qaraganda ko'proq vaqt talab etadi, shuning uchun ularni faqat biron bir xususiyatni amalga oshirishning sodda usuli bo'lmaganda yoki quyida sanab o'tilgan muammolar mavjud bo'lganda yaratish tavsiya etiladi. Custom View komponentasi quyidagi holatlarda qo'l keladi:

- Ishlashi. Agar sizda bitta layout – fayli bo'lib, unda ko'plab View – komponentlari bo'lsa va siz yagona Custom View – komponentasini yaratib uni optimallashtirmoqchi bo'lsangiz;
- Ishlashi va xizmat ko'rsati qiyin bo'lgan View – komponentalarni katta ierarxiyasida;
- Qo'lda chizishni talab qiladigan to'liq moslashtirilgan View komponenti.

Har bir View – komponenti konstruktor bilan boshlanadi. Va bu bizga uni tayyorlash, turli xil hisob-kitoblarni amalga oshirish, standart qiymatlarni sozlash, biz uchun zarur bo'lgan barcha imkoniyatlarni beradi. Custom View – komponentasini aylana shakliga keltirishni aynan konstruktor ichida amalga oshiramiz.

View – komponentalari o'zining sinfi (class) orqali aniqlanganligi uchun Custom View – komponentasi uchun MyCircleImageView sinfini yaratib olamiz. Custom View – komponenta sinfini uning ishlash imkoniyatidan kelib chiqib biror bir View - komponentasi sinfidan vorish olish talab etiladi. Bu yerda biz rasmlar bilan ishlash imkoniyati bo'lishi uchun AppCompatImageView sinfidan voris olamiz. So'ng ota sinfdagi zarur bo'lgan konstuktorlarni yaratamiz:

```
public class MyCircleImageView extends AppCompatImageView {

   public MyCircleImageView(Context context) {
      super(context);
   }

   public MyCircleImageView(Context context, AttributeSet attrs) {
      super(context, attrs);
   }

   public MyCircleImageView(Context context, AttributeSet attrs, int defStyleAttr)
{
      super(context, attrs, defStyleAttr);
   }
```

MyCircleImageView View – koponentasi attributlari uchun values ichida yangi fayl yaratib uni "attrs.xml" deb nomlanadi. Ushbu faylda bir nechta Custom View – komponentlarining atributlari bo'lishi mumkin. Faylning ichiga biz yaratmoqchi bo'lgan View – komponentasi nomi uchun MyCircleImageView va uning atributi nomiga *image*, qiymatiga "reference" ni yozamiz:

```
<declare-styleable name="MyCircleImageView">
   <attr name="image" format="reference"/>
</declare-styleable>
```

View – komponentasidagi kiritilgan atributlarini olib ularga ishlov berib aylana shaklidagi View ni hosil qilish uchun sinf ichida Drawable sinfidan ob'ekt olib *drawable* maydoni hamda attributlarni initsializatsiya qilish uchun *init(AttributeSet attrs, int defStyleAttr)* metod va Drawable toifasida kiritilgan rasmni aylana shakliga keltirib Drawable *o*b'ektini qaytaradigan *getCircleImage(Drawable drawable)* metodlarini yaratamiz.

View – komponentasiga qo'yilgan rasmni aylana shakliga keltirish uchun quyidagi aylana tenglamasidan foydalanamiz:

$$R = \sqrt{(x-a)^2 + (y-b)^2} \qquad (1)$$

View – komponentasiga qo'yilgan rasmni aylana shakliga keltirishning mohiyati shundaki, qo'yilgan rasmni aylana tenglamasi sohasiga tegishli

bo'lmagan maydondagi piksellarining ranglarini o'chirib chiqishdan iborat. Va uning kodi quyidagicha bo'ladi:

```
private Drawable getCircleImage(Drawable drawable) {

    Drawable circleImage = null;
    Bitmap bitmap = ((BitmapDrawable) drawable).getBitmap();
    Bitmap bitmap1 = bitmap.copy(bitmap.getConfig(), true);

    int a = bitmap1.getWidth() / 2;
    int b = bitmap1.getHeight() / 2;
    int r = Math.min(a, b);

    for (int i = 0; i < bitmap1.getWidth(); i++) {
        for (int j = 0; j < bitmap1.getHeight(); j++) {
            if (Math.pow(a - i, 2) + Math.pow(b - j, 2) > r * r) {
                bitmap1.setPixel(i, j, Color.TRANSPARENT);
            }
        }
    }

    circleImage = new BitmapDrawable(getResources(), bitmap1);
    return circleImage;
}
```

Aylana shakliga kelgan rasmni *init()* metodi ichida View – komponentasiga o'rnatish uchun kiritilgan atributlarni olib, ulardagi qiymatlarni tekshirib so'ng View – komponentasiga o'rnatiladi:

```
private void init(AttributeSet attrs, int defStyleAttr) {
    final TypedArray a = getContext().obtainStyledAttributes(
            attrs,
            R.styleable.MyCircleImageView,
            defStyleAttr,
            0
    );

    if (a.hasValue(R.styleable.MyCircleImageView_image)) {
        drawable = a.getDrawable(R.styleable.MyCircleImageView_image);
        setImageDrawable(getCircleImage(drawable));
    }
}
```

Oxirgi bosqichda *init()* metodini yuqorida qayta yuklangan ota sinf konstruktorlarida chaqirib qo'yiladi:

```
public MyCircleImageView(Context context) {
    super(context);
    init(null, 0);
```

```
}

public MyCircleImageView(Context context, AttributeSet attrs) {
    super(context, attrs);
    init(attrs, 0);
}

public MyCircleImageView(Context context, AttributeSet attrs, int defStyleAttr) {
    super(context, attrs, defStyleAttr);
    init(attrs, defStyleAttr);
}
```

Bunda, har bir konstruktorga berilgan qiymatni *init()* metodiga yuboriladi, agar konstruktorga qiymat berilmagan bo'lsa vaziyatdan kelib chiqib *null* yoki *0* qiymat beriladi. Shu yo'sinda Custom View – komponentasi foydalanishga tayyor bo'ldi. Endi undan foydalanish uchun *activity_main.xml* layout ichida biz yaratgan Custom View ya'ni MyCircleImageView ni chaqirib uning *image* atributiga rasm joylab tekshiramiz:

```
<com.example.mycustomview.MyCircleImageView
    android:layout_width="200dp"
    android:layout_height="200dp"
    app:image="@drawable/img"
    app:layout_constraintTop_toTopOf="parent"
    app:layout_constraintBottom_toBottomOf="parent"
    app:layout_constraintLeft_toLeftOf="parent"
    app:layout_constraintRight_toRightOf="parent"/>
```

Birinchi urinishda View ko'rinish bermaydi. Chunki biz yangi Custom View yaratdik va bunda tizimning fayllar bo'limida o'zgarishlar sodir bo'lgan. Shuning uchun Build bo'limiga kirib, Rebuild project ni bosib loyihani qayta qurish kerak. So'ng View – komponentasining ko'rinishi quyidagi bo'ladi:

1-rasm.

Ushbu yaratilgan Custom View – komponentasini dastur loyihasinig istalgan qismida foydalanib rasmlarni aylana shakliga keltirib foydalanish mumkin. Bundan tashqari ushbu View – komponentasi AppCompatImageView sinfidan voris olganligi uchun AppCompatImageView barcha xususiyatlari bizning Custom View – komponentasiga o'tgan va bu xususiyatlar unda to'liq ishlaydi.

Adabiyotlar

1. *Ноутон П., Шилдт Г. JAVA 2. Наиболее полное руководство СПб.: БХВ-Петербург, 2008.*
2. *Г. Н. Яковлев, Геометрия I Қисм. Тошкент – "Ўқитувчи". г-1981.*
3. *Custom View Components: https://developer.android.com/guide/topics/ui/custom-components*

Android Notes for Professionals: https://www.twirpx.com/file/2472440/

COMPARATIVE CLASSIFICATION ALGORITHM ANALYSIS ON VARIOUS DATASETS USING WEKA

Misbah Akbar, Student, Department of Computer science at RIPHAH international University Lahore, Pakistan.

Abstract. *Data mining is the imminent area of research for solving distinct problems, and in data mining one of the key issues is classification. This paper is based on two classification algorithms of the Weka interface NAIVE BAYES and LOGISTICS. It can be used to evaluate a multitude of datasets. Naive Bayes and Logistics' performance has been treated to select the best algorithm based on data sets accustom. The datasets are taken from Kaggle. In machine learning systems, naïve Bayes is used to measuring the latest data or check data, and is established on the principle of Bayes. Logistics regression is an algorithm for binary classification. The input variables are assumed to be numeric and they have got a 'Gaussian distribution'. It only supports problems with binary classification, with multi-class classification implementation of weka has been modifying to support problems.*

Keywords. *Classification, Naive Bates, Logistics, Data mining, weka tool*

1. INTRODUCTION

Data mining is the impending area of research in which we can solve different problems, and classification is one of the very calumniating problems of data mining. In this paper we deal with two Weka classification algorithms named 'NAIVE BAYES' and 'LOGISTICS'. It can be used to classify a multitude/myriad of datasets. The performance of 'Naive Bayes' and 'Logistics' can be analyzed to select the best algorithm that is based on the condition of data sets. The datasets are picked from Kaggle.

Naive Bayes is used in a machine learning system to measure/calculate new data or audit data, and this is placed on the principle of Bayes.

Logistics algorithm is an algorithm that is used for the binary classification. The input variables are numeric and they have a Gaussian distribution. To support problems with multi-class classification, weka implementation has been modified and only binary classification is supported.

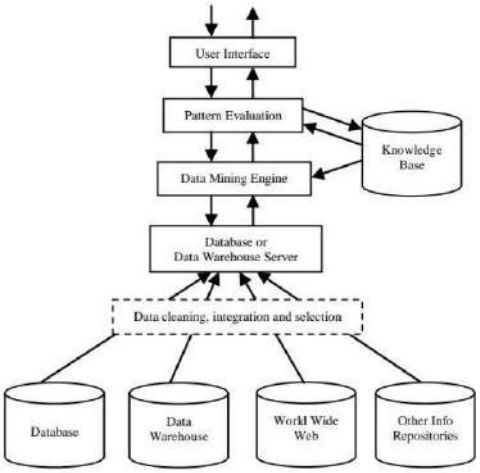


Fig 1: Architecture of a Typical Data Mining System

Related work:

On the performance and efficiency of the decision tree and the backpropagation, recent studies have been done. In machine learning and data processing, classification is a simple problem. Few researchers have developed data clustering algorithms and some introduced new algorithms, and some just examined and compared existing clustering algorithms [2]. Decision trees are much famous because they are realistic and forward-looking. From decision trees rules can be extracted easily. Many algorithms, just like Naive Bayes and Logistics are designed for decision making by using tree construction.

II. RESEARCH QUESTIONS

1. Which algorithm of machine learning give us better accuracy or performance using five different dataset?

2. Which algorithm gives better results either Naive Bayes or Logistics based on different clustering datasets?

III. METHODOLOGY

Datasets

We have five (5) datasets on which we are working and we take these datasets from Kaggle. Explanation of every dataset is in the table.

Table no.1: Explanation of datasets

Dataset	Instances	Attributes	Number of classes	Sort
Contact lens	24	4	3	Numeric
Credit	1000	20	2	Numeric
Soybean	683	35	2	Numeric
Super market	4627	8	2	Nominal
Weather	14	5	2	Numeric

The first dataset named ‘contact lens’ have 24 instances and three classes, its type is numeric and it has four attributes.

The second dataset named ‘credit’ contains 1 thousand (1000) instances, 20 attributes, and only two classes, its type is numerical and it is used for

Classification. The same methodology is used by this dataset as the prior dataset used. The attribute of the following dataset is inexpensive or commercial [3].

The 3rd dataset named 'Soybean' contains 683 instances and 35 attributes, have 2 classes and its type is numerical.

The 4th dataset named 'Supermarket' contains 4627 instances, 35 attributes, have 2 classes and its type is numerical.

The 5th dataset named 'weather' contains 14 instances and 5 attributes and has 2 classes and its type is numerical.

Weka interface

Weka stands for 'Waikato Environment for Information Analysis'. It can occur in Java machine learning software and developed at the University of Waikato in New Zealand [1]. Weka suite includes the real-time tools and algorithms for analyzing and predicting data, as well as gives management tool for accessing existing applications. Java Applications is an end-to-end TCL / TK software solution that uses algorithms routed in other programming languages, as well as C-level software and machine learning tools. These fields are used, especially for educational and research purposes. Following are the some advantages of weka:

- It is available for free concealed by the GNU General Public. Because of its implementation in the programming language of Java it is very scalable and it can almost run on any architecture.
- It can be an incredible range of modeling skills pre-processing.
- It is very easy to use because of its Graphical GUI.

Naive Bayes Classification algorithm

In machine learning, naïve Bayes classifiers are "probabilistic classifiers". It based on the Bayes theorem with vigorous assumptions of independence between the features. In the Bayesian network they are the simplest model. By estimating the Kernel density they can be combined and attain the advanced levels of precision.

Naïve Bayes algorithm has been universally researched since the 1960s. In the early 1960s, it was popularized in the text retrieval community and endure a simple (baseline) tool for text grouping, the controversy of defining documents is an association of one or another category, for example, spam or legitimate, sports or politics, etc. with the frequencies of the word as a feature. With more mature or practical approaches in this area like with sufficient pre-processing promote the vector machines. In automated medical diagnosis, it is also useful.

Logistics classification algorithm

Logistic algorithms are worked to identify binaries. In the logistics algorithm we can suppose that input variables are numeric and it has a Gaussian distribution that is bell-curved. Logistics algorithm can produce good results if the data don't have the Gaussian. Some input attributes have a Gaussian-like distribution in the case of the Ionosphere dataset but others don't. This algorithm determines a coefficient for every input value, which is reconstructed using a logistic (s-shaped) function and linearly connected into a regression function. Logistic regression can

be prosperous and successful to solve certain issues because it is a quick and easy technique. Only binary classification problems can be sustained by Logistic regression. Multi-class classification issues have been modified by the weka implementation.

Choose the 'Logistic regression algorithm':

Press the "Choose" button and pick "Logistic" under the division of "Functions."[1].

Demonstration of the configuration algorithm, click on the algorithm name.

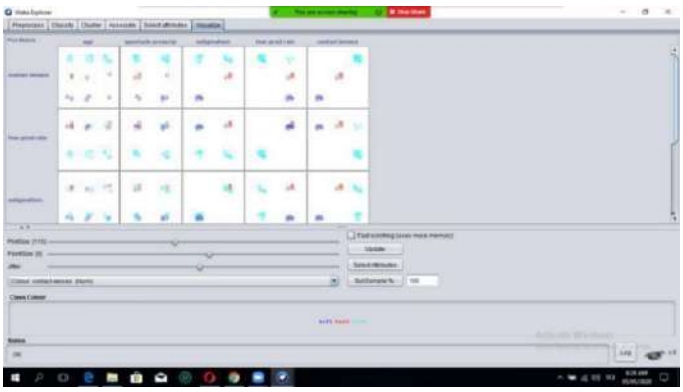
IV. RESULTS

In this chunk, I'll discuss the results which I got from all of the above-mentioned work. I picked five datasets named "Contact lens, Credit, Soybeans, Supermarket, weather". On these datasets, I applied two algorithms, First is 'NAVIE BAYES' and the second one is 'LOGISTICS'. The findings will be given/shown in the following tables:

Table 2. Accuracy of dataset 'Contact lens'

S.N.	PARAMETERS	NAÏVE BAYES	LOGISTICS
1	TP Rate	0.708	0.667
2	FP Rate	0.305	0.316
3	Precision	0.691	0.663
4	Recall	0.708	0.667
5	F-Measure	0.698	0.662
6	ROC Area	0.870	0.781

Figure 2: Accuracy chart of dataset 'Contact lens'



The accuracy parameters of the dataset 'Contact lens' has displayed in Table 2 and Figure 2. The above chart shows that all parameter of algorithm 'Naive Bayes' has high accuracy as compared to 'Logistics' except of one parameter that is 'FP rate'. So, we can say that the Naïve Bayes algorithm is a better algorithm for the dataset 'Contact lens'.

Table 3. Accuracy of dataset 'Credit'

S.N.	PARAMETERS	NAÏVE BAYES	LOGISTICS
1	TP Rate	0.754	0.752
2	FP Rate	0.393	0.398
3	Precision	0.743	0.741
4	Recall	0.745	0.752
5	F-Measure	0.746	0.744
6	ROC Area	0.787	0.785

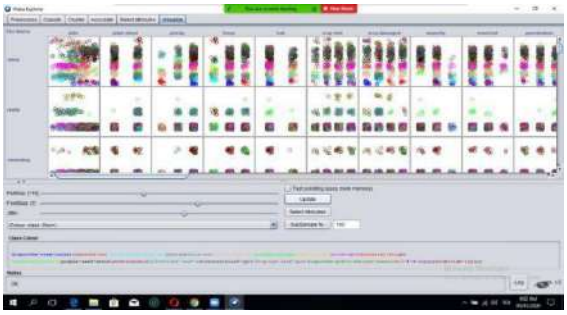
Figure 3: Accuracy chart of dataset 'Credit'

The accuracy parameters of the dataset 'Credit' has displayed in Table 3 and Figure 3. The above chart shows that it has almost equal accuracy measures except the 'ROC Area' measure in which logistics has a higher accuracy rate as compare to the Navie Bays algorithm. So, we can say that the logistics algorithm is a better algorithm for dataset 'credit'.

Table 4. Accuracy of dataset 'Soybean'

S.N.	PARAMETERS	NAÏVE BAYES	LOGISTICS
1	TP Rate	0.930	0.977
2	FP Rate	0.009	0.002
3	Precision	0.938	0.997
4	Recall	0.930	0.977
5	F-Measure	0.929	0.977
6	ROC Area	0.994	0.1000

Figure 4: Accuracy chart of dataset 'Soybean'

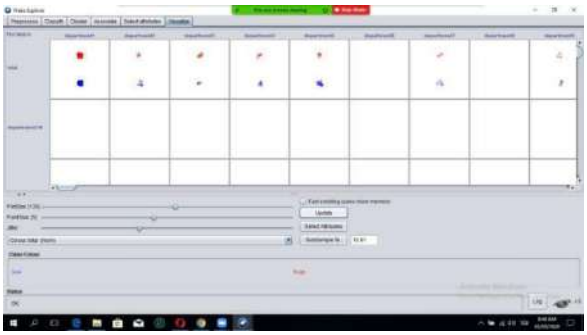


The accuracy parameters of the dataset 'Soybean' has displayed in Table 4 and Figure 4. The above chart shows that the Logistics algorithm has higher accuracy results as compared to the Naive Bayes algorithm except for the 'FP rate' parameter, in which Native Bayes has a higher accuracy. So, we can say that the logistics algorithm is a better algorithm for dataset 'soybean'.

Table 5. Accuracy of dataset 'Super Market'

S.N.	PARAMETERS	NAÏVE BAYES	LOGISTICS
1	TP Rate	0.637	0.637
2	FP Rate	0.637	0.637
3	Precision	0.637	0.637
4	Recall	0.637	0.637
5	F-Measure	0.778	0.778
6	ROC Area	0.499	0.499

Figure 5: Accuracy chart of dataset 'Super Market'

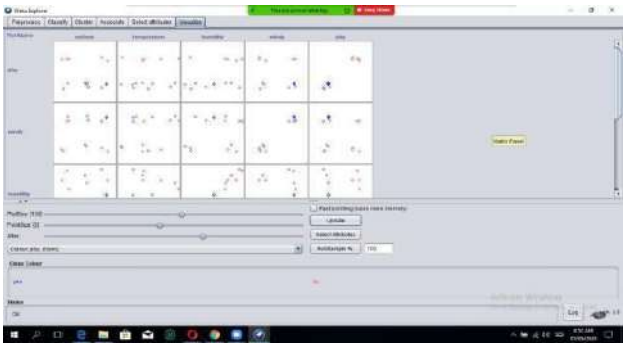


The accuracy parameters of the dataset 'SuperMarket' has displayed in Table 5 and Figure 5. The above chart shows that the Logistics and Naive Bayes algorithm, both have same accuracy measures in all parameters.

Table 6. Accuracy of dataset 'Weather'

S.N.	PARAMETERS	NAÏVE BAYES	LOGISTICS
1	TP Rate	0.643	0.571
2	FP Rate	0.554	0.505
3	Precision	0.607	0.571
4	Recall	0.643	0.571
5	F-Measure	0.592	0.571
6	ROC Area	0.444	0.351

Figure 6: Accuracy chart of dataset 'Weather'



The accuracy parameters of the dataset 'Weather' has displayed in Table 6 and Figure 6. The above chart shows that all parameter of the algorithm 'Naive Bayes' has high accuracy as compared to 'Logistics'. So, we can say that the Naïve Bayes algorithm is a better algorithm for the dataset 'Weather'.

V. CONCLUSION

We use different accuracy measurement parameters just like " TP rate, FP rate, Precision, Recall, F-measure and ROC area " in this paper. By using the Naive Bayes and Logistics algorithms we appraise the performance in terms of 'Classification'. The precision of each dataset is measured or calculated.

In the dataset 'Contact lens' Naïve Bayes gives better accuracy as compared to Logistics, but the FP rate of logistics is more accurate. In Datasets 'Credit' and 'Super Market' both algorithms Naïve Bayes and Logistics giving the same results. But in the dataset 'credit' the accuracy of parameter FP rate is high in the Logistics algorithm.

In the dataset 'Soybean' Logistics is giving us more accurate and good results.

In the dataset 'weather' Naïve Bayes is giving us more accurate results.

VI. FUTURE WORK

Many classification algorithms can be implemented in the future and many other data sets can also be taken. We also get the industry's real dataset and take the actual impact of algorithm results.

In extension, the level of research concerning the number of attributes and the number of instances can be taken into account for the performance and achievement of the algorithm.

REFERENCES

[1] R. Arora, "Comparative analysis of classification algorithms on different datasets using WEKA," International Journal of Computer Applications, vol. 54, no. 13, 2012.

[2] P. Singh and A. Surya, "Performance analysis of clustering algorithms in data mining in weka," International Journal of advances in engineering & technology, vol. 7, no. 6, p. 1866,2015.

[3] R. Duriqi, V. Raca, and B. Cico, "Comparative analysis of classification algorithms on three different datasets using WEKA," in 2016 5th Mediterranean Conference on Embedded Computing (MECO), 2016: IEEE, pp.335-338.

АВТОМАТИЗАЦИЯ ЗАМЕРА И РЕГИСТРАЦИИ ГИДРОГЕОЛОГИЧЕСКИХ ПАРАМЕТРОВ ПОДЗЕМНОЙ ГИДРОСФЕРЫ

Джуманов Ж.Х., Сайфуллаева Н.А., Ишанходжае О.А., Ёсинов С.Д., Ирмухамедова Н., Ташкентский университет информационных технологий, Узбекистан.

***Аннотация**. Мақолада ер ости суви мониторингини юритилиши ва ахборот таъминотида ер ости гидросферасининг гидрогеологик тавсилотларини ўлчашда автоматлаштирилган қурилмаларнинг қўлланилиши масалалари ҳақида фикр юритилади.*

***Калит сўзлар**: автоматлаштирилган ўлчов қурилмалар, ахборот таъминоти, сув ресурслари, гидрогеологик тавсилотлар.*

***Аннотация**. В статье рассматриваются вопросы применения устройств автоматизированного замера гидрогеологических параметров подземной гидросферы для ведения мониторинга подземных вод и информационное обеспечение.*

***Ключеные слова**: устройства автоматического замера, информационное обеспечение, водные ресурсы, гидрогеологические параметры.*

***Annotation**. In the article are considered problems of automated measuring instruments of hydrogeological parameters of underground hydrosphere for information ensuring in the conducting of groundwater monitoring.*

***Key words**: automation measuring instruments, information support, water resources, hydrogeological parameters.*

Введение. Основной функцией мониторинга подземных вод (ПВ) является регулярное выполнение с определенной периодичностью наблюдений за уровнем, температурой и регистрации в скважинах применяются гидрогеологические рулетки, электроуровнемеры и др. Выполнение этих наблюдений традиционными методами связано с большими трудозатратами, субъективными и объективными ошибками, что естественно экономически не эффективно. Для выполнения работ по наблюдению на гидрогеологических станциях, ручной трудоемкий процесс замера необходимо заменить датчиками, т.е. автоматизировать процесс мониторинга ПВ [1].

При изучении сложных гидрогеологических процессов методом математических моделей основными исходными данными являются: данные мониторинга подземных вод (режимных наблюдений); результаты геологоразведочных работ (на перспективных участках); материалы опыта эксплуатации (на участках действующих водозаборов); что позволяет определить основные факторы формирования, структуру потока и т.д. Состав показателей, расположение наблюдательных пунктов и периодичность

замеров должны быть достаточны для отражения и объяснения изучаемых процессов. Данная работа посвящена вышеописанным вопросам, т.е. изучению водных ресурсов, в частности к разработке устройствам связи автоматизированного измерения уровня (АИУ) и регистрации параметров подземной гидросферы.

Основная часть. Автоматизированные устройства связи с гидрогеологическим объектом – это комплекс устройств и математическо-программного обеспечения, осуществляющих измерение и регистрацию гидрогеологических режимообразующих параметров объекта т.е., взаимодействие гидрогеологических объектов с ЭВМ.

Задача разработки и создание опытного образца устройства АИУ представляется в виде последовательности логических действий: регистрация – накопление – обработка – представление результатов. Каждый этап включает в себя множество операций, позволяющие существенно упростить измерительный процесс. Основное предназначение приборов взаимо-действие с гидрогеологическими объектами реального времени в темпе процессов, протекающих в этих объектах. На основе схемы устройства предусматривается применение принципиально новых сборочных единиц и методов работы. Конструирование устройства было связано с проведением опытно-конструкторских и научно-исследовательских работ всех частей устройства и их взаимодействия в заданных параметрах.

Устройства АИУ состоят из следующих основных электронных, информационно-технических частей (рис. 1.), в которую входят: измерительая часть (I), определяющая физическое явление и регистрирующая импульсы в аналоговый сигнал. Часть устройства связи (II), аналогового - цифровые преобразователи считывания, прием-передачи связей и вычислительная т.е. функциональная часть (III), программно- математическое обеспечение, сбор, анализ и отчет к которой относятся взаимосвязанные программы, автоматизирующие функции управления, и устанавливающие взаимодействие с компьютерами [2].

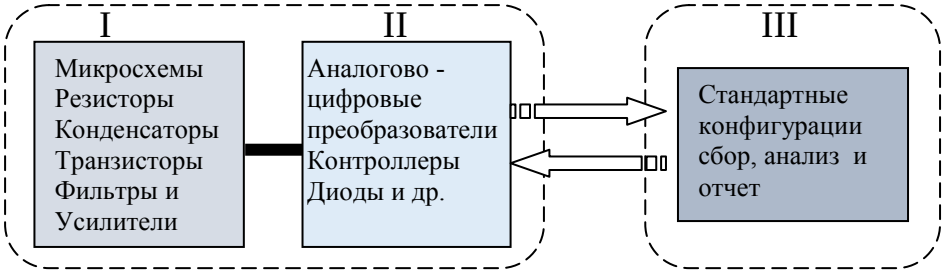


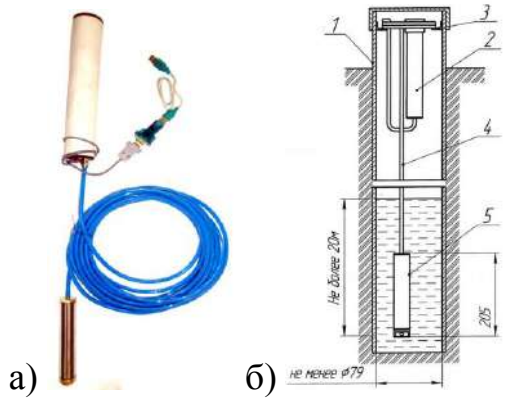
Рис. 1. Структура прибора АИУ

По требованию к конструктивному исполнению подготовлен испытательный и измерительный стенд системы мониторинга. Новизна и сложность разработок определялись по критериям научных и конструкторских работ всех этапов разработки, согласования и утверждения технической документации, изготовление, испытание опытных образцов по

«Типовым нормам времени на разработку конструкторской документации» и основополагающими документами Узгосстандарта последних лет.

Состав и структуры АИУ состоят из следующих основных частей гермоотсека с блоком электроникой и блоком электропитания; блока датчиков уровня и температуры; специального кабеля для соединения электронного блока с датчиками для установки АИУ в скважине (рис.2).

Метрологические характеристики соответствуют приводимым значениям -глубинность измерения уровня, в м, -100; диапазон измерений уровня, в м, от 0 до 20; предел допускаемой основной погрешности измерений уровня, в %, ±2,5; диапазон измерений температуры, в °С, от 5 до 50; ПО датчиков на компьютере должно соответствовать настоящее время программам операционных систем [2]. Выходной сигнал АИУ-1: интерфейс USB или COM4; Габаритные размеры: длина 27 см, диаметр и масса АИУ с кабелями 5 кг, должны позволять транспортировку автомобилем; питание АИУ-1 осуществляется от блока электрических батарей напряжением 4,5В; время непрерывной работы АИУ не менее 365 суток; климатическое исполнение АИУ, У5 по ГОСТ 15 150-69 при температуре окружающей среды (рабочие условия применения) от минус 10 до 50°С, а также при относительной влажности воздуха в скважине 100% при температуре 35°С и более низких температурах, с конденсацией влаги; защита от воздействия пыли и воды АИУ по ГОСТ 17185-72; средний срок службы АИУ до списания - 6 лет. Информационное обеспечение составляет содержательную основу, хранящуюся в базе данных не только для мониторинга, а также для её последующего анализа, обработки, оценки, многоцелевого поиска и выдачи. Данные собираются как из наблюдательных сетей мониторинга, так и из отдельных скважин. Поступающая информация из АИУ должна быть унифицирована, т.е. приведена в вид, удобный для её дальнейшего использования в базе данных. Это важный вопрос, особенно при создании разветвлённых локальных сетей мониторинга [3].



а) б)

Рис. 2. Устройство автоматизированного измерения уровня температуры подземных вод:

а - общий вид прибора; б - схема установки в наблюдательную скважину.

В базе данных ГИС информация хранится в виде атрибутивных данных, соответствующих информационным формам, каждая из которых выступает в качестве определенного свойства объекта. Для определения текущей

ситуации на наблюдательном пункте целесообразно собрать и анализировать необходимую совокупность данных, позволяющих оценить степень влияния отдельных показателей. Основными конструктивными элементами информационных моделей являются сущности, связи между объектами и их параметры, свойства – атрибуты, которые вводятся в «Базу данных» ГИС для дальнейшей обработки. Для ведения государственного кадастра подземных вод и автоматизированного мониторинга разрабатывается база данных ГИС.

В заключение хотелось бы отметить эффективность системы АИУ, действительно, эффективность в явном виде просчитать достаточно трудно, анализ непрост, или невозможен. Простой пример - использование режимных станций системы мониторинга в процессе строительства скважин. Деньги потрачены, внедрение проведено, все прекрасно функционирует, но полученная в процессе мониторинга информация денег не приносит, ничего вещественного мониторинг не производит, и получается, что автоматизации дают результатов; первый- улучшение производственной дисциплины на объекте мониторинга. Второе-сокращение времени принятия решений за счет оперативной информации независимо от удаленности объекта. Третье - улучшение качества принятых решений за счет удаленной работы высококвалифицированных специалистов. Четвертое - снижение затрат инженерного персонала за счет автоматизации работы с информацией. Пятое-полная сохранность и доступность всем специалистам информации, формируемой в базе данных и т. др.

Заключение. Высокая эффективность процесса сбора данных в моделировании гидрогеологических условий на 80 % эффективнее традиционных методов, более высокая точность данных благодаря расширению и использованию проверенных и надежных технологий. Вклад в системе мониторинга, а также в снижение риска здоровья наблюдателя при замера вреднее веществ в пункте наблюдений или аварийности скважин, значительно высок. Внедрение автоматизация измерений геологических процессов ведет, прежде всего, к повышению эффективности работы мониторинговых службах в целом, в составе информационно-коммуникационной технологии в гидрогеологии, образуются основу для формирования инфраструктуры пространственных данных РУз.

Литературы

1. *Джумано в Ж.Х. Применение автоматизированных измерительных датчиков для ведения мониторинга подземных вод Республики Узбекистан// Геология и минеральные ресурсы. -2011. №4. С.23-28.*
2. *Усманов Р.Н., Джуманов Ж.Х., ДигаевА.Х., Сеитназаров К.К., Отениязов Р.И. Программный комплекс для устройства автоматизированного измерения уровня и температуры подземных вод в гидрогеологических скважинах/ Свидетельство о депонировании объектов*

интеллектуальной собственности Рег.№ 1764 от 06.03.2014г до 06.03.2014г. г.Ташкент.

Умаров У.У. Автоматизированная информационно-поисковая система «Мелиоративная гидрогеология» и постоянно действующие модели. –Т.: «Фан», 1978. 120 с.

Методы организации компьютерных сетей на основе распределенных геоинформационных систем

Ишанходжаев О.А., Жамолов Х.М., Омонбоев О.А., Ташкентский университет информационных технологий, Узбекистан.

Аннотация. *Мақолада моделлаштириш, техник хизмат кўрсатиш ва ахборот таъминоти харажатларини камайтириш мақсадида созланадиган объект моделларининг ақлли тизимларини яратиш учун тақсимланган географик ахборот тизимлари асосида компьютер тармоқларини ташкил қилиш усуллари муҳокама қилинади.*

Калит сўзлар: *географик ахборот тизимлари, ахборот таъминоти, компьютер алоқа тармоқлари, вазиятларни таҳлил қилиш ва синтез қилиш.*

Аннотация. *В статье рассматриваются вопросы методы организации компьютерных сетей на основе распределенных геоинформационных систем для создания интеллектуальных систем настраиваемой объектной модели в целях снижения затрат на моделирование, сопровождение и информационное обеспечение.*

Ключеные слова: *геоинформационные системы, информационное обеспечение, компьютерных сетей связи, анализа и синтеза ситуаций.*

Annotation. *The article discusses the methods of organizing computer networks based on distributed geographic information systems to create intelligent systems of customizable object models in order to reduce the cost of modeling, maintenance and information support.*

Key words: *geographic information systems, information support, computer communication networks, analysis and synthesis of situations*.

Эффективность использования геоинформационных систем (ГИС) в различных сферах трудовой деятельности человека определяется прежде всего тем фактом, что 85% информации, с которой сталкивается человек в своей жизни, имеет территориально- распределенного характера. ГИС технологии активно внедряются в различных областях управления, информационно-коммуникационных технологии, промышленности, транспорта, геологии, экологии, здравоохранения и т.д.

В настоящее время ГИС все чаще начинают применяться для решения задач проектирование компьютерных сетей и связи, мониторинга и моделирования процессов, а также анализа и синтеза ситуаций, что позволяет говорить о новом классе интеллектуальных геоинформационных систем. Однако, при разработке и использовании таких систем, особенно предназначенных для работы в распределенной среде, неизбежно возникает ряд проблем, унаследованных от традиционных топологически-ориентированных и объектно-ориентированных подходов к построению ГИС.

Целью данной работы является разработка методики организации компьютерных сетей и создании интеллектуальных геоинформационных систем на основе настраиваемой объектной модели предметной области в целях снижения затрат на моделирование, создание и сопровождение ГИС.

Интеграция ГИС технологий для использования и анализа развития территориальных систем компьютерных сетей является перспективной при комплексном прогнозировании развития регионов, что представляет собой довольно представительные групп информационно-ресурсного характера.

В управление данными ГИС хранят пространственные и атрибутивные данные для их дальнейшего анализа и обработки, а в *запросах и анализа данных ГИС* выполняют запросы о свойствах объектов, расположенных на схеме или карте, и автоматизируют процесс сложного анализа, сопоставляя множество параметров или устройств сети для получения сведений или прогнозирования явлений.

*В визуализации данных у*добное представление ГИС данных непосредственно влияет на качество и скорость их анализа. Пространственные данные в геобазыданных предстают в виде интерактивных схем. Информации о состоянии объектов могут быть построены в виде графиков, диаграмм, трехмерных изображений.

Возможности ГИС работать как с векторными, так и с растровыми моделями данных, эффективны при решении различных задач, касающихся адекватного моделирования информации пространственного характера. ГИС тесно связаны с другими информационными системами и используют их данные для сбора, систематизации, обработки и анализа.

Разработка распределенных геоинформационных систем и распределенных баз данных, отражающих организационные структуры предприятий, позволяет сделать данные, поддерживаемые каждым из существующих подразделений, общедоступными, обеспечив при этом их сохранение именно в тех местах, где они чаще всего используются. Подобный подход расширяет возможности совместного использования информации, одновременно повышая эффективность доступа к ней.

Предложенный подход применения ГИС для решения задач формирования, анализа и эксплуатации компьютерных сетей, основан на разработках лаборатории «Интеграции геоинформационных технологий использования и распределения ресурсов компьютерных систем» кафедры «Компьютерные системы», факультета «Компьютерный инжиниринг» Ташкентского университета информационных технологий. Такой подход позволяет, объединить в себе технологии и инструменты для выполнения различных задач, актуализировать данные по объектам сети, оказывать информационные, технологические, учетно-аналитические, опытно-конструкторские и интерактивные услуги, что является перспективным для принятия управленческих решений.

Использование ГИС в телекоммуникационной сфере, прежде всего, подразумевает использование для учета объектов телекоммуникационных

сетей и решение актуальных задач связанных с планированием развития сети и оптимизацией текущей инфраструктуры, с целью решения фундаментальных, прикладных задач коммуникационных технологий, использования и распределения ресурсов компьютерных систем, организации и планирования лабораторных и опытно-конструкторских работ.

Использование ГИС для учета объектов и картографического моделирования телекоммуникационных сетей продиктовано планированием развития сети и оптимизацией текущей инфраструктуры, что подразумевает решение следующих первоочередных задач: анализ логических схем; анализ территориально-распределенных систем; анализ экономической обстановки крупных регионов; разработка программных продуктов в среде ГИС; разработка инструкций, апробация и ее реализации.

Географическая информационная система телекоммуникационных сетей обеспечивает соответствующей технической документацией системы, и представляются в полном объеме с широким набором атрибутивной информации по каждому объекту.

Внедрение ГИС в сетях телекоммуникации обеспечивает:

– качественный мониторинг параметров сети;
– оперативное диспетчерское управление;
– прогнозирование поведения сети связи в различных условиях;
– инвентаризация оборудований сети и средств связи;
– мониторинг транспорта, построение маршрутов, оптимизация трафиков;
– планирование развития сетевой инфраструктуры.

Система включает три основных этапа, позволяющих проанализировать широкий круг вопросов: стратегический (анализ путей оптимизации телекоммуникационных сетей); тактический (анализ экономической среды реализации ГИС проектов) и оперативный (оформление ГИС проектов и программных ресурсов, а также организация процесса их реализации).

Для реализации этой логической схемы исследований необходимо проведение с адекватного расширения и модификаций математическо – геоинформационного аппарата изучений. Решение задач стратегического этапа разработки ГИС проекта развития телекоммуникационной сети формируется как оптимизационно - территориальная региональная модель. При решении задач реализации телекоммуникационной сети применяется принципиально иная группа моделей (имитационная модель координации интересов, поведенческая модель реализации программы), соответственно, отработка схемы проведения согласованных расчетов с использованием этих моделей. При решении задач анализа ГИС проекта решаются задачи актуализации данных по объектам телекоммуникации (сотовая связь, традиционные сети, стратегическое планирование телекоммуникационных сетей, выбор оптимального расположения антенн, ретрансляторов и др., определение маршрутов прокладки кабеля, мониторинг состояния сетей, оперативное диспетчерское управление) и т.д.

Дополнительные возможности интеграции ГИС схемы в сфере телекоммуникационной системы, позволяет получить полный доступ к геоинформационному описанию каждого объекта непосредственно через геобазы данных, а также взаимосвязь с учетными данными из других фактографических баз данных. Эти данные используются для анализа с целью выявления нарушений и отклонений в работе телекоммуникационной сети. Такой подход имеет свои несомненные преимущества в областях связанных с коллективным использованием накопленных данных и их интеграцией в единый информационный массив, что является важным фактором повышения эффективности эксплуатации геоинформационных систем.

В лаборатории «Интеграции геоинформационные технологий использования и распределения ресурсов компьютерных систем» кафедры «Компьютерные системы» разработаны ГИС проекты нижеследующим направлениям:

– ГИС модели телекоммуникационных систем;
– ГИС модели систем электроснабжения;
– ГИС модели сетей теплоснабжения;
– ГИС модели сетей водоснабжения и водоотведения;
– ГИС модели по использованию аудиторного фонда ТУИТ;
– ГИС модели территориального расположения и площади природного озеленения;
– ГИС модели взаимосвязи ТУИТ с филиалами.

Анализ результатов показывает, что ГИС объединяют данные, накопленные в различных подразделениях или даже в разных областях и сфере деятельности. Имеет свои преимущества коллективное использование накопленных данных и их интеграция в единый информационный массив дает существенные конкурентно способные показатели и повышает эффективность телекоммуникационных систем.

Разработка и внедрение ГИС сетей связи отличают: развитые аналитические функции; возможность управлять большими объемами данных; инструменты для ввода, обработки и визуального отображения геопространственных данных.

Выводы. Внедрение современных геоинформационных технологий для анализа пространственных данных позволяет ускорить и повысить эффективность процедуры принятия решений.

Схематизация пространственных данных на ГИС основе, в том числе в трехмерном измерении, позволяет, упростить организации запросов и их последующий анализ.

Удобное представление данных территориального характера, повышает качество, и ускоряет их анализ. Пространственные данные в геоинформационных системах представляют в виде интерактивных карт. Состояния объектов в ГИС представляются в виде графиков, диаграмм, трехмерных изображений, тем самым дают возможность, существенно

экономить временные ресурсы, оптимизировать процесс работы со схемами, картами и созданием 3D модели местности.

MS WORDDA MA'LUMOTLAR BA'ZASIDAN FOYDALANIB AVTOMATLASHGAN MA'LUMOTLAR YARATISH.

Yunusov, O.P. Ortiqov Z.U., Andijon davlat universiteti, Akbarova F.Sh. Andijon mashinasozlik instituti Akademik litseyi, O'zbekiston.

Annotation: *This article discusses and illustrates how to organize the flow of data in Excel to a document created in Word.*

Keywords*: computer, Word, Excel, database, data flow, calendar.*

Аннотация: *В этой статье обсуждается и иллюстрируется, как организовать поток данных в Excel для документа, созданного в Word.*

Ключевые слова: *компьютер, Word, Excel, база данных, поток данных, календарь.*

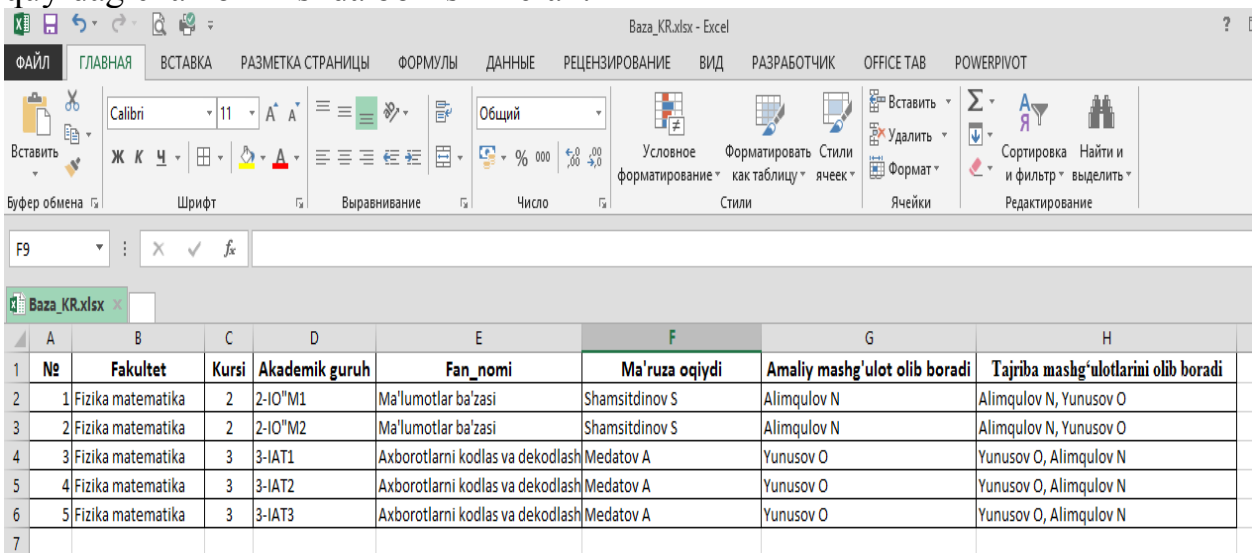
Annotatsiya:*Mazkur maqolada Word dasturida yaratilgan xujjatga Excel dasturidagi ma'lumotlar oqimini qanday qilib, avtomatk tarzda oqib o'tishini tashkil etish masalasi ko'rib chiqildi va yoritildi.*

Kalit so'zlar: *kompyuter, Word, Excel, ma'lumotlar bazasi, ma'lumotlar oqimi, kalendar-reja.*

Keyingi davrlarda axborot texnologiyalarini juda tez va shiddat bilan rivojlanishida ma'lumotlar ba'zasidan foydalanib, turli hil ma'lumotlar va xujjatlar tayyorlash juda dolzarb masalalardan biri hisoblanadi. Shuning uchun biz ham mana shunday masalalardan birini ma'lumotlar ba'zasidan foydalanib, juda oson hal qilishga xarakat qildik.

Bu masalani qoyilishi quyidagicha. Bilamizki Oliy Ta'lim tizimida xar bir Professor–o'qituvchi xar o'quv yili boshlanishida o'zining dars taqsimoti bo'yicha xar bir fandan kalendar ish rejasini tayyorlashga duch keladi.

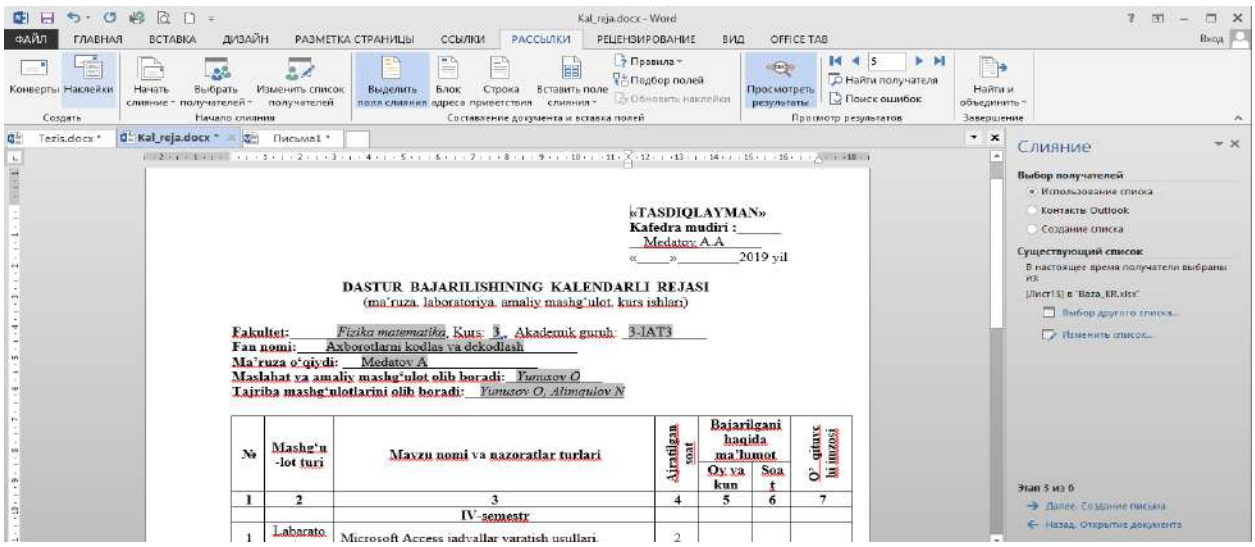
Mana shu masalani oson hal etish mumkin. Uning uchun MS Exsel dasturida bitta ma'lumotlar bazasi sifatida ma'lumot tayyorlab olinadi. Masalan, u quyidagicha ko'rinishda bo'lishi kerak:



	№	Fakultet	Kursi	Akademik guruh	Fan_nomi	Ma'ruza oqiydi	Amaliy mashg'ulot olib boradi	Tajriba mashg'ulotlarini olib boradi
2	1	Fizika matematika	2	2-IO"M1	Ma'lumotlar ba'zasi	Shamsitdinov S	Alimqulov N	Alimqulov N, Yunusov O
3	2	Fizika matematika	2	2-IO"M2	Ma'lumotlar ba'zasi	Shamsitdinov S	Alimqulov N	Alimqulov N, Yunusov O
4	3	Fizika matematika	3	3-IAT1	Axborotlarni kodlas va dekodlash	Medatov A	Yunusov O	Yunusov O, Alimqulov N
5	4	Fizika matematika	3	3-IAT2	Axborotlarni kodlas va dekodlash	Medatov A	Yunusov O	Yunusov O, Alimqulov N
6	5	Fizika matematika	3	3-IAT3	Axborotlarni kodlas va dekodlash	Medatov A	Yunusov O	Yunusov O, Alimqulov N

1-rasm. Excel da hosil qilingan ma'lumotlar ba'zasi.

So'ngra MS Word dasturidan foydalanib birorta fanning kalaendar-reja namuna fayli hosil qilib olinadi. U masalan quyidagi ko'rinishda:



2-rasm. Wordda hosil qilingan namuna fayli sahifasi.

Keyin Word dasturining “РАССЫЛКИ” bo’limiga o’tildi hamda “Выбрать получателей” tugmasi orqali ‘’Использоватт существующий список” bandi tanlanib, oldingi hosil qilingan Baza.xls fayli ochiladi. Keyingi qadamda esa “Начать слияние” tugmasi bosilib undan “Пошаговый мастер слияния” bandi tanlanadi. Endi esa “Вставить поле слияние” tugmasi orqali 2-rasmda ajratib ko’rsatilgan maydonlar 1-rasmda korsatilgan mos maydonlar bilan birma-bir bo’g’lab chiqiladi. So’ngra “Просьотреть результаты” tugmasi orqali hosil bo’lgan Kalendar-reja sahifalarini faqat bir sahifani o’zida ko’rish mumkin bo’ladi. Oxirgi qadamda esa “Найты и объединить” tugmasini bosilib undan “Изменить отдельные докуьенты” bandi tanlanib avtomatik tarzda hosil qilingan xujjatlarni alohida-alohida sahifalarda ko’rish mumkin bo’ladi. Bu ma’lumotlar Письма1.docx nomli faylda avtomatik tarzda hosil bo’ladi.

Shunday qilib yuqorida qoyilgan masalani xal etib maqsadga erishgan bo’lamiz. Bu masalani xal etishda Word dasturida hosil qilinadigan xujjatga, Excel dasturidagi ma’lumotlar oqimini oqib o’tishini tashkil etish orqali erishdik.

Adabiyotlar

1. *O`.T.Haitmatov va b. Informatika va axborot texnologiyalari. O’quv qo’llanma. T. TKTI. 2005 y.*
2. *O`.T.Haitmatov va b. Informatika va axborot texnologiyalari fanidan laboratoriya ishlarini bajarish ushun uslubiy qo’llanma. T. TKTI. 2005 y.*
3. *Aripov M., Xaydarov A. Informatika asoslari T. “O`qituvchi” 2002y.*
4. *Holmatov T.X.,Toyloqov N.I. Amaliy matematika,dasturlash va kompyuterning dasturiy ta’minoti. T.Mexnat, 2000 y.*

EXCEL jadval hisoblagichida iqtisodiy va moliyaviy masalalami yechish : R.X. Ayupov ; 0 ‘zbekiston Respublikasi Oliy va o‘rta maxsus ta’lim vazirligi. - T.: ≪Tafakkur-Bo‘stoni≫, 2012-yil, 148 bet.

ANALYSIS OF EFFICIENT ALGORITHM FOR WEB NEWS MINING BY USING DATA MINING TOOLS

Talha Tariq, Riphah International University, Department of Computer Science, Lahore Campus, Pakistan.

Abstract. Web news mining is an application of data mining techniques to find information patterns and relationship from the web data. Web mining is applied to less structured data or unstructured data, dynamic or large scale data. The purpose of this paper is to find efficient algorithm for web news mining with the objectives of analysis of web news data using data classification techniques. Analyze the best way of application of algorithm on web news data is carried out in this paper as well as determining the accuracy of web news dataset using data mining tools and techniques.

Keywords: web news mining, classification algorithms, stringtowordvector filter, weka

INTRODUCTION

Web mining is defined as the use of data mining techniques to automatically discover and extract information from the web document and services [2]. Web mining referred as Web Content Mining, Web Structure Mining and Web Usage Mining. For discovering useful data (videos, tables, audio, images etc.) from the web different techniques and tools are used.

Figure 1: Web Minning

Information over the internet is huge and increasing with passage to time due to which size of databases are also growing. Discover knowledgeable information and analyzing the data sets for relevant data is much difficult because data over the internet in not in plain text. It could be semi-structured and unstructured data which is quite difficult to handle.

Purpose of this paper is to describe Web news mining, it portrays the examination of web news and is a remarkable kind of open information which has unusual characteristics [4]. The presence of different reliable news sources and quick news invigorates are two fundamental complexities. Henceforth, new approaches, advancements and devices ought to be made remembering the ultimate objective to achieve the unmistakable targets proposed around there.

Collected Dataset from different-different news websites is not appropriate for experiment work. It must be preprocessed by data cleaning and transformed into appropriate form of data for exploratory work.

So, First step of proposed approach is data preprocessing. By using data mining tools and its classification techniques to classify the web news data and clustering techniques for clustering. The goal of this paper is to find efficient algorithm for web news mining with the objectives of analysis of web news data using Data Clustering and Classification Techniques.

LITERATURE REVIEW

In this paper researcher present a basic outline of web mining techniques which changed the scenario of today world and help in retrieval of information. Web Mining actually referred as mining of interesting pattern by using set of tools and techniques from the vast pool of web [1].

When the web user access the network a large amount of data is generated and it is stored in web log files. The web log file contains previous user navigation data or historical data. This web access pattern is used to find the user access behavior. Through this behavior analysis it is much easier to predict the next set of pages the user going to visit. This approach is based on session, Time and frequency based analysis using comparison of different classification techniques such as c4.5, AdaBoost M1, SVM, Rule Part. Finally predict the accuracy result for web page recommendations [2].

In this paper researcher examine the role of information mining for information extraction in web page, structure and usages mining in current web models, and also the outlines the method of extracting patterns from information. This paper in addition gift processing primitives, from that processing question languages are designed. Concerning of problems a way to integrate an information mining system with a database or data warehouse are mentioned. Also looking for a classification of knowledge mining systems, and its troublesome analysis issues for building processing tools of the long-standing time [3].

In this paper researcher present a novel application intended to identify the nation targeted by a specific web page. The goal is to be able to automatically distinguish websites targeting a specific nation, using both the URL and the content of a web page. Addresses the issue of identifying Algerian-interest web pages using a machine learning approach. And present the process of gathering data for the supervised learning phase and adapting it into a usable dataset, as well as using it to construct three distinct classifiers using different parts of the data [5].

This article aims at designing and carring out a web application to extract news information using new features such as geolocation and time information as well as showing a comparative study on three different mining techniques. Application can run on different devices including Tablets, Smartphones and Laptops. Moreover, the application can retrieve information features accordingly. Then, the obtained information could be used as a basis for starting or as input for the data-mining techniques, including K-Nearest-Neighbor (k-NN), decision tree

and deep-learning recurrent neural network (such as Long Short-Term Memory 'LSTM'). These techniques are separately implemented and they are compared in terms of time/space complexity and classification accuracy [6].

In this paper researcher examine that users prefer World Wide Web more to upload and download data. Data is increasing gradually over the internet, so it is getting difficult and time consuming for discovering informative knowledge and patterns. Digging knowledgeable and user queried information from unstructured and inconsistent data over the web is not an easy task to perform. Different mining techniques are used to extract relevant information from web (hyperlinks, contents, web usage logs). Web data mining is a sub discipline of data mining which mainly deals with web [7].

In this paper researcher examines that the primary goal of the website owner is to provide the relevant information to the users to fulfill their needs. Web mining techniques are used to categorize users and pages by analyzing users behavior, the content of pages and order of URLs accessed. Web Structure Mining plays an important role in this approach. In this paper researcher discuss and compare the commonly used algorithms i.e. PageRank, Weighted PageRank and HITS [8].

Research Questions

In [4], authors uses WEKA data mining tool. And use its classification techniques namely Naive Bayes Multinomial technique, J48 technique and SMO technique to classify the web news data and K-means clustering techniques for clustering. The goal of this paper is to find efficient algorithm for web news mining with the objectives of analysis of web news data using Data Classification Techniques.

- The first limitation of their work they performed experiment on 20 days (20-12-2017 to 10-01-2018) data.
- Secondly they use only classification techniques namely Naive Bayes Multinomial technique, J48 technique and SMO technique to classify the web news data.
- There is need to find efficient classifiers for web news mining using Weka data mining tool.

METHODS AND MATERIALS

Dataset is collected from different news websites is not appropriate for experiment work. Data must be pre- processed by data cleaning and data must be transformed into appropriate form of data for exploratory work. So, First step of proposed approach is data preprocessing. For this study, all news text files are stored in a folder. With the help of Text Directory Loader class of WEKA, data is loaded for pre-processing. After preprocessing of data in WEKA, it shows text and class two attributes. For data classification, we have to convert text attribute into collection of words that is word vector. For Data classification WEKA's String to Word Vector unsupervised filter is used. With the help of this filter we convert string attribute into set of attributes representing word based on which tokenizer

used. In this paper classification techniques are applied on processed dataset. In first phase classification technique is applied on dataset by using different classification algorithms like Naive Bayes Multinomial, SMO (Sequential Minimal Optimization) , Decision Tree (J48) and Random Forest. After that performance of each algorithm is calculated and compare. Now performance of each algorithm is calculated and outputs of these algorithms are compared.

Methodology

The figure 2 shows the methodology to achieve the objectives of this paper.

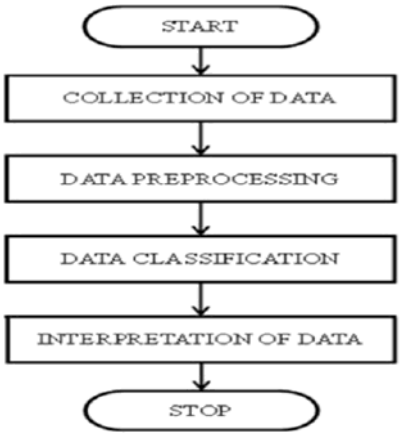


Figure 2: Depicts the methodology of proposed work

Data Collection

All the data required for this study is collected from different-different news websites i.e. The Hindu, Times of India.

Preprocessing

Data which is collected from different-different news websites is not appropriate for experiment work. Data must be pre-processed by data cleaning and transformed into appropriate form of data for exploratory work. For this study all news are stored in folder named as news dataset. By using Text Directory Loader class of WEKA, present study load data for pre-processing. When data is loaded in WEKA, it shows two attributes named as text and class. For text classification, it is needed to convert text attribute into collection of words that is word vector. For this purpose WEKA's String To Word Vector unsupervised filter is used. This filter convert string attribute into set of attributes representing word based on which to tokenizer used. To increase the efficiency of the system stop words should be removed. Stop words such as a, an, the, is, am, are etc. And also eliminate the frequently usage words such as conjunctions, names, prepositions, verbs, numbers etc. To remove these words WEKA's unsupervised remove filter is used.

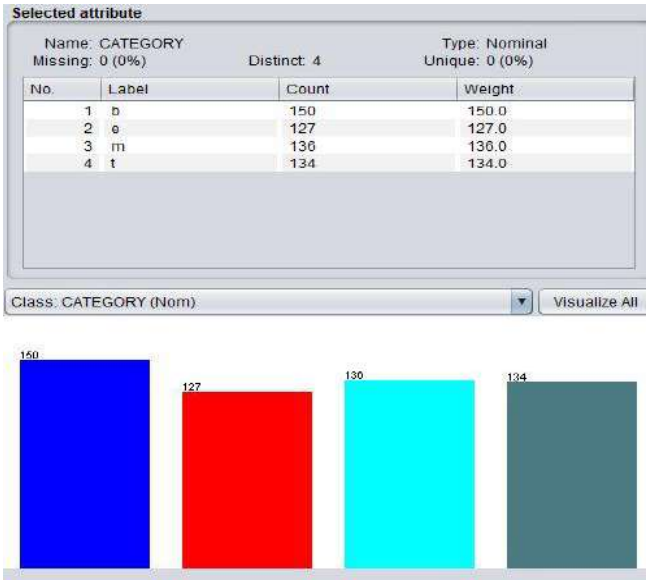


Figure 3: Dataset .arff File

RESULTS AND DISCUSSIONS

This work uses WEKA data mining tool. And used its classification techniques namely Naive Bayes Multinomial technique, J48 technique, Random Forest and SMO technique to classify the data. In this phase of proposed approach all these three classification techniques are applied on the preprocessed news dataset one after another. And the result of each algorithm is calculated and analyzed. Then compare the result of these algorithms with each other.

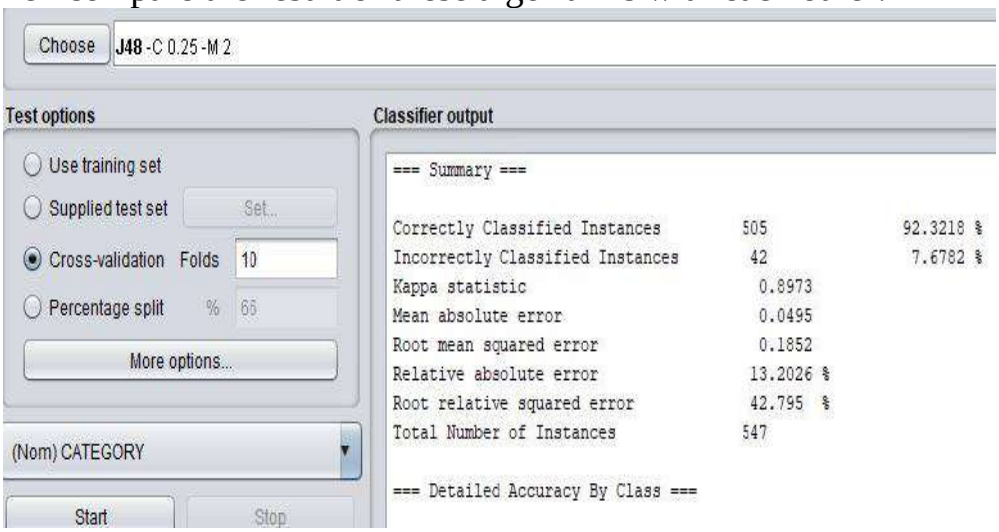


Figure 4: J48 Classifier

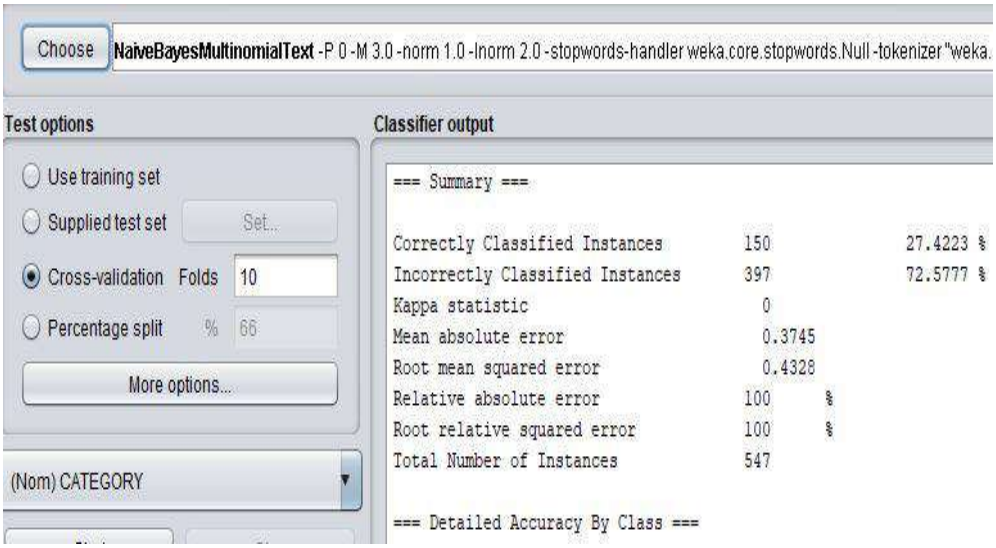


Figure 5: Naïve Bayes Multinomial Text Classifier

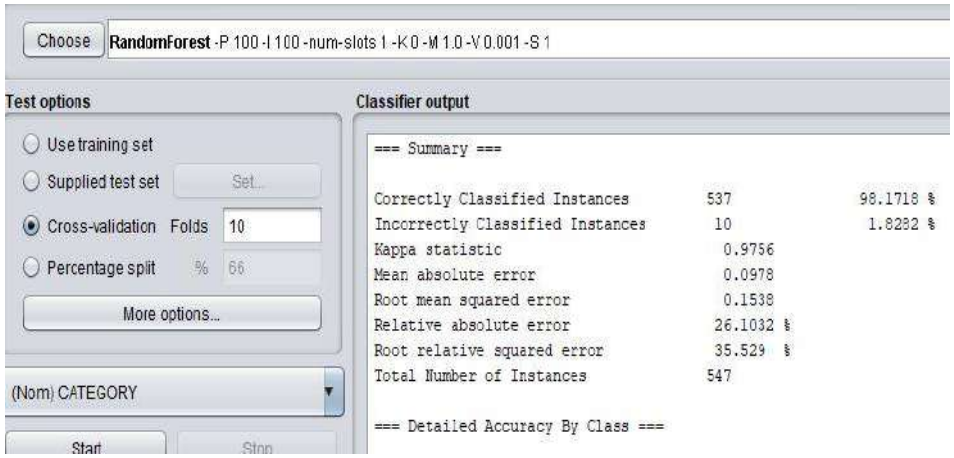


Figure 6: Random Forest Classifier

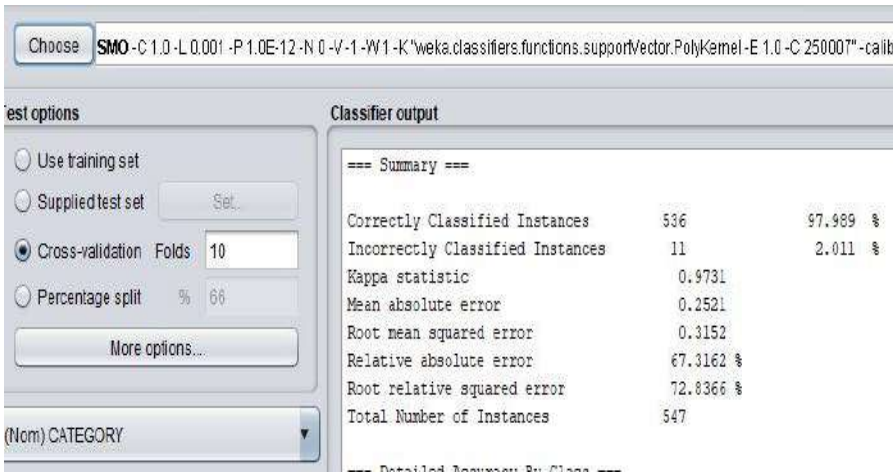


Figure 7: SMO Classifier

The below table shows that Random Forest technique has the best accuracy of 98.17% among other classifiers. And the result of these algorithms are analyzed

and compared. Overall analysis of decision tree (J48 technique),Random Forest, Naive Bayes multinomial and SMO classification techniques is analyzed.

Algorithms	Accuracy
Decision tree (J48)	92.32 %
Naive Bayes Multinomial	27.42 %
SVM(SMO)	97.98 %
Random Forest	98.17%

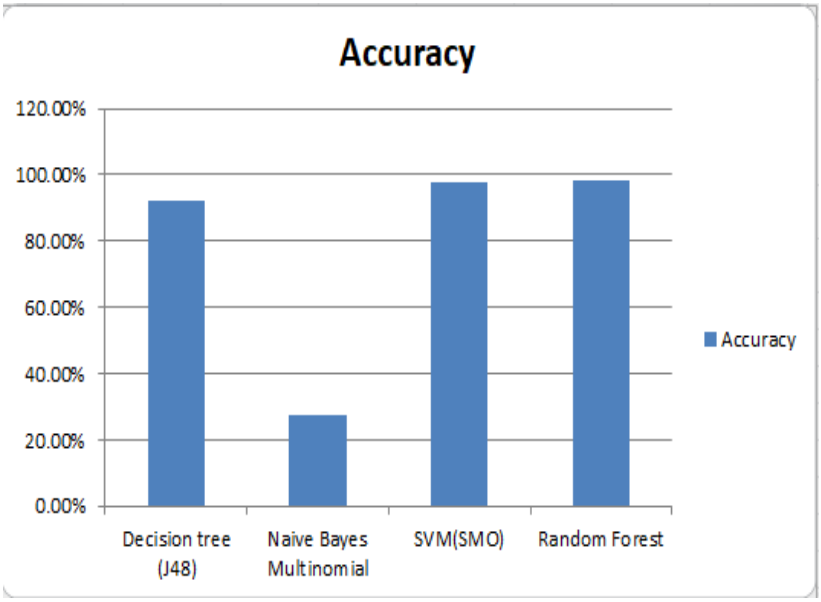


Figure 8: SMO Classifier

CONCLUSION AND FUTURE SCOPE

Conclusion

There are many problems arise to handle the huge amount of data. For this problem different classification algorithms are applied on processed dataset and compare them. Naïve Bayes Multinomial, J48, Random Forest and SMO algorithms are used to classification of news. Proposed work analyzed result on the basis of accuracy. Hence, the proposed work concludes that SMO and Random Forest technique is best for news classification.

Future Scope

- By finding efficient classifier we can apply that algorithm for fake new detection.
- The flow of proposed approach could be applied to other types of web news dataset.

REFERENCES

[14] Irfan, S., & Ghosh, S. (2018). Web Mining for Information Retrieval

[15] Rajeswari, B. & Dr. Nisha, S. (2018). Web Page Prediction Using Web Mining

[16] Patil, P. (2017). Application for Data Mining and Web Data Mining Challenges, IJCSMC, Vol. 6, Issue. 3, pg.39 – 44

[17] Verma, K. & Malviya, K. (2019). An Efficient Way of Handling Large Scale Web News Data using Data Mining Techniques, 2nd International Conference On Advanced Computing And Software Engineering.

[18] Abdessamed, O. & Zakaria, E. Web Site Classification based on URL and Content: Algerian Vs. non-Algerian Case.

[19] Halgurd, S. & Maghdid , (Member, IEEE). (2019) A Comparative Study on Web News Mining Using New Features.

[20] Hamid, J. (2018). Data Mining: Web Data Mining Techniques, Tools and Algorithms: An Overview (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 9, No. 6, 2018

[21] Jain, R. & Purohit, N. (2011). Page Ranking Algorithms for Web Mining, International Journal of Computer Applications (0975 – 8887) Volume 13– No.5, January 2011.

[22] Jebakumari, A. & Goldena, N. (2019). A Survey on Web Content Mining Methods and Applications for Perfect Catch Responses

[23] Sharma, S. & Chauhan, A. & Sharma, P. (2019). Novel Pattern Classification Techniques for Web Mining, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8, Issue-7S, May 2019.

[24] Belgiu, M., Drgu, L.: Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing 114, 24– 31 (2016)

[25] Breiman, L.: Random forests. Machine learning 45(1), 5–32 (2001)

[26] Goyal, A., Mehta, R.: (2012). Performance comparison of Nave Bayes and J48 classification algorithms. International Journal of Applied Engineering Research 7(11), (2012)

CLASSIFYING HANDWRITTEN DIGITS BY USING DIFFERENT CLASSIFIERS IN WEKA

Talha Tariq, Riphah International University, Department of Computer Science, Lahore Campus, Pakistan.

Abstract. Recognition of handwritten digits is important issue in pattern recognition applications. The aim of the paper is to evaluate the ability to develop an effective classifier that can recognize hand written digits and which is submitted by users through tablets, and other digital devices. After applying the image filter on minist dataset, several machines learning algorithm namely, multilayer perceptron, support vector machine, random forest, and j48 has been used for the recognition of digits using weka. The results indicate that the random forest algorithm performs best in classifying handwritten digits images, while the random tree performs worst with the lowest classification accuracy.

Keywords: *Handwritten Digits Recognition, Classification Algorithms, Image Filters, MNIST Dataset, WEKA*

INTRODUCTION

Identifying digits written by human using a machine learning approach is Handwritten digits recognition. Machine learning models can be applied to recognition of handwritten digits problem, however, research issue is that the recognition of handwritten digits, many researches have been carried out and no one of them has suggested a perfect method to recognize handwritten digits.

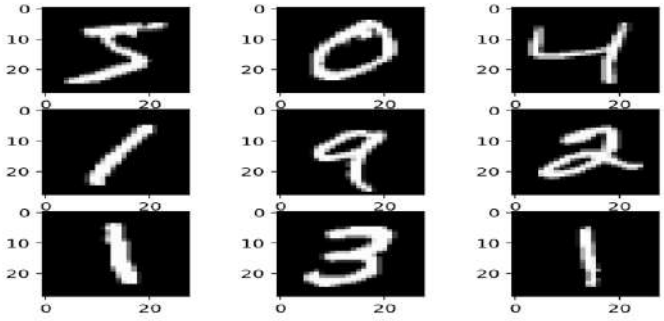


Figure 1: Handwritten Digits Images

The standard dataset MNIST have simplified the recognition of handwritten digits problem. Datasets have huge number of sample images of handwritten digits which can be used in machine learning models like Multilayer Perceptron, Support Vector Machine, Random Forest, and J48 has been used for the digits recognition using WEKA.

Purpose of this paper is to describe handwritten digits recognition, to evaluate the ability to develop an efficient classifier that can recognize hand written digits and which is submitted by users through tablets, and other digital devices.

MINIST Dataset is taken for experiment work. It must be preprocessed by applying image filters and transformed into appropriate form of data for exploratory work. After applying the image filter on MINIST dataset, Several machines learning algorithm namely, Multilayer Perceptron, Support Vector Machine, Random Forest, and J48 has been used for the digits recognition using WEKA.

So, First step of proposed approach is data preprocessing. By using classification techniques to classify the handwritten digits. The goal of this paper is to find, which algorithm best classifies the handwritten digits.

LITERATURE REVIEW

The purpose of this paper is to present different classifiers based on statistical techniques. In which that shows a comparison between a multivariate and a probabilistic approach, concluding that methods provides same results in form of test error rate. Experiments are performed on MNIST and USPS databases in binary-level image. Total error rate achieved for the MNIST dataset is 3.6% and 4.4% for the USPS dataset [1].

In this paper researcher used digits dataset for recognition of handwritten digits using a decision tree classifier. And accuracy is evaluated of the model against digits. The machine was trained with a kaggle dataset and obtained accuracy is 83.4%, the experiment result indicates that the decision tree classifier is effective in recognition of handwritten digits [2].

The purpose of this work to develop a model that can recognize handwritten digit from its image with better accuracy by using the concepts of CNN and MNIST dataset. The objective is to develop a model that can recognize the digits. Results of experiment proves that CNN far better than other classifiers. The results can be made better by increasing the convolution layers and number of hidden neurons [3].

In this paper researcher present an approach to off-line handwritten digit recognition based on different machine learning technique. The objective of this paper is to ensure reliable approaches for recognition of handwritten digits. Machines learning algorithm namely, Multilayer Perceptron, SVM, NB, Bayes Net, J48, Random Forest and Random Tree has been used for the recognition of handwritten digits. The maximum accuracy 90.3% is achieved in the recognition process by Multilayer Perceptron. [4].

In this paper, researcher describes the methodology of using WEKA DeepLearning4j for image classification on larger datasets. In conclusion, the best MNIST dataset, we can see that using CNN5 layer gave the best result for MNIST performance measure. [5].

The purpose of the paper is to evaluate the ability of different decision tree classifiers for cultural heritage image classification involving a small sample, based on three types of extracted image features. Decision tree algorithms includes Hoeffding Tree, J48, Random Forest and Random Tree. Results indicate that the Random Forest algorithm performs better in classifying

a small amount of cultural heritage images, while the Random Tree performs worst with the lowest accuracy [6].

Researcher aims to use the concepts of Convolution Neural Network. The pupose of this study is to open the way to digitalization. Researcher aims to learn and apply the concepts of Machine Learning and Neural Networks. Results of experiment proves that CNN far better than other classifiers. Results can be made more better by increasing the convolution layers and number of hidden neurons [7].

In this paper researcher used three classifiers to recognize the handwritten which is Support Vector Machine, KNN and Neural Network. From the result, KNN and SVM forecast correctly the dataset but for Multilayer perceptron Neural Network that is some mistake to predict the number 9 [8].

In this paper classifier was tested on two image datasets. The first dataset is the MNIST dataset The second dataset contains 441 images of the assembly microdevice. There are many results for the MNIST dataset in the literature. In the best cases, the error rates are 0.7, 0.63 and 0.42%. Error rate of 0.61% by LIRA classifier gives as a mean value of 3 trials. The classifier LIRA also shows sufficiently good results in task of pin–hole position estimation [9].

In this paper, researcher implements a model that contains multi-layer fully connected neural network with 1 hidden layer for recognition of handwritten digits. Testing has been performed on MNIST handwritten dataset. MNIST dataset contains 28000 digits images for training and 14000 digits images for test. Our multi-layer artificial neural network has an accuracy of 97.6% with test performance [10].

Research Questions

In [4], authors uses WEKA tool. The objective of this paper is to ensure reliable approaches for recognition of handwritten digits. Machines learning algorithm namely, Multilayer Perceptron, SVM, NB, Bayes Net, Random Forest, J48 and Random Tree has been used for the recognition of digits using WEKA. The maximum accuracy 90.37% is achieved in the recognition process by Multilayer Perceptron. The goal of this paper is to find which algorithm is effective for handwritten digits recognition using Data Classification Techniques.

- The first limitation of their work that, they applies different classifiers without filtering the dataset.
- Secondly they use only classification techniques namely Multilayer Perceptron, SVM, NB, Bayes Net, Random Forest, J48 and Random Tree has been used for the recognition of digits using WEKA. The maximum accuracy 90.37% is achieved in the recognition process by Multilayer Perceptron.
- There is need to improve the accuracy of algorithm for handrittendigits recognition.
-

METHODS AND MATERIALS

MNIST dataset is taken for experiment work. Which has 840 images of handwriiten digits. Data must be pre- processed by applying filters and data must be transformed into appropriate form of data for exploratory work. So, First step of proposed approach is data preprocessing. For this study, all handwritten digits images are stored in a folder. And make .arff file of images folder. Then load the .arff file in weka which contains filename and Class attributes. For data classification, we have to load images folder for this purpose select the colorLayout filter and set the path of images folder. For Data classification WEKA's ColorLayoutFilter is used. After applying the image filter on MINIST dataset, Several machines learning algorithm namely, Multilayer Perceptron, Support Vector Machine, Random Forest, and J48 has been used for the recognition of digits using WEKA. Now performance of each algorithm is calculated and outputs of these algorithms are compared.

Methodology

The figure 2 shows the methodology to achieve the objectives of this paper.

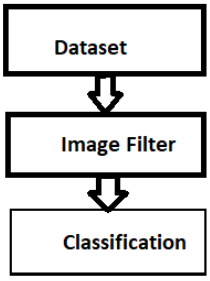


Figure 2: Depicts the methodology of proposed work

Data Collection

MNIST dataset is taken from github. Which contains images of handwritten digits in folder.

Figure 3: Dataset

And make the .arff file of images folder, which contains images name and labels.

mnist.meta.minimal - Notepad

File Edit Format View Help

```
@relation mnist.meta.minimal
@attribute filename string
@attribute class {0,1,2,3,4,5,6,7,8,9}
@data
img_12829_0.jpg,0
img_32870_0.jpg,0
img_28642_0.jpg,0
img_742_0.jpg,0
img_19838_0.jpg,0
img_2344_0.jpg,0
img_16786_0.jpg,0
img_28050_0.jpg,0
img_22624_0.jpg,0
img_21227_0.jpg,0
img_1779_0.jpg,0
img_29822_0.jpg,0
```

Figure 4: Dataset .arff File

Preprocessing

Data must be pre- processed by applying filters and data must be transformed into appropriate form of data for exploratory work [12]. So, First step of proposed approach is data preprocessing.

When we load the .arff file, Before applying the filter that contains filename and class labels.

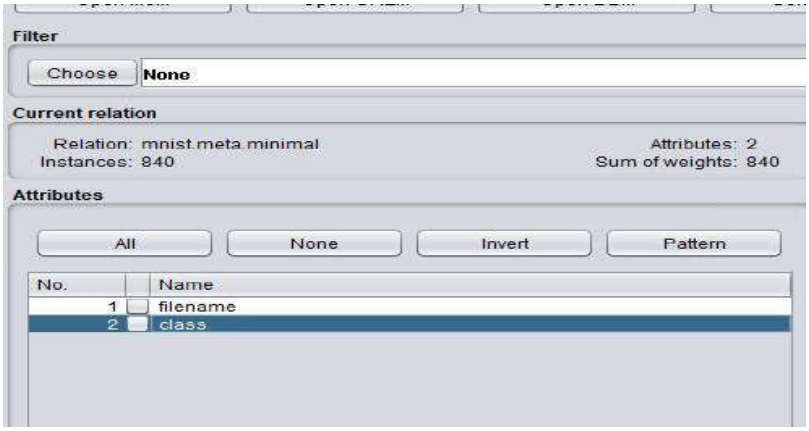


Figure 5: Before Apply Filters

For data classification, we have to load images folder for this purpose select the colorLayout filter and set the path of images folder.

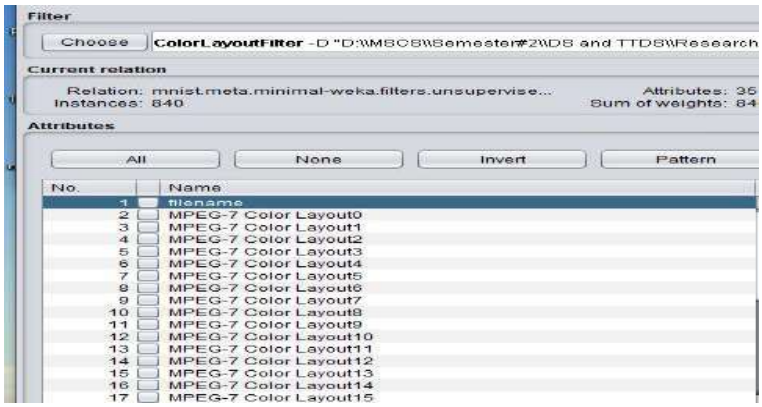


Figure 6: ColorLayout Filter

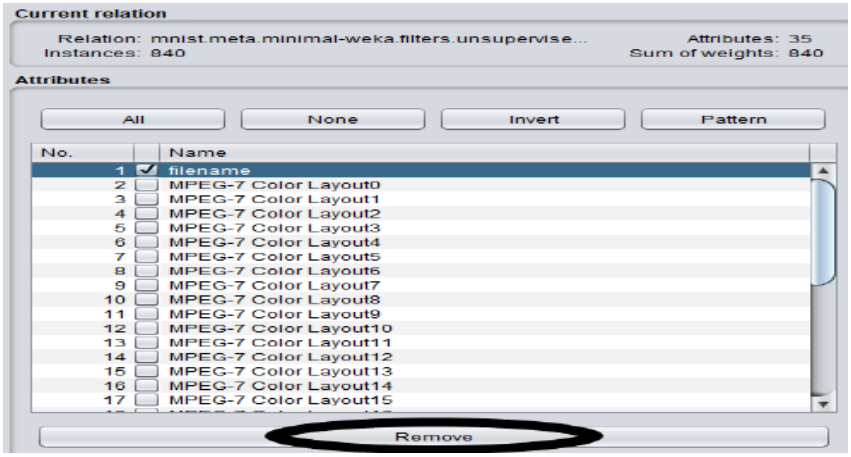


Figure 7: Remove Filename

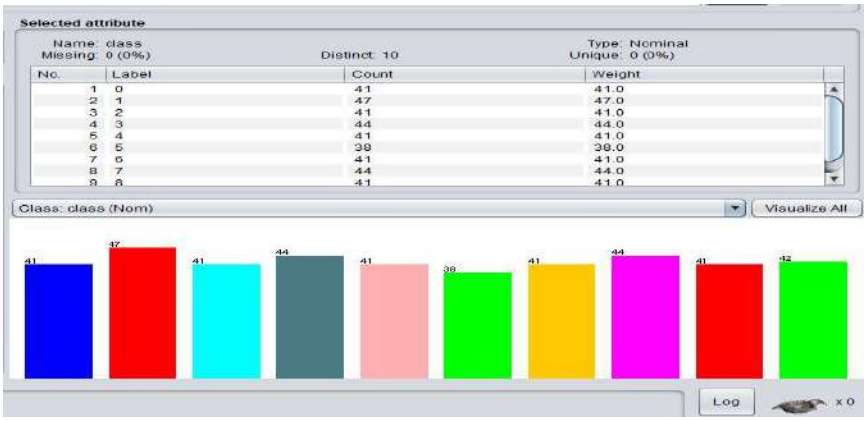


Figure 8: Dataset Classes

RESULTS AND DISCUSSIONS

This work uses WEKA data mining tool. And used its classification techniques namely Multilayer Perceptron, Support Vector Machine, Random Forest, and J48 has been used for the recognition of handwritten digits. In this phase of proposed approach all these four classification techniques are applied on the preprocessed MNIST dataset one after another . And the result of each algorithm is calculated and analyzed. Then compare the result of these algorithms with each other.

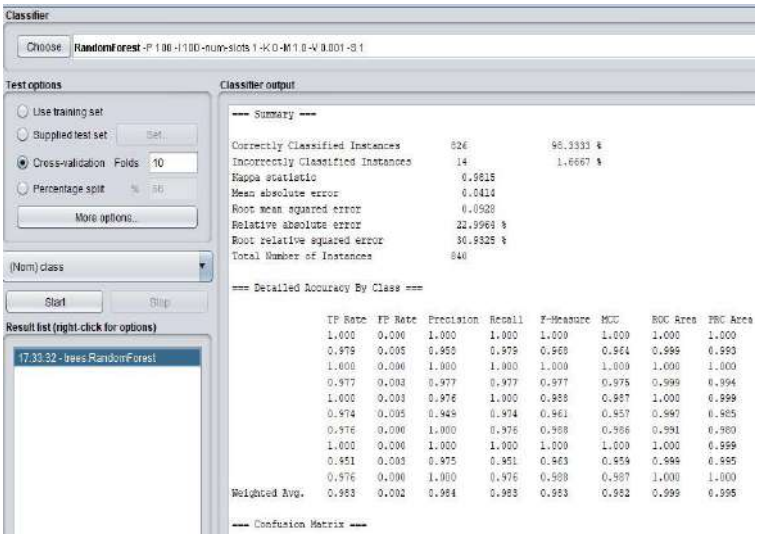


Figure 9: Random Forest Classifier

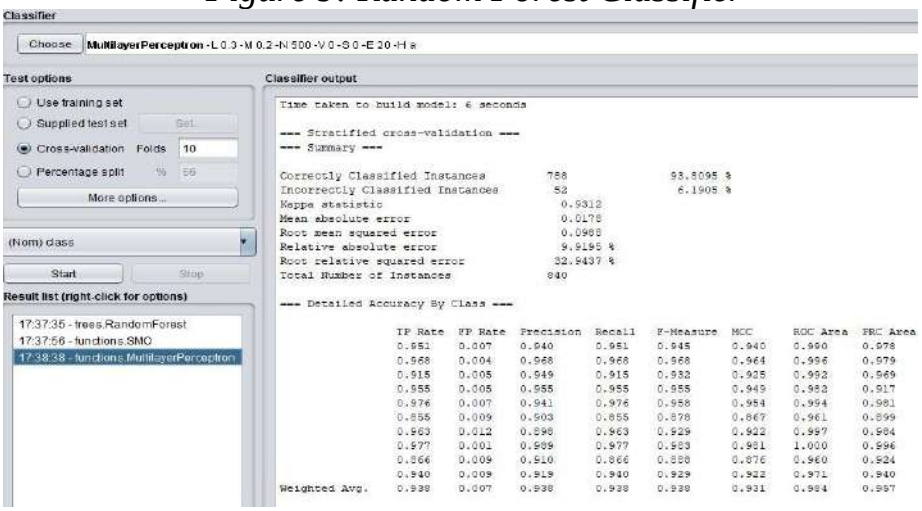


Figure 10: Multilayer Perceptron Classifier

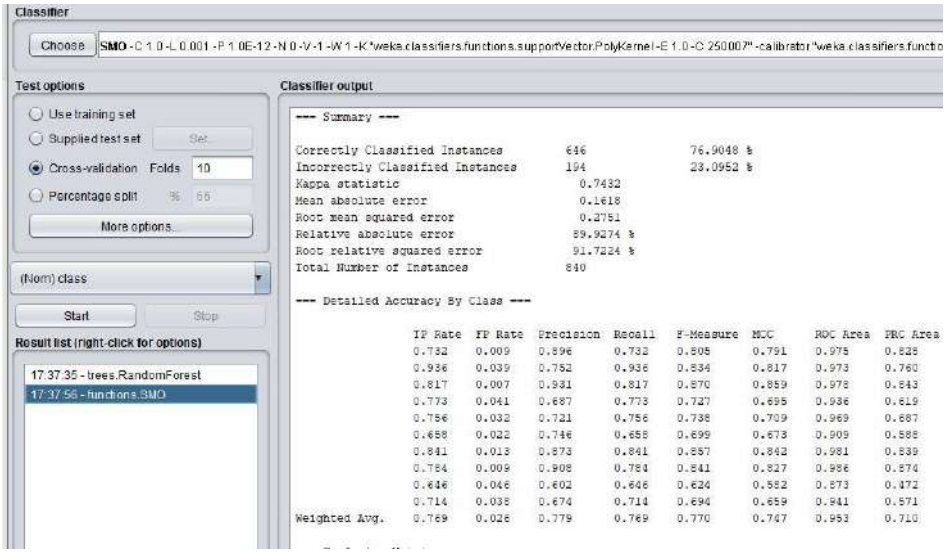


Figure 11: SMO Classifier

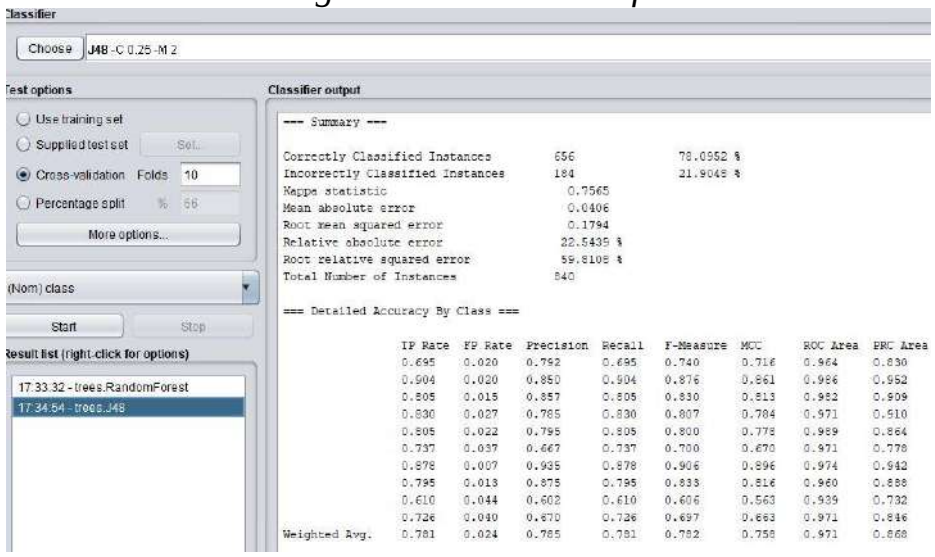


Figure 12: J48 Classifier

Algorithms	Correctly Classified Instances(% Value)	Incorrectly Classified Instances(% Value)
Random Forest	98.3333%	1.6667%
Multilayer Perceptron	93.8095%	6.1905%
SMO	76.9048%	23.0952%
J48	78.0952%	21.9048%

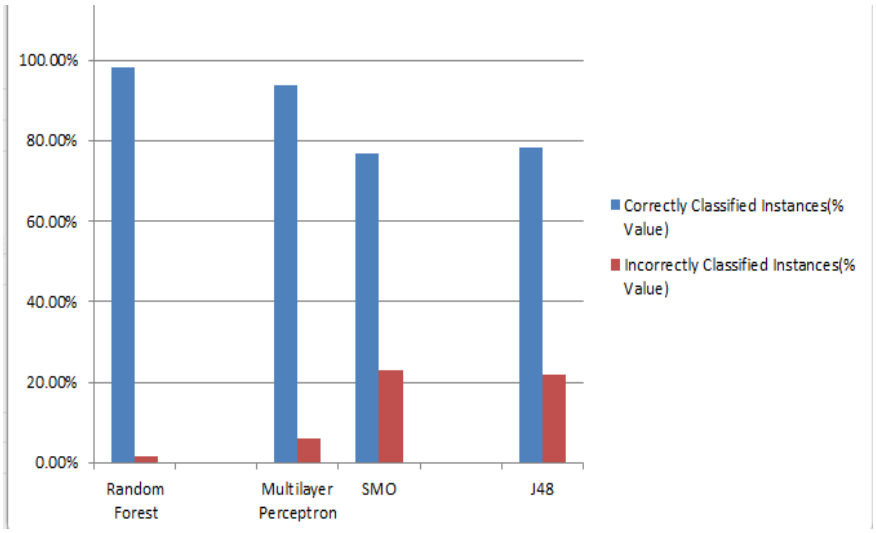


Figure 13: Results Comparison Graph

The table shows that Random Forest technique has the best accuracy of 98.3333% among other classifiers. And the result of these algorithms are analyzed and compared. Overall analysis of decision tree (J48 technique), Random Forest, Multilayer Perceptron and SMO classification techniques is analyzed.

CONCLUSION AND FUTURE SCOPE

Conclusion

The objective of this paper is to find a effective classifier for recognition of handwritten digits. In this paper we used different machine learning algorithms for recognition of handwritten digits. Classifiers accuracy improves by applying the image filter. The highest accuracy 98.3333% is obtained in by Random Forest for recognition of handwritten digits.

Future Scope

By finding efficient classifier we can apply that algorithm for handwritten Characters[a,b,c etc] recognittion.

REFERENCES

[27] Giuliodori, A. & Lillo, R. and Peña, D (2011). HANDWRITTEN DIGIT CLASSIFICATION

[28] Admassu, T. & Sekharan, P. (2019). Handwritten digits recognition with decision tree classification: a machine learning approach, Vol. 9, No. 5, October 2019, pp. 4446~4451 (IJECE)

[29] Anwar, H & Mohon, MD. (2019) Recognition of Handwritten Digit using Convolutional Neural Network (CNN) Volume 19 Issue 2 Version 1.0 Year 2019

[30] Shamim, S. & Mohammad Alam, B. etal (2018) Handwritten Digit Recognition using Machine Learning AlgorithmsVolume 18 Issue 1 Version 1.0 Year 2018

[31] Zainudin, Z. & Mariyam, S. (2019) Deep Learning for Image Processing in WEKA Environment Vol. 11, No. 1, March 2019

[32] Jankovi, R. Classifying Cultural Heritage Images by Using Decision Tree Classifiers in WEKA

[33] Jain, N. & Rahul, K. & Khamaru. I. & Ghosh, A. (2017). "Hand Written Digit Recognition using Convolutional Neural Network (CNN)", International Journal of Innovations & Advancement in Computer Science, IJIACS,ISSN 2347 – 8616,Volume 6, Issue 5.

[34] Norhidayu binti Abdul Hamid & Nilam Nur Binti Amir Sjarif Handwritten Recognition Using SVM, KNN and Neural Network

[35] Kussul, E. & Baidyk, T. (2004). "Improved method of handwritten digit recognition tested on MNIST database". Image andVisionComputing,22(12):971–981

[36] Islam, T. & Mujtaba, G. "Handwritten Digits Recognition with Artificial Neural Network," Proc. of the International Conference on Engineering Technologies and Technopreneurship (ICE2T 2017),Sep2017

[37] Belgiu, M., Drgu, L.: Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing 114, 24– 31 (2016)

[38] Breiman, L.: Random forests. Machine learning 45(1), 5–32 (2001)

[39] Goyal, A., Mehta, R.: (2012). Performance comparison of Nave Bayes and J48 classification algorithms. International Journal of Applied Engineering Research 7(11), (2012)

DELPHI TILIDA DINAMIK MASSIVLAR VA TADBIQLARI

Ahmadaliyev Sh.Sh., Xasanov X.M., M.M.Botirov,
Qo'qon DPI, O'zbekiston.

Annotasiya: *Ushbu maqolada zamonaviy dasturlash tillaridan Delphi dasturlash tilining massiv, dinamik massiy imkoniyatlari misollar orqali yoritib berilgan.*

Аннотация: *В данной статье освещены примеры динамических массивов возможностей современного языка программирования Delphi.*

Annotation: *In this article, the dynamic mass capabilities of the modern Delphi programming language are covered by examples.*

Tayanch so'z va iboralar*: Massiv, dinamik massiv, indeks, StringGrid.*

Ключевые слова и выражения*: Массив, динамический массив, индекс, StringGrid.*

Keywords and expressions*: Massive, dynamic array, index, StringGrid.*

Insonlar hayotlari davomida juda ko'p turdagi jadvallardan foydalanadi. Jumladan, dars jadvali, shaxmat yoki futbol o'yinlari bo'yicha musobaqa jadvali, Pifagor (karra) jadvali, kelishiklar jadvali va boshqalar. Jadvalni tashkil etuvchilar uning elementlari deyiladi. Jadval ko'rinishidagi miqdorlar bir o'lchovli (chiziqli), ikki o'lchovli (to'g'ri to'rtburchakli), uch o'lchovli (parallelopipedli) va hokazo bo'ladi.

Chiziqli jadvallar satr yoki ustun shaklida ifodalanadi. Ikki o'lchovli jadvallar ustunlar va satrlardan tashkil topadi. Ularning elementlari ustun va satrlar kesishgan kataklarda joylashadi.

Dasturlash tilda jadvallar bilan ishlash uchun massiv tushunchasi kiritilgan. Massiv – jadval ko'rinishidagi miqdor bo'lib, u aniq sondagi bir turli va tartiblangan elementlar majmuidan iborat. Massiv elementlarining tartib raqami butun sonlarda ifodalanadi, ya'ni ular manfiy son bo'lishi ham mumkin.

Dinamik massiv deb – dastur bajarilishi jarayonida, ya'ni dastur kompyuter operativ xotirasiga uzatilgandan so'ng o'z o'lchamini o'zgartirish jarayoniga aytiladi.

Dasturlash tilarida har bir massiv o'z nomiga ega bo'lib, ularni nomlash o'zgaruvchilarni nomlash kabidir. Masalan: a5, dars_jadvali, tub_sonlar. Massiv elementlarining tartib raqami indeks deb ataladi va indeks kvadrat qavs ichida yoziladi. Masalan, a[5] yozuvi – a nomli massivning beshinchi elementini bildiradi.

Ikki o'lchovli massiv elementlari ikkita indeks orqali aniqlanib, ular o'zaro vergul bilan ajratib yoziladi va birinchi indeks satr tartib raqamini, ikkinchi indeks ustun tartib raqamini bildiradi. Masalan, D[4,3] yozuvi – D nomli massivning 4-satri va 3-ustuni kesishgan katakda joylashgan elementini bildiradi.

Yuqoridagi singari Delphi dasturlash tilida ham jadval kattaliklari bilan ishlash uchun StringGrid jadval komponentasi mavjud bo'lib, StringGrid jadval

komponentasi yordamida bir yoki ikki o‘lchovli, masalan, matritsa elementlari qiymatini ekranda jadval ko‘rinishida tasvirlash, ular qiymatini kiritish va tahrirlash uchun ishlatiladi. Jadvalning qator va ustun nomerlari noldan boshlanadi. Jadvalning ustun va qatorlar sonini keraklicha o‘zgartirishi mumkin. Bu uning xossasi yordamida aniqlanadi. Jadval har bir kesishgan ustun va satri yacheyka deyilib, unga kiritilgan ma’lumot simvol qatori bo‘lib aniqlanadi. Masalan, (3,5) yacheyka to‘rtinchi ustun va oltinchi qatorda joylashgan.

StringGrid jadval komponentasi quyidagi asosiy xossalarga ega [1]:

ColCount – jadvaldagi ustunlar sonini aniqlaydi;

RowCount – jadvaldagi satrlar sonini aniqlaydi;

FixedCols – fiksirlangan ustunlar sonini aniqlaydi;

FixedRows – fiksirlangan satrlar sonini aniqlaydi;

Options –jadvad holatini aniqlaydi;

ColWidths – jadvaldagi har bir ustun kengligini aniqlaydi;

DefaultColWidth – jadvalning boshlang‘ich ustunlar kengligini aniqlaydi;

DefaultRowHeight – jadval satrining boshlang‘ich balandligini aniqlaydi;

FixedColor – fiksirlangan yacheyka rangini aniqlaydi;

RowHeights – jadval satri balandligini aniqlaydi;

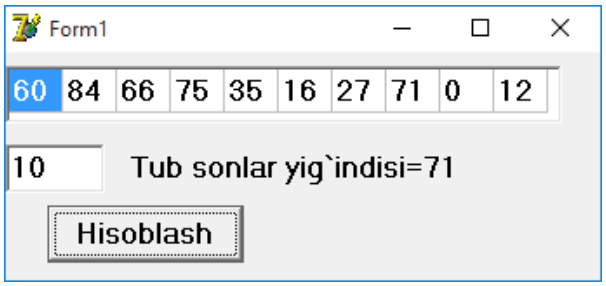
Cells – simvol qatorli ikki o‘lchamli massivni aniqlaydi.

Misol. Kiritilgan N ta sondan tublari yig‘indisini topish dasturini tuzing.

Buning uchun delphi dasturlash tilini yuklaymiz va formaga Stringgrid, Edit, Label, Button komponentalarini joylashtiramiz va Button tugmasi uchun quyidagi dastur kodini kiritamiz.

```
procedure TForm1.Button1Click(Sender: TObject);
var     n,i,j,k,w,s:integer;
        a:array[0..500] of integer;
begin
randomize;
n:=strtoint(edit1.Text );
Stringgrid1.ColCount :=n;
Stringgrid1.Width :=n*32;
form1.Width:=n*35;
for i:=0 to n-1 do
stringgrid1.Cells [i,0]:=inttostr(random(100));
For i:=0 to n-1 do
Val(StringGrid1.cells[i,0], a[i], w);
S:=0; k:=0;  For i:=0 to n-1 do begin
for j:=2 to round(sqrt(a[i])) do
  if a[i] mod j = 0 then k:=k+1;
  if k=0 then s:=s+a[i];
k:=0;
end;
label1.Caption :='Tub sonlar yig`indisi=' + inttostr(s);
end;
```

end.



1-rasm.

Adabiyotlar.

1.Sh.A.Nazirov va bosh. "Delphi tilida dasturlash asoslari" T. 2007 y.
2. Arxangelskiy A.Ya. Delphi 7. Spravochnoe posobie — M.: 2003 g.

ВИРТУАЛ ЎҚУВ ЖАРАЁНИНИНГ АРХИТЕКТУРАСИ, АЛГОРИТМИ ВА ДАСТУРИЙ ТАЪМИНОТИ

Беккулов Ж.Ш., Раҳимов А. М., Қарши муҳандислик-иқтисодиёт институти, Ўзбекистан.

Бугунги кунга келиб таълим тизимида, жумладан Олий таълим тизимида ҳам замонавий ахборот технология имкониятлари ҳамда инновацион технологиялардан фойдаланган ҳолда ўқув жараёнларини олиб бориш жадал тус олмоқда. Шу қаторда Олий таълим тизимида талабаларга билим бериш самарадорлигини ошириш мақсадида ўқув жараёнини виртуаллаштириш масаласи муҳим аҳамият касб этади.

Шу мақсадда виртуаллаштирилган ўқув жараёнининг дастурий таъминотини ишлаб чиқиш мақсадида катта ишлар амалга оширилмоқда. Ягона интерфейсга бирлаштирилган виртуаллаштирилган ўқув жараёнининг дастурий таъминотини ишлаб чиқиш ва ундан тармоқ орқали фойдаланишни таъминлашдан иборат.

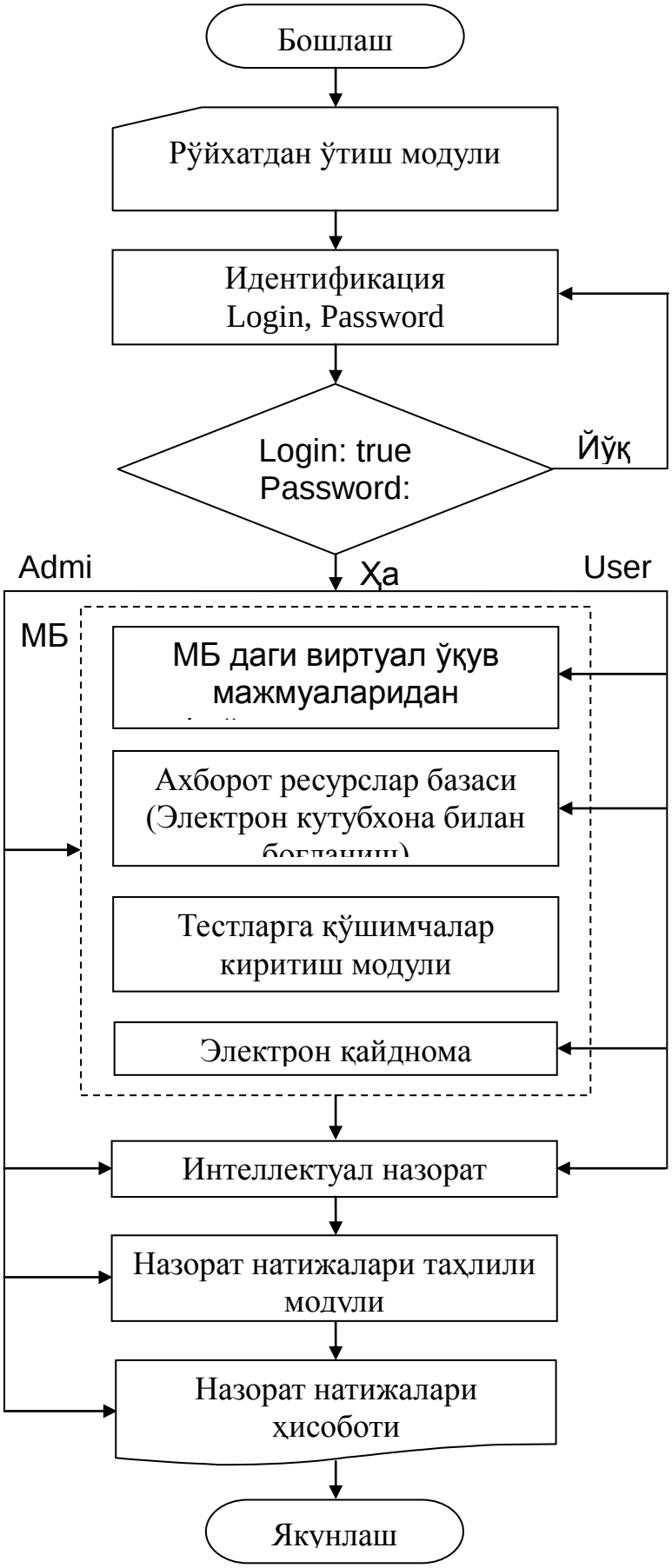
Бунда ўқув жараёнини реал аудиторияда виртуал ўқув қўлланмалари (виртуал маъруза, амалий ва тажриба машғулотлари бўйича) асосида олиб борилиши ҳамда олинган билимни назорат қилиш ва ўқув-машғулотларини қай даражада ўзлаштирилганлигини таҳлил қилиб берувчи виртуал назорат тизимидан фойдаланиш кўзда тутилган. Бундан ташқари дастурий таъминот таркибида ахборот ресурслар базаси мавжуд бўлиб, талабалар мустақил равишда мутахасислик фанлари бўйича кутубхонада мавжуд жамланган адабиётлар рўйхати ва электрон қўлланмаларни қидириб топиш ва фойдаланиш имкониятига эга бўладилар. ЎЖнинг компьютер моделини ишлаб чиқишда унинг ташкил этувчиларининг барчасини ягона интерфейсга бирлаштирилган бўлиб, бу фойдаланувчиларга қулайлик туғдиради. [2]

1.1-расм. Виртуаллаштирилган ўқув жараёни архитектураси

Схемада виртуал ўқув жараёнини компьютер моделинининг архитектураси акс эттирилган бўлиб (1.1-расм), асосий ташкил этувчиларининг функционал вазифаларини кўриб чиқамиз:

Рўйхатдан ўтиш ва Авторизация бўлимида ҳар бир фойдаланувчи ўзи тўғрисида сўралган маълумотларни ҳамда тизимга киришни таъминловчи паролни киритади. Бу маълумотлар эса ўз навбатида МБга жамланиб боради ҳамда тестлаштириш натижаларини таҳлил қилишда қўл келади.[1]

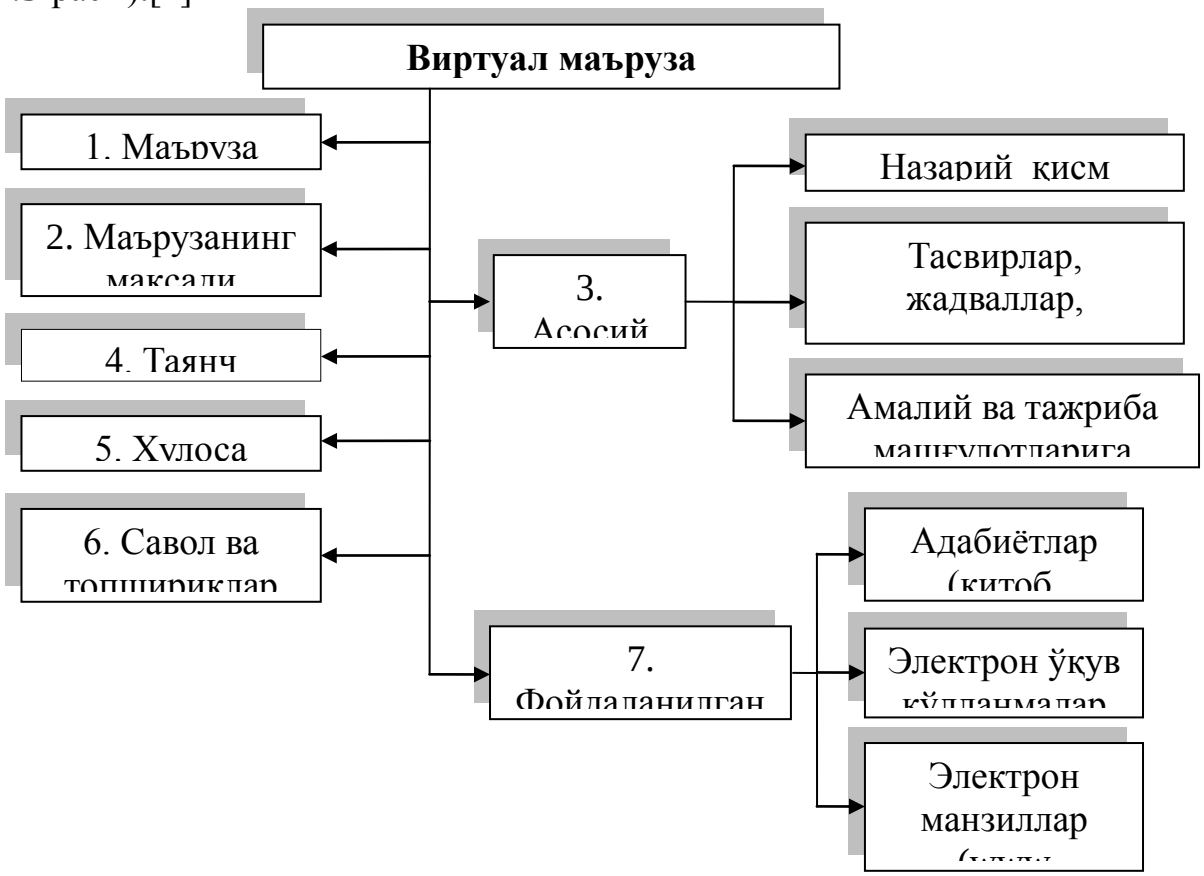
Виртуал ўқув жараёни ўз навбатида электрон маъруза, шу маъруза асосида тузилган виртуал амалий ва тажриба машғулотлари, мустақил таълимда фойдаланиш учун ахборот ресурслар базаси ҳамда олинган билимни текшириш мақсадида виртуал назорат блокларидан иборат бўлиб, бу жараёнлар бир-бирига боғлиқ ҳолда ишлайди (1.1-расм).



1.2-расм. Виртуал ўқув жараёни дастурий муҳитининг блок-схемаси

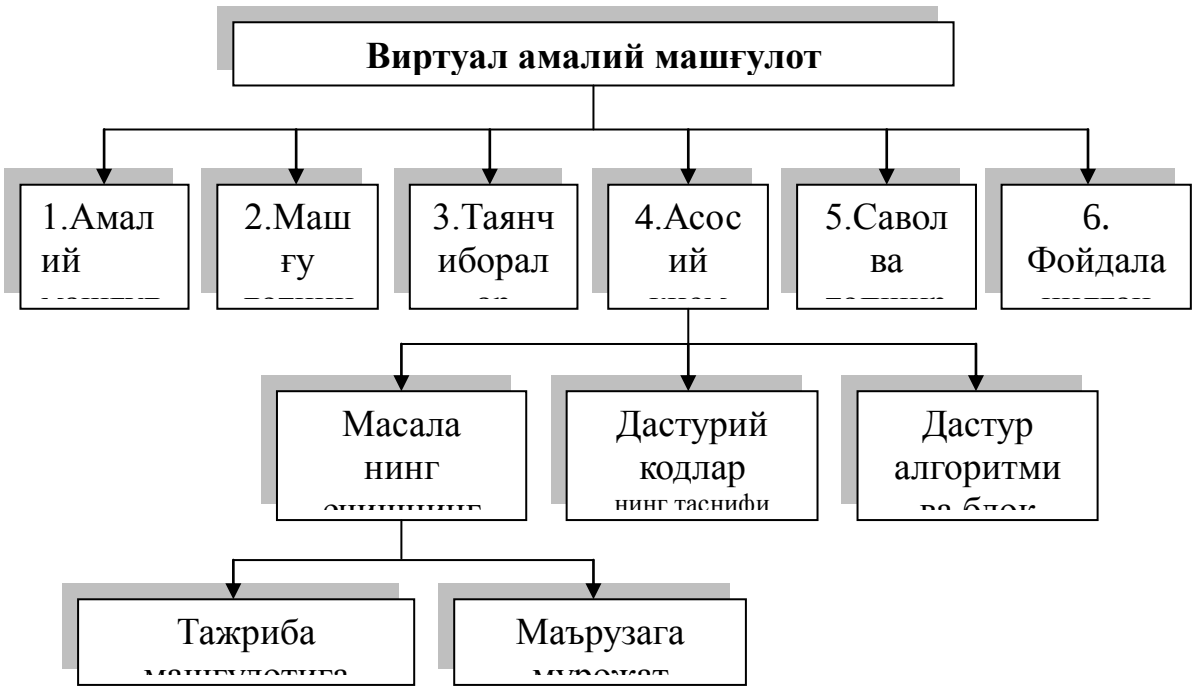
Ўқув жараёнининг компьютер модели орқали унинг барча ташкил этувчиларини ўзаро боғловчи ва администратор томонидан бошқарилувчи ягона интерфейс ташкил этилган. Шунингдек дастурда таҳлил натижалари асосида ўқув машғулот соатларини қайта тақсимлаш ёки тузилган машғулотларни ҳамда тест вариантларини қайта кўриб чиқиш орқали ўқув мажмуалари ва таълим сифати янада ортади. Бундан ташқари дастурий таъминотдан локал тармоқ орқали реал аудитория ЎЖда, ҳамда масофавий таълимда фойдаланиш мумкин.[3]

Электрон маърузада дарс жараёни асосан анимациялар, тасвирлар, графиклар, жадваллар ва схемалар ёрдамида тушунтириб берилган, ҳамда амалий ва тажриба машғулотларига бевосита мурожаат қилиш мумкин. Виртуал маъруза жараёнининг тузилиши қуйидаги схемада ўз аксини топган (1.3-расм).[4]



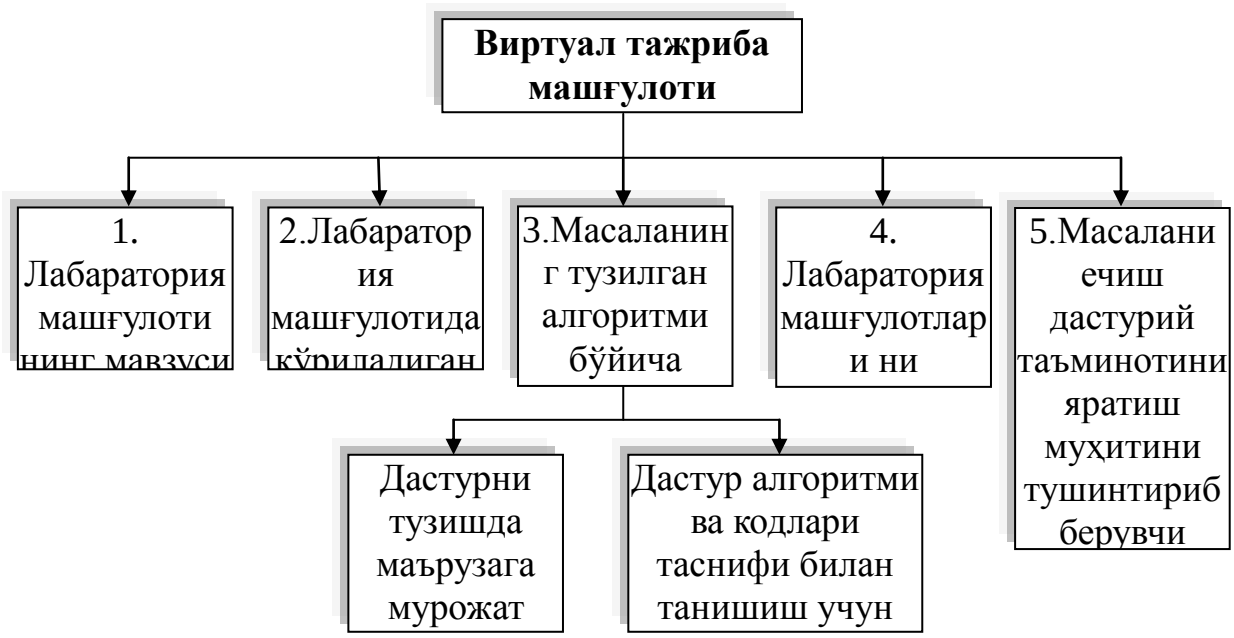
1.3 - расм. Виртуал маъруза структураси

Маърузага асосан виртуал амалий ва тажриба машғулотлари тузилган бўлиб, амалий машғулотида асосан дастур кодларининг таснифи ва дастур ечимининг алгоритмик блок схемалари келтирилган. Амалий машғулотларда ҳам ўз навбатида маъруза ва тажриба машғулотига мурожат қилиш мумкин. Виртуал амалий машғулотининг тузилиши қуйидаги схемада ўз аксини топган (1.4-расм).



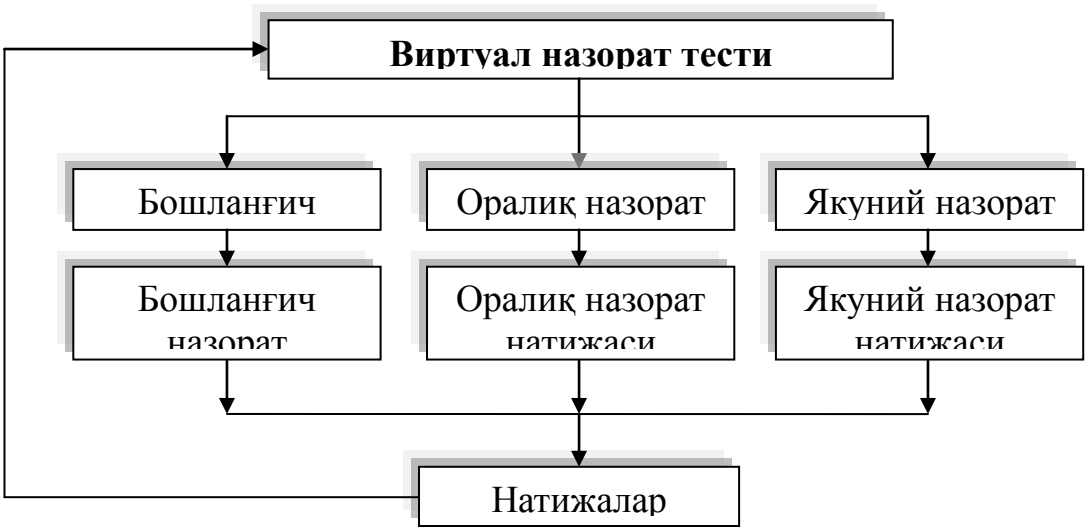
1.4- расм. Виртуал амалий машғулот структураси

Виртуал ўқув жараённинг яна бир асосий қисми бўлган виртуал тажриба машғулотларини яратишда алоҳида меҳнат талаб этилиб, бунда асосан мультимедиа воситалари асосида анимацион ва овозли ўқув қўлланмалар яратилади. Виртуал тажриба машғулотида ечиладиган масала ёки тузиладиган дастур бажарилишининг компьютер моделлари яратилган (1.5-расм).[5]



1.5- расм. Виртуал тажриба машғулоти структураси

Ўқув жараённинг виртуал назорат блоки махсус дастурлар асосида яратилган тестлардан иборат бўлиб, тест якунида талабалар олган билимини таҳлил қилинади (1.6-расм).



1.6 - расм. Виртуал ўқув жараёнида талабалар билимини назорат қилиш дастури структураси

Виртуал ўқув-услубий қўлланмалар Олий ва ўрта махсус таълим вазирлиги томонидан тасдиқланган намунавий дастурлар асосида, Web ҳамда анимацион дастурий воситалар ёрдамида ишлаб чиқилган бўлиб, Олий ва ўрта таълим тизимида тўлиқ фойдаланиш мумкин.[4]

Адабиётлар

1. Агеев В.Н. Электронные учебники и автоматизированные обучающие системы. – М.: 2001. – 79 с.

2. Аджемов А.С. Единое образовательное пространство на основе инфотелекоммуникационных технологий // Сети и системы связи, 2001, №11. – с. 20-23.

3. Александров Г.Н. Программированное обучение и новые информационные технологии обучения. // Информатика и образование, 1993, №5. – с. 7-19.

4. Андерсон Дж. Р., Рейзер Б. Дж. Учитель Лиспа. / В сб. "Реальность и прогнозы искусственного интеллекта" под ред. Стефанюка В.Л. / Пер. с англ. – М.: Мир, 1987. – с. 27-47.

5. Булгаков М.В., Якивчук Е.Е. Инструментальные системы для разработки обучающих программ / В кн. "Компьютерные технологии в высшем образовании". / Ред. кол.: А.Н. Тихонов, В.А. Садовничий и др. – М.: Изд-во Моск. ун-та., 1994. – с. 153-162.

ФУНКЦИЯНИ ХААР ВЕЙВЛЕТЛАРИ ЁРДАМИДА ИНТЕРПОЛЯЦИЯЛАШ

Жўраев Ж.Ў., Самарқанд давлат университети, Ўзбекистан.

Аннотация. *Ушбу мақолада функцияни Хаарнинг бўлак- ўзгармас вейвлетида рақамли ишлаш алгоритми ишлаб чиқилди, ҳамда интерполяциялаш хатолигини баҳолаш натижалари келтирилди. Хатоликни баҳолашда функцияни бўлак- ўзгармас вейвлетининг абсолют хатолиги ва нисбий хатолиги дастур натижасига асосан келтирилди. Функцияни бўлак-ўзгармас вейвлетлари ёрдамида ўзгартириш натижасида ортонормал вейвлетлар ҳосил қилинди ва унинг хатоликларини баҳолаш амалга оширилди.*

Калит сўзлар: *вейвлет ўзгартириш, интерполяция, интерпояциялаш хатолиги, Хаар вейвлети, бўлак- ўзгармас вейвлет.*

Аннотация. *В этой статье был разработан алгоритм численной обработки функции в фрагментарном вейвлете Хаара, а также результаты оценки ошибки интерполяции. При оценке погрешности абсолютная погрешность и относительная погрешность кусочно-постоянные функции были заданы на основе результата программы. Известно, что в результате модификации функции с помощью фрагментарных постоянных переменных формируются ортонормированные вейвлеты и оцениваются ее ошибки.*

Ключевые слова: *вейвлет-модификация, интерполяция, погрешность интерполяции, вейвлет Хаара, кусочно-постоянные вейвлет.*

Annotation. *In this article, an algorithm for the numerical processing of a function in a fragmentary Haar wavelet was developed, as well as the results of estimating the interpolation error. When estimating the error, the absolute error and relative error piecewise constant functions were set based on the result of the program. It is known that as a result of modifying a function using fragmentary constant variables, orthonormal wavelets are formed and its errors are estimated.*

Keywords: *wavelet modification, interpolation, interpolation error, Haar wavelet, piecewise constant wavelet.*

Ҳозирги вақтда вейвлетлардан тасвирларни таниб олиш масалаларида, турли хил сигналларни, масалан, нутқни қайта ишлаш ва синтез қилиш вақтида, табиатдаги ҳар хил расмларни таҳлил қилишда(кўзнинг рангдор пардаси, буйракнинг рентгенографияси, булутнинг ёки сайёра сиртининг сунъий йўлдош расмлари ва бошқалар бўлиши мумкин) гирдобсимон майдонларнинг хусусиятларини ўрганишда ва бошқа ҳолларда фойдаланилмоқда [2].

Хаар-вейвлетининг тўлқин чизиқлари функция графиги билан бирга вақт ўқи бўйлаб чўзилади. Хаар-вейвлетининг графиги кўп ҳолларда функция бўйлаб бир томонлама тўлқин чизиқлар шаклида функцияга яқинлашади, бу эса баъзи бир сигналларни сиқишда яхши натижа беради. Унинг математик

талқини-тўлқин ҳолатларини турли хил частотада таҳлил қилишга имкон беради. Хаар-вейвлет функцияси графигининг амплитудаси нолгача пасайиб тебранувчи тўлқинларни ҳосил қилади.

Сигналларга рақамли ишлов беришда Хаар вейвлет моделларини қўлланилиши фан техникани, ҳамда иқтисодий жараёнларни ривожланишида долзарб масалалардан ҳисобланади [1].

Ушбу ишда гастроэнтерологик сигнални таҳлил қилиш мақсадида дастлабки эксперементал маьлумотлар олинди ва шу маьлумотлар асосида Хаарнинг бўлак- ўзгармас вейвлет модели қурилиб уларнинг хатоликлари баҳоланди. Хаарнинг бўлак- ўзгармас вейвлет моделини қурилиши функцияларни унга бирор муайян маънода яқин ва тузилиши соддароқ бўлган функцияларга алмаштириш ғоясига асосланган.

Сигналларни вейвлет ўзгартириш жараёни икки кўринишдаги функциялардан фойдаланишга таянади: вейвлет функция ва масштаблаш функцияси, яъни улар битта оналик вейвлети $\psi(t)$- ни сигнал бўйлаб вақт бўйича силжитиш b ва вақт масштабини a ўзгартириш йўли билан қурилади[3]:

$$\psi_{ab}(t)=\frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right),\quad (a,b)\in R,\ \psi(t)\in L^2(R) \tag{1}$$

Сигналларга рақамли ишлов беришда сигналларнинг деталларини ва локаллик хусусиятларини ажратиш учун вейвлет функциялардан, сигналларни аппроксимациялаш учун эса масштаблаш функциясидан фойдаланилади.

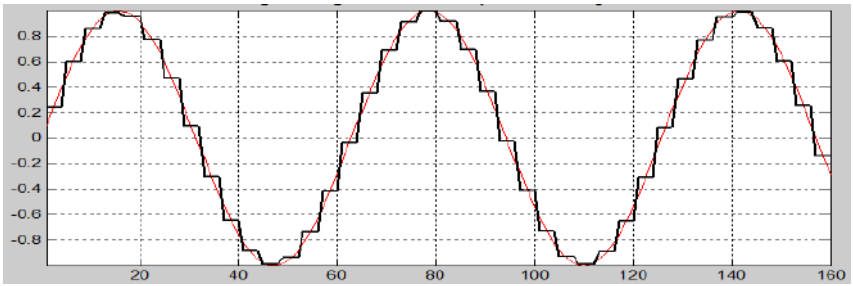
Хаарнинг бўлак- ўзгармас вейвлет модели [4]:

$$f(x)\cong\sum_{n=0}^{\infty}C_n\phi_n(x) \tag{2}$$

1-жадвал

Т/р	Вейвлет турлари	Абсолют хатолик	Нисбий хатолик
1.	Хаарнинг бўлак-ўзгармас вейвлетлари	0,20317	0,75872 %

Келтирилган модел асосида *y=sinx* функцияни олиб бўлак-ўзгармас вейвлетларида рақамли ишлаш амалга оширилди (1-расм).



1-расм. *y=sin(x)* функцияни бўлак- ўзгармас вейвлетларида рақамли ишлаш.

Бўлак-ўзгармас ва бўлак-чизиқ вейвлетларида интерполяциялаш хатолигини келтирамиз.

$[a,b]$ да аниқланган $f(x)$ узлуксиз функция берилган бўлсин [2]. $[a,b]$ сегментни

$$a \le x_0 < x_1 < \ldots < x_i < \ldots < x_n \le b$$

тугун нуқталарга ажратиб оламиз.

$$h = x_{i+1} - x_i = const \quad (3)$$

h- тугун нуқталар орасидаги масофа.

Хар хил даражали полиномлар учун интерполяциянинг методик хатоликларини аниқлаш формулалари мавжуд. Масалан, нолинчи дааражадаги полиномлар учун (бўлак- ўзгармас функциялар учун) хатоликни баҳолаш формуласи қуйидагича ифодаланади[5]:

$$\left|P(x) - f(x)\right| \le \frac{1}{2}\max\left|f'(x)\right|h \quad (4)$$

ХУЛОСА

Хаар-вейвлет ўзгартириш ёрдамида $y=f(x)$ функцияни бўлак- ўзгармас вейвлетида интерполяциялаш моделини қуриб 1-жадвалга кўра унинг хатоликларини баҳолаш амалга оширилди. Хулоса қилиш мумкинки Хаар-вейвлет ўзгартириш модели содда бўлганлиги учун функцияни интерполяциялашда кўпроқ фойдаланилади. Келажакда ушбу усулдан сигналларга рақамли ишлов бериш, сигналларни сиқиш, фильтрлаш, ўзгартириш ва сигналлардан шовқинларни ажратиш масалаларини ечишда ижобий натижалар олиш мумкин.

Адабиётлар

[1] Акимов П.А., Мозгалева М.Л., "Некоторые элементы кратномасш табного вейвлет-анализа. часть 2. анализ и синтез", Вестник МГСУ, 2012, № 8, 60–65

[2] Астафьева Н.М. Вейвлет-анализ: Основы теории и примеры применения // Успехи физических наук, 1996, т.166, № 11. С. 1145– 1170.

[3] Ахметханов Р.С., Дубинин Е.Ф., Куксова В.И., "Применение вейв-лет преобразований для анализа экспериментальных данных", Проблемы машиностроения и автоматизации, 2012, №4, 39–45

[4] Зайнидинов Х.Н. Методы и средства обработки сигналов в кусочно полиномиальных вейвлетах. // «Ташкент», 2015. 70 стр.

[5] Зайнидинов Х.Н., Сплайны в задачах цифровой обработки сигналов //Ташкентский университет информационных технологий-Т.: «Fan va texnologiya», 2015, 208 с.

CREDIT CARD FRAUD DETECTION USING MACHINE LEARNING

Muhmmad Umar[a], Qaiser ali[b] Imran Ahmed. [a]Riphah international university lahore. [b]Dept. Of Computer science, Lahore, Pakistan.

Abstract: *Credit card fraud is frequent and can lead to significant financial loss. credit card fraud detection, data mining cases, are questioned for two main reasons: fraudulent behavior in common practice, and second, credit card fraud data collection. variable selection and data set identification techniques. this article investigates the performance of logistic regression on disproportionate data, close neighbors, and credit card fraud. each trick is created using a set of engine training models, which is chosen by the best diagnostic method. three techniques are used for raw and pre-generated data. this project is implement in python. the work of this technology is estimate on the basis of accuracy, duration and level of permanent diagnosis. results of this shows of best accuracy for logistic regression and random forest classifier, classifiers are 99.92%, and 99.96% respectively. the comparative results show that logistic regression performs better than naïve bayes and knearest neighbour techniques.*

Keywords: *machine learning, logistic, regression random forest classifier.*

INTRODUCTION

Credit card fraud is the major problems in today's business. However, in order to effectively prevent fraud, it is important to understand the mechanisms that implement fraud. Credit card fraud is a great way to get rid of fraud. Many procedures have been adopted. To sum up, credit card fraud is defined as: If a person uses another person's card owner card holder not know about issuer no card. - Except for use, the card holder has no connection with the card holder or issuer, and does not contact the cardholder or pay for the purchase. This perform in the following ways:

- illegal fraud (intentional misuse) uses unauthorized accounts and / or personal information
- Use your account illegally or for personal gain. no
- Misrepresentation of account information to obtain goods and/or services.

Credit fraud is divided into two types:

- Offline.
- Online.

Offline fraud is done using an original card stolen from your store or call center. In most cases, the card issuer can block the card before using it for fraud. Online scams are not available for the web, phone purchases or card holders. You only need the card information, no manual signing or card printing at the time of purchase.

Location: Purchase made from different location

Items you buy: If you deviate from your regular buying pattern or time

Frequency: Make a large number of transactions in short period of time
Amount: Suddenly if the costly items are purchased.

lITERATURE REVIEW

1. You Dai, et. al: This article discusses the forest algorithm used to identify fraud. Random forests consist of random trees and wagon-based random forests. They describe 91.96% and 95.77% and accuracy respectively. The second type of paper is better than the other.
2. Suman Arora: In this article, many machine learning monitoring algorithms apply to 70% training and 30 test test datasets. Random Forest,Decision Tree, Beginner Algorithm and KNN (compare to each other), ie 94.59%, 95.27%, 94.59%, 93.54%, 93.24%,90.54% and 94.25%. To summarize this article, the highest rating for SVM is 0.5360 FPR and the lowest rating for the overall range is 0.0335.
3. Kosemani Temitayo Hafiz: In this paper, they describe flow chart of fraud detection process i.e., data Acquisition, data pre-processing, Exploratory data analysis and methods or algorithms are in detail. Algorithms are K- nearest neighbour (KNN), random tree, AdaBoost and Logistic regression accuracy are 99.92%, 94.32%, 57.73% and 98.24% respectively.

RESEARCH QUESTIONS

1) How is credit card fraud detected?
2) How to control large amount of people truncations through Credit Card?
3) Is it possible to find out Credit Card Fraud Detection Using Machine learning?

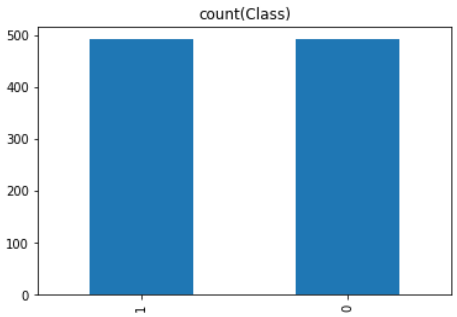
METHODOLOGY

MACHINE LEARNING AND ITS ALGORITHMS

A. Machine learning

Machine learning is a set of methods that can automatically identify patterns

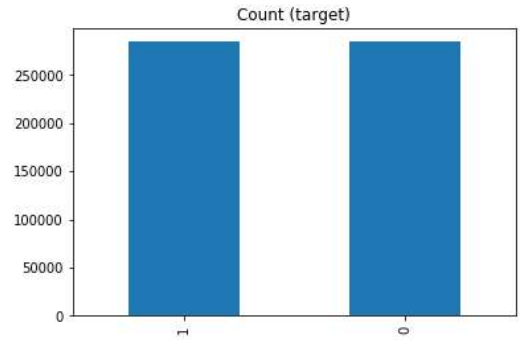
```
Random under-sampling:
1    492
0    492
Name: Class, dtype: int64

: <matplotlib.axes._subplots.AxesSubplot at 0x1e05963c7f
```



in your data and then use those patterns to predict future results or make other types of decisions under certain conditions. There are a number of algorithms in machine learning, which allow the machine to be based on a machine that can understand its current state and make the right choices. Machine learning works independently and decides for itself. There are two basic types of machine learning: supervised learning and non-supervised learning. This is called learning to learn. When models are applied to new datasets, learn from the training dataset, create models, and provide predictive results. Examples of learning to monitor, such as important trees, are the basis of the proposal. Unskilled education: Non-social education is a place where there is only input data and no matching variables. The main task of informal education is to write your own class labels. An unbiased survey algorithm can be used to detect correlations between statistics to determine whether the data produces grouping features. This group is known as clusters. Unsupervised learning can be also described as cluster analyses‖. K Means Clustering, KNN etc. are the examples of unsupervised learning.

Using target the value

```
Random over-sampling:
1    284315
0    284315
Name: Class, dtype: int64
```



A. Selected online dataset

Analyze this data by grouping it into multiple transactions. In addition, analysis of key components of PCA is performed to transform the data into two-dimensional space. The dataset had a transaction with European card holders

through credit cards in September 2013, which took two days, 492 out of 284,807 fake. Dataset is too inconsistent. In all trades, the positive class (fraud) is 0.172%. It only includes numeric input variables that are the result of PCA changes. Unfortunately, due to privacy concerns, we are unable to provide real data and a lot of background information about this data. The V1, V2, and V28 attributes are the main components acquired by the PCA, and the only features not changed by the PCA are 'time' and 'quantity'. The 'Time' function includes the last seconds between each transaction and the first transaction in the data set. The 'money' feature is a transaction count, and this feature can be used as an example of cost-sensitive learning.

C. Amount per transaction by classis

There are to type of transaction are made by the difference of two types

- Normal
- Fraud

Normal that is real transactions made by the card holder by its own

Fraud transactions that is made by theft or unknown person that is not made by card holder and thy are not now bot that transactions

C. Transaction VS Amount classis

Transaction time that are doing in that time not in pattern time means that is not made by organal card holder on the basis of time check real and fraudent transactions

This data set train on 201109 samples, validate on 56962 *samples*

D. Select algorithm for implementing

The literature reviews many algorithms on the detection of fraud. The questionnaire base includes 22, logistic regression, and more. The nearest neighbor is better than any other algorithm that detects fraud.. randomforest classfier ,logistic regression are used in this articals and implementing the fraud detection using these models

1) RANDOM FOREST CLASSIFIER

The Random Forest Classifier Classification. The algorithm organizes the theory of bias. The Indiana Ada algorithm Yang Sedrana Dan Sangat Kot.

The accuracy of Random Forest Classifier

```
[[113714     10]
 [    41    158]]
Accuracy: 99.96%
```

Basic Theorem: The Prayer Theorem Cummington Tragedy Cateca Peristua Lane MoundP (A/B) = (P (B/A) P (A)) / P (B)

Where,

P (A) – Priority of A P (B) – Priority of B

P (A/B) – Posteriori priority of B

Naïve Bayes algorithm is easy and fast. This algorithm need less training data and highly scalable.

2) LOGISTIC REGRESSION

Logistic regression using the functional approach to estimate the probability of a binomial response based on one or more changes (attributes). Find the best parameter for a nonlinear function called sigmoid.

The vector Z is the input data and the optimal coefficient is W. This coefficient is multiplied by each factor and then added to get a number that determines the classification of the target category. A sigmoid value greater than 0.5 is considered 1. 0 otherwise. The optimization method is used to train the classification and find the most suitable criteria. Gradient climbing (9) and modified stochastic gradient climbing compatibility methods experimented to evaluate for their performance in classifier.

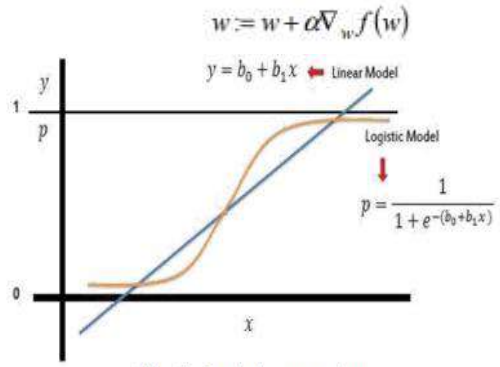


Fig. 1. Logistic regression

```
[[170557     29]
 [   105    194]]
Accuracy: 99.92%
```

This algorithm is similar to the linear regression algorithm. However, linear regression is used for predictive values and logistic regression is used to classify.

- Linear regression classified as
- Binomial – 2 Possible types (i.e. 0 or 1) only

results

After implementing the algorithm, the logistic regression gives the highest accuracy. The timeframe for logistic regression is quite high, but in this case accuracy is primarily considered as a result. Random forest classifier result 99.96%, logistic regression 99.92%, Beas, Logistic Relativity. Therefore, the results of the comparisons show that random forest classifier is more efficient than sophisticated bias and near-peripheral techniques. Then you can use logistic regression technology to detect credit cards.

REFERENCES

[1] J. Hand, G. Blunt, M.G. Kelly, and N.M. Adams, "Data Mining for Fun and Profit," Statistical Science, vol. 15, no. 2, pp. 111-131, 2000.

[2] "Statistics for General and On-Line Card Fraud," http://www.epaynews.com/statistics/fraud.html, Mar. 2007.

[3] S. Ghosh and D.L. Reilly, "Credit Card Fraud Detection with a Neural-Network," Proc. 27th Hawaii Int?l Conf. System Sciences:Information

Systems: Decision Support and Knowledge-Based Systems, vol. 3, pp. 621-630, 1994.

[4] M. Syeda, Y.Q. Zhang, and Y. Pan, "Parallel Granular Networks for Fast Credit Card Fraud Detection," Proc. IEEE Int?l Conf. FuzzySystems, pp. 572-577, 2002.

[5] S.J. Stolfo, D.W. Fan, W. Lee, A.L. Prodromidis, and P.K. Chan, "Credit Card Fraud Detection Using Meta-Learning: Issues and InitialResults," Proc. AAAI Workshop AI Methods in Fraud and Risk Management, pp. 83-90, 1997.

[6] S.J. Stolfo, D.W. Fan, W. Lee, A. Prodromidis, and P.K. Chan, "Cost-Based Modeling for Fraud and Intrusion Detection: Results fromthe JAM Project," Proc. DARPA Information Survivability Conf. and Exposition, vol. 2, pp. 130-144, 2000

[7] Abdelhalim, A, and I Traore. "Identity Application Fraud Detection using Web". International Journal of Computer and Network Security1, no. 1 (October 2009): 31-44.

[8] E. D. Y. Sahin, "Detecting credit card fraud by decision trees," in

[9] Proceedings of the international multiconference of engineers and

[10] computer science, Hong Kong, 2011.

[11] 18. P. M.,. S. H. Geoffrey F.Miller, "Designing Neural networks using

[12] genetic algorithms," [Online]. Available:

[13] M. H. O. Ekrem duman, "Detecting credit card fraud by genetic

[14] M. K. A. N. Alireza Pouramirarsalani1, "Fraud detection in E-

[15] Banking by using the hybrid feature selection and evolutionary

[16] A. T. P. K. M. G. P. N. Priya Chougle, "Genetic K- Means Algorithm

[17] A. O. A. S. A. O. John o. Awoyemi, "credit cars fraud detection using

[18] science and information technologies, vol. 6, pp. 1724-1727, 2015.

[19] O. O. O. A.,. W. Stephen Fashoto, "Hybrid Methods for credit card

[20] fraud detection," Kampala International University, Kampala,

[21] Uganda, University of Abuja, Nigeria, Redeemer's University, Ede,

[22] Osun State, Nigeria, [Online]. Available:

[23] E. Aji M. Mubarek, "Multilayer perceptron neural network technique

[24] (Caelen, Multiple perspectives HMM-based feature engineering for credit, 2019)

МОДЕЛИРОВАНИЕ ПРОЦЕССА РАСПРОСТРАНЕНИЯ ВРЕДНЫХ ВЫБРОСОВ В АТМОСФЕРЕ С УЧЕТОМ ВЕРТИКАЛЬНОГО ПРОФИЛЯ ВЕТРА

Равшанов Н., Ахмедов Д.Д., Научно-инновационный центр информационно-коммуникационных технологий при Ташкентском университете информационных технологий имени Мухаммада ал-Хоразмий, Узбекистан

Аннотация. *В работе рассматривается влияние вертикального профиля ветра на процесс распространения вредных выбросов в атмосфере на основе разработанного математического обеспечения и привлечения возможностей современных ГИС-технологий. Приводятся результаты вычислительных экспериментов по определению характеристик воздушного потока на заданных высотах при различных значениях приземной скорости ветра и коэффициента шероховатости подстилающей поверхности, а также результаты численных расчетов, демонстрирующих степень этих параметров на распределение концентрации взвешенных аэрозольных примесей.*

Ключевые слова: *загрязнение атмосферы, уравнения переноса и диффузии, вертикальный профиль ветра, степенной закон, коэффициент шероховатости, ГИС.*

Abstract. *The paper deals with study the influence of the vertical wind profile on the process of the spread of harmful emissions in the atmosphere based on the developed mathematical tools and GIS technologies. The results of computational experiments on determination the air flow characteristics at given heights and at various values of the ground-level wind speed and the roughness coefficient are presented in the paper, as well as the results of numerical calculations demonstrating the influence of these parameters on the distribution of suspended aerosol impurities in the considered region.*

Keywords: *atmospheric pollution, transport and diffusion equations, vertical wind profile, power law, roughness coefficient, GIS.*

ВВЕДЕНИЕ: Возможность достаточно точной оценки степени рассеивания загрязнений в атмосфере, в первую очередь, опирается на изучение аэродинамической структуры воздушного потока [1].

Вблизи земной поверхности скорость движения воздушной массы всегда ниже за счет воздействия силы трения о шероховатости подстилающей поверхности – неровности рельефа, тип почвы, растительность и строения. С удалением от земной поверхности влияние силы трения уменьшается и скорость ветра возрастает. Такое распределение скорости ветра по высоте относительно земной поверхности называется вертикальным профилем ветра, и всегда определяется для каждой конкретной местности.

Опытное измерение вертикальных профилей ветров выполняется с помощью средств аэрологического радиозондирования или лидарного сканирования [2]. Однако, при решении различных прикладных задач, связанных с атмосферными процессами, обычно применяют математические методы описания пространственной изменчивости ветра. Математические модели, построенные с учетом фактических метеорологических данных и орографии местности, позволяют достаточно точно воспроизводить параметры воздушного потока в заданной расчетной области.

Основной целью данной работы является изучение, с использованием возможностей ГИС-технологий, влияния вертикального распределения скорости ветра, в зависимости от шероховатости наземного покрова, на процесс распространения вредных выбросов в атмосфере.

ПОСТАНОВКА ЗАДАЧИ: Непосредственно сам процесс переноса и диффузии вредных примесей в атмосфере с учетом метеорологических факторов, рельефа местности и других внешних возмущений, может быть описан следующим многомерным дифференциальным уравнением в частных производных [3]:

$$\frac{\partial\theta}{\partial t}+u\frac{\partial\theta}{\partial x}+v\frac{\partial\theta}{\partial y}+\left(w-w_g\right)\frac{\partial\theta}{\partial z}+\sigma\theta=\frac{\partial}{\partial z}\left(\kappa\frac{\partial\theta}{\partial z}\right)+\mu\Delta\theta+Q\delta\left(\bar{r}-\bar{r}_0\right). \quad (1)$$

Здесь θ – концентрация вредных веществ в атмосфере; x, y, z – система координат; u, v, w – скорость ветра по трем направлениям; w_g – скорость осаждения частиц; σ – коэффициент поглощающей способности атмосферы; μ, κ – соответственно, коэффициенты диффузии и турбулентности; Q – мощность точечного стационарного источника выброса; δ – функция Дирака; Δ – оператор Лапласа.

Для сокращения объема производимых вычислений на ЭВМ, имеет смысл редуцировать фактическую область решения задачи от традиционного параллелепипеда до полусферы. Таким образом, уравнение (1) можно рассмотреть в области

$D=\left\{R_1<r<R_2,\ \lambda_1<\lambda<\lambda_2,\ \eta_1<\eta<\eta_2, R_1<R_2, \lambda_2-\lambda_1<2\cdot\pi,\ \eta_2-\eta_1<\pi\right\}$,

представляющей собой полусферу с радиусом R, а само уравнение (1) и соответствующие краевые условия записать в следующем виде [3]:

$$\frac{\partial\theta}{\partial t}+\sigma\cdot\theta=\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\kappa\frac{\partial\theta}{\partial r}\right)+$$
$$+\mu\left(\frac{1}{r^2\sin\theta}\cdot\frac{\partial}{\partial\eta}\sin\eta\frac{\partial\theta}{\partial\eta}+\frac{1}{r^2\sin^2\eta}\frac{\partial^2\theta}{\partial\lambda^2}\right)+Q\delta\left(\bar{r}-\bar{r}_0\right); \quad (2)$$

$$\left.\begin{aligned}
&\alpha_{1,1} r^2 \kappa \frac{\partial \theta}{\partial r} - \beta_{1,1} \cdot \theta = -\mu_{1,1} \qquad \text{при} \quad r = R_1; \\
&-\alpha_{1,2} r^2 \kappa \frac{\partial \theta}{\partial r} - \beta_{1,2} \cdot \theta = -\mu_{1,2} \qquad \text{при} \quad r = R_2; \\
&\alpha_{2,1} \sin\eta \frac{\partial \theta}{\partial \eta} - \beta_{2,1} \cdot \theta = -\mu_{2,1} \qquad \text{при} \quad \eta = \eta_1; \\
&-\alpha_{2,2} \sin\eta \frac{\partial \theta}{\partial \eta} - \beta_{2,2} \cdot \theta = -\mu_{2,2} \quad \text{при} \quad \eta = \eta_2.
\end{aligned}\right\} \qquad (3)$$

МЕТОД РЕШЕНИЯ: Для решения задачи (2)-(3) нами был разработан консервативный численный алгоритм, основанный на разностном методе [3-5]. Данный численный алгоритм не вызывает особых затруднений.

Как было сказано выше, при решении задачи прогнозирования распространения вредных выбросов в атмосфере, большое значение имеет описание вертикального профиля скорости ветра. К настоящему времени, разработано большое количество математических методов расчета профилей скорости ветра [6]. Наиболее широко распространены аппроксимации в виде степенной и логарифмической функций [7].

Описание вертикального профиля ветра в виде степенной функции имеет вид

$$V_1 = V_0 \left(\frac{H_1}{H_0} \right)^k, \qquad (4)$$

где V_1 – искомая скорость ветра на высоте H_1; V_0 – известная скорость ветра на контрольной высоте H_0; k – безразмерный показатель степенного закона изменения скорости ветра с высотой, являющийся либо константой, либо переменной величиной [7].

При решении прикладных инженерных задач, значение k зачастую принимается постоянным. Между тем, в местах, где на поверхности присутствуют структуры, препятствующие движению воздушной массы, использование стандартного значения $1/7$ может давать весьма ошибочные оценки. Например, даже при условии безразличной стратификации, над открытой поверхностью водоемов значение $k = 0{,}11$ является более подходящим, чем $k = 0{,}143$.

По этой причине нами применялись возможности ГИС-технологий. А именно, для территории рассматриваемого региона из открытых баз данных в Интернет выполнялся импорт снимков радарной топографической съемки (SRTM) и мультиспектральных спутниковых изображений [17, 18] (рис. 1). Первые обеспечивают данные цифровой модели рельефа, а вторые – возможность автоматического распознавания типов земной поверхности и построения тематической карты степенного показателя k.

Во многих случаях (особенно для небольших областей) тематические карты целесообразно формировать вручную, путем обводки и обозначения

участков поверхности в режиме цветности снимков RGB. В результате создаются так называемые шейп-файлы – простой формат для хранения геометрического местоположения и атрибутивной информации географических объектов, которые в нашем случае представлены полигонами (рис. 2).

Классификация шероховатости подстилающей поверхности осуществлялась согласно таблице в [8], значения коэффициента изменяются от 0,0 – для поверхности воды до 0,44 – для больших городов с высокими зданиями и небоскрёбами.

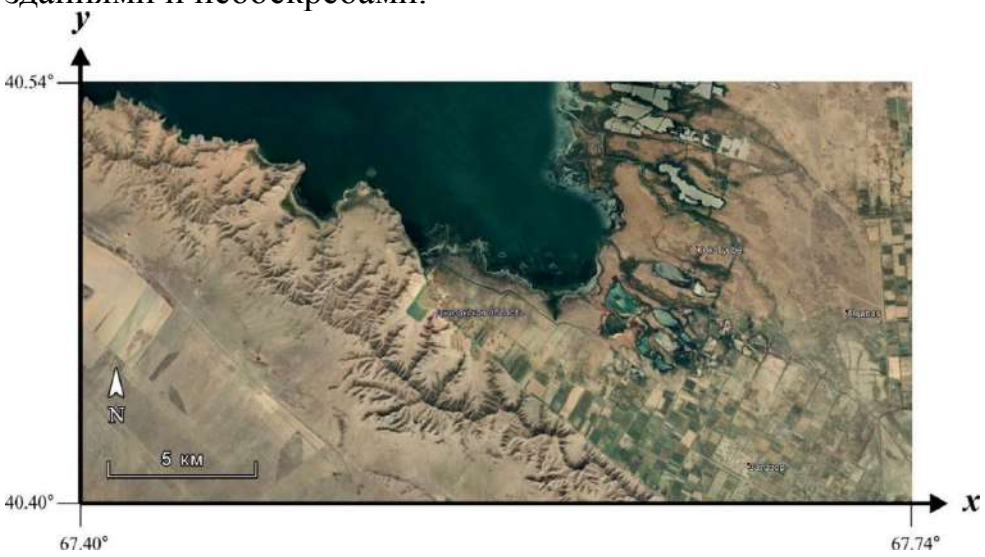


Рис. 1. Спутниковое изображение части территории Джизакской области РУз вблизи озера Айдаркуль.

Рис. 2. Тематическая карта шероховатости наземного покрова рассматриваемой области, приведенной на рис. 1.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ: На основе разработанного математического инструмента (модель, численный алгоритм) и возможностей обработки данных в ГИС, было создано программное средство на языке Python.

Вычислительные эксперименты для исследования характеристик воздушного потока на различных высотах были проведены при различных значениях приземной скорости ветра и коэффициента шероховатости подстилающей поверхности.

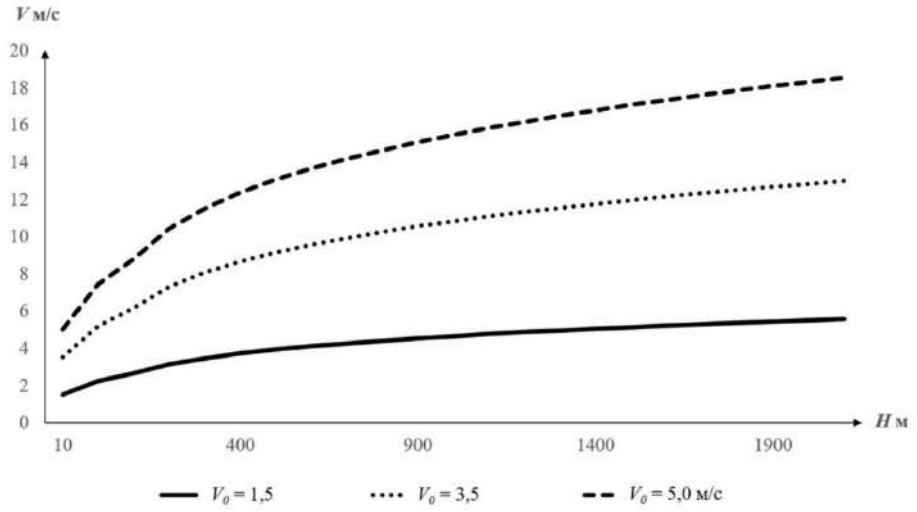


Рис. 3. Изменение скорости ветра по высоте пограничного слоя атмосферы при коэффициенте шероховатости поверхности $k = 0,245$ и различных приземных скоростях ветра.

Из результатов численных расчетов (рис. 3) ожидаемо следует, что на увеличение скорости воздушной массы по высоте пограничного слоя атмосферы, главным образом влияет скорости ветра у повехности земли (здесь уровень флюгера – 10 м) и класса шероховатости наземного покрова.

Вычислительными экспериментами (рис. 4) установлено, что с ростом коэффициента шереховатости (при $V_0 = 3,5$ м/с) скорость ветра с высотой растет по логарифмическому закону. Такой рост происходит до определенной высоты $H \approx 2000$ м, а после скорость ветра изменяется незначительно и влияние эффекта шероховатости исчезает.

Как можно видеть из кривых на рис. 4, характер движения воздушной массы при проходе над водной гладью ($k = 0$) практически не изменяется с высотой, а в случае прохода вдоль участков суши со значительной неоднородностью рельефа ($k = 0,405$) – наблюдается градиент скорости.

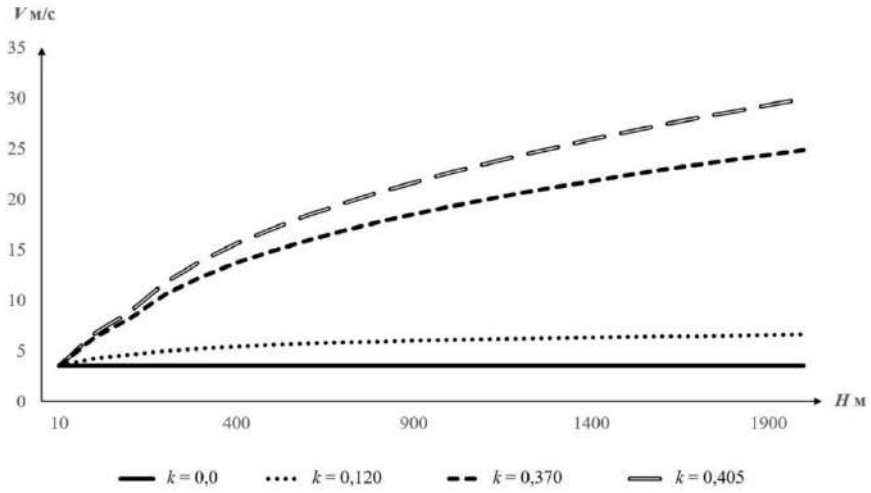


Рис. 4**.** Изменение скорости ветра по высоте пограничного слоя атмосферы при $V_0 = 3{,}5$ м/с.

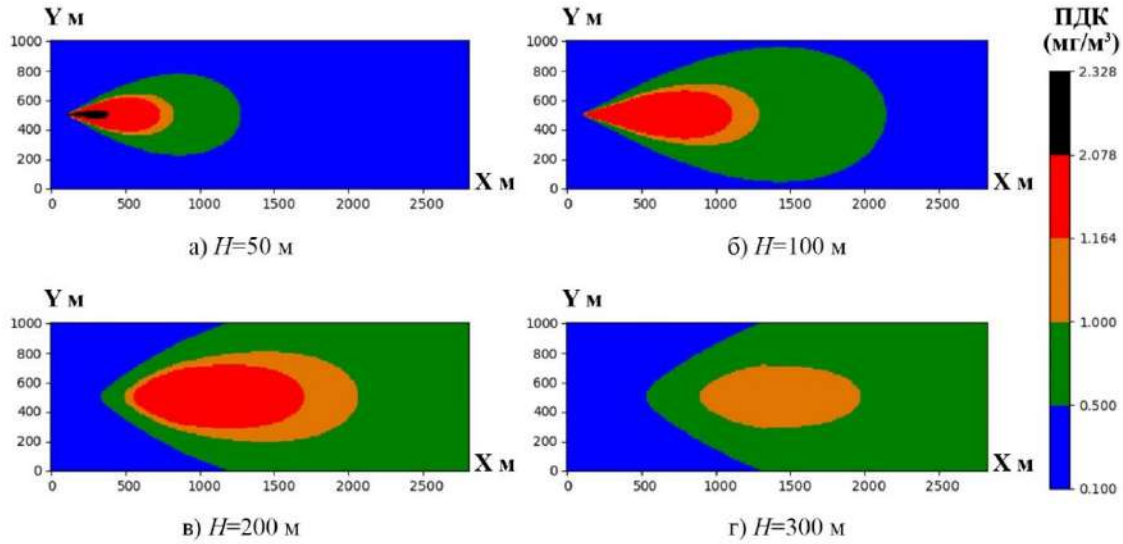


Рис. 5. Изменение концентрации нитрата аммония со временем на различных высотах при скорости ветра у поверхности земли $V_0 = 0{,}2$ м/с.

Анализ результатов численных расчетов показывает, что распределение концентрации взвешенных аэрозольных примесей в рассматриваемом регионе в значительной мере зависит от горизонтальной составляющей скорости воздушной массы в пограничном слое атмосферы. С ростом горизонтальной скорости ветра область воздействия концентрации взвешенных веществ пропорционально расширяется со временем, а непосредственно уровень концентрации аэрозольных частиц уменьшается по мере удаления от источника выброса вредных веществ (рис. 5).

Также, проведенными расчётами была установлена зависимость процесса переноса и диффузии аэрозольных частиц от таких факторов как: метеорологические условия, свойств частиц и распределение коэффициента турбулентности по вертикали.

4. Заключение

Знание и учет закономерности изменения скорости ветра с высотой в зависимости от типа наземного покрова и коэффициента шероховатости подстилающей поверхности являются необходимым условием для адекватности результатов математического моделировании процесса рассеяния вредных выбросов в атмосфере. Как показывает анализ опубликованных научных работ и полученные результаты данного исследования – использование постоянного значения данного параметра для различных типов земной поверхности может приводить к существенным ошибкам в оценках режима течения ветрового потока даже при условии безразличной стратификации атмосферы.

Рассмотренный подход к построению полей ветра на разных высотах, основанный на применении существующих математических методах и ГИС технологий, позволяет достаточно точно описать процесс переноса и диффузии примесей в атмосфере и оценить распределение концентрации вредных частиц.

Литература

[1] *Млявая Г.В. Влияние параметров шероховатости под стилающей поверхности на скорость ветра // Известия Академии Наук Молдовы : Науки о жизни. – 2014. – № 2(323). – С. 181-187.*

[2] *Сапунов М.В., Мельникова И.Н., Донченко В.К., Самуленков Д.А., Кузнецов А.Д. Сопоставление вертикальных профилей скорости и направления ветра, полученных на основе лидарных и аэрологических измерений // Современные проблемы дистанционного зондирования Земли из космоса. – 2016. – Т. 13, № 1. – С. 149-160.*

[3] *Ravshanov N.2, Akhmedov D. Air Quality Dispersion Modeling in Spherical Coordinates // Techno-Societal 2018 / Eds. P. Pawar et.al. - Springer, 2020. – P. 149-156.*

[4] *Sharipov D., Muradov F., Akhmedov D. Numerical Modeling Method for Short-Term Air Quality Forecast in Industrial Regions // Applied Mathematics E-Notes. – 2019. – № 19. – P. 575-584.*

[5] *Ravshanov N.3, Shafiev T.R. Nonlinear mathematical model for monitoring and predicting the process of transfer and diffusion of fine-dispersed aerosol particles in the atmosphere // Journal of Physics: Conference Series. – 2019. – Vol. 1260. – P. 102014.*

[6] *Werapun W., Tirawanichakul Y., Waewsak J. Wind Shear Coefficients and their Effect on Energy Production // Energy Procedia. – 2017. Vol. 138. – P. 1061-1066.*

[7] *Симанков В.С., Шопин А.В., Бучацкнй П.Ю. Моделирующий комплекс поступления энергии для оперативного управления автономными фотоветроэнергетическими системами // Труды ФОРА. – 2002. – № 7. – С. 13-20.*

Голубенко Н.С., Довгалюк С.И., Фельдман А.М., Худик В.Б. О зависимости скорости ветра от высоты с учетом рельефа местности / ПКТБ КОНКОРД. – 2005. – URL: https://bit.ly/34tP8kK.

FACIAL EXPRESSION RECOGNITION – A CLUSTERING BASE APPROACH

Azfar Wajid Siddiqui. Department of Computer Science and Technology Riphah International University, Lahore, Pakistan.

Abstract. *The automatic analysis of facial expressions has drawn the attention of several researchers in* ***several*** *fields* ***like*** *psychology,* ***computing****, neuroscience* ***within the*** *last years, the researchers have introduced and evaluated several approaches for face detection, recognition, and analysis. The* ***facial expression*** *recognition* ***is often*** *used for human-computer interface (HCI),* ***special effects*** *animation, and security. This paper describes* ***a*** *clustering-based feature extraction method for* ***facial expression*** *recognition. Facial expressions are the facial changes in response to a person's internal emotional states, intentions, or social communications. Facial expression analysis has been* ***a*** *research topic for behavioral scientists [1].*

Keywords: *Facial expression Recognition, HCI, Clustering based approach*

1. INTRODUCTION

Facial expression recognition delivers good information about human emotion and plays a crucial role in human communications. For intelligent and natural human-computer interaction, it's essential to recognize facial expression automatically. Various techniques are developed for automatic facial expression recognition, which differs in data used, feature extraction methods, and classifiers used [2].

For facial expression recognition from image sequences, optical flow estimation is usually wont to extract features. Here we go to describe facial expression recognition how useful for us. We use in medical science and Artificial intelligence. For facial expression recognition we use Cohn Kanade (CK +) shown in Fig 1.

Fig 1. Shows some expression from cohn – kanade data set.

1.1 Expression of Facial:

Dataset and extract the facial expression through Matlab Code. Matlab [3] is good software in Data science and Artificial Intelligence field. We use Facial expression Recognition Master Code For extract the features we got good results. But in future we provide more efficient results but in this research we want cluster based approach.

Clustering is that the task of dividing the population or data points into variety of groups such data points within the same groups are more almost like other data points within the same group and dissimilar to the info points in other groups. When we apply clustering algorithms on facial expression dataset then we use Weka. The Simple K-Means algorithm uses Euclidean distance measure to compute distances between instances and clusters [1].

To perform clustering, select the "Cluster" tab in the Explorer and click on the "Choose" button. This results in a drop list of available clustering algorithms. Some Facial Expression Recognition results shown in Fig 2 [1].

Fig 2. Emotion-specified facial expression (posed images from dataset [1] 1.anger, 2. Happy, 3. Fear, 4. Sadness, 5. Surprise

2. LITERATURE REVIEW

2.1 Principals of Facial Expression Analysis:

Facial expression analysis includes both measurements of facial motion and recognition of expression. The overall approach to automatic expression analysis (AFEA) [4] consists of three steps: face acquisition, facial data extraction and representation, and expression recognition. After the face is found, the subsequent step is to extract and represent the facial changes caused by facial expressions.

In facial feature extraction for expression analysis, there are mainly two sorts of approaches: geometric feature-based methods and appearance-based methods. The geometric countenance presents the form and locations of facial components including (Nose, Mouth, Head, Lips) Fig3.

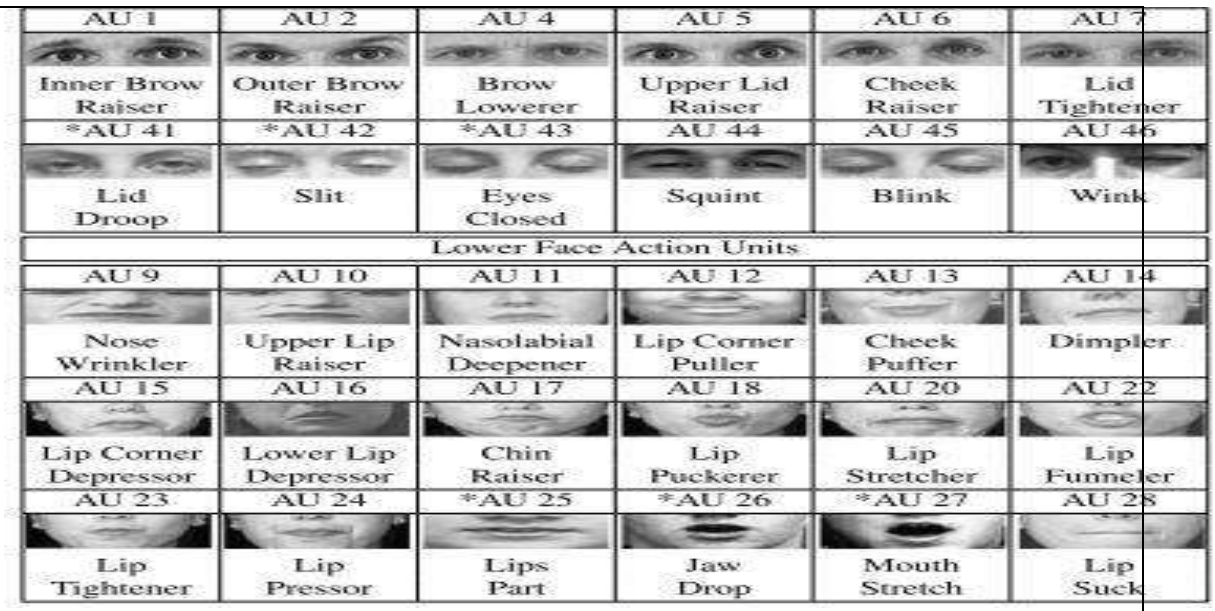


Fig 3. Location of facial expressions

The facial components or facial feature points are extracted to make a feature vector that represents the face geometry. Counting on the various facial feature extraction methods, the consequences of in-plane head rotation and different scales of the faces can be eliminated by face normalization before the feature extraction or by feature representation before the step of expression recognition. Facial expression recognition is that the last stage of AFEA systems.

2.2 Clustering based discriminant analysis:

Each image is often treated as a feature vector by concatenating the rows of the image together, using each pixel as one feature. Thus, each image are often represented by an n-dimensional vector. Where n is that the number of pixels in each image. We use Cluster different models like (Use Training Set, Supplied Test Set, Percentage Split, Classes to cluster split) and apply different algorithms. But we shows only K-Means algorithms results [1].

3. METHODOLOGY

3.1 Research Questions:

Different techniques are researched for automatic facial expression recognition, which differs in data used, feature extraction methods [5], and classifiers used. For facial expression recognition from image sequences, optical flow estimation is usually wont to extract features. Here we go to describe facial expression recognition how useful for us. We use in medical science and Artificial intelligence [5].

3.2 Statements of Problems:

Although humans recognize facial expressions vertually without effort or delay, reliable Expression recognition by machine is still a challenge. The problems that have haunted the pattern recognition community at large still require

attention for facial expression recognition we use Cohan Kanade (CK +) [6] Dataset and extract the facial expression through Matlab Code.

3.3 Methodology of FER:

There are a number of problems in facial expression recognition due to the position of facial expression across the human population and to the context-dependent variation even for the same individual. Even we human beings may make mistakes. On the other hand, facial expression recognition by computer is very useful in many applications such as human behavior a particular way and human-computer interface.

An automatic facial expression recognition system needs to solve the following problems detection and location of faces in a cover scene, facial feature extraction, and facial expression classification. Fig 3 [6].

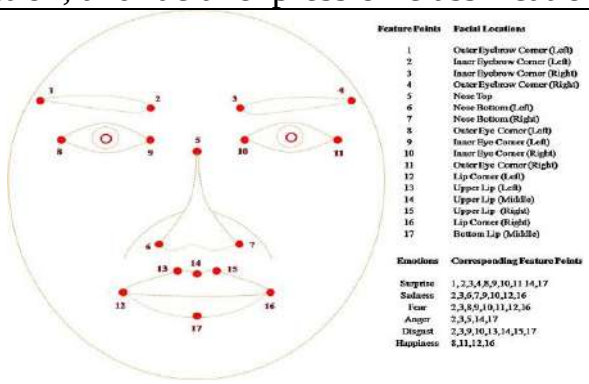


Fig 3. Clustering based landmark detection

3.4 Research Methodology.

- Input face Image
- Face detection
- Face feature localization
- Feature extraction
- Feature Selection
- Emotion Classification
- Facial expression result

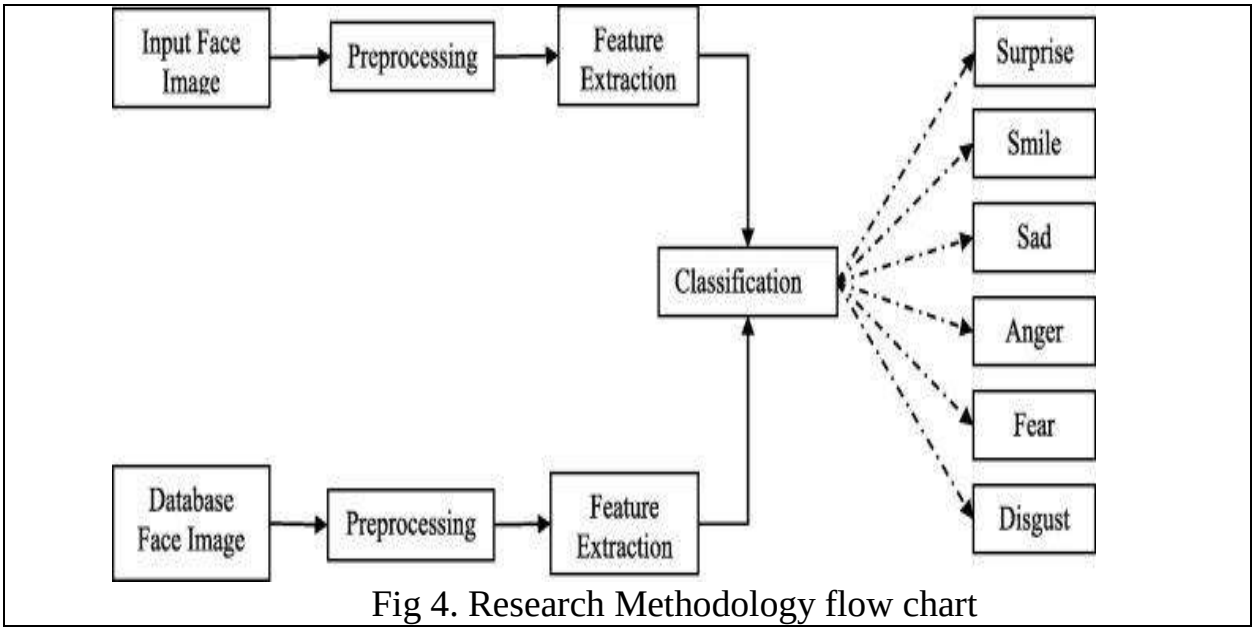


Fig 4. Research Methodology flow chart

3.5 Research Methodology exclusive representation

Face detection has been studied by many researchers, and it to give the impression of being that most successful systems are based on neural networks. Once a face is detected in the image directly related to something region is featured out, and is usually simplified to have the same size (for example, the same distance between two eyes) and the same gray level. Facial feature extraction attempts to find the most exclusive representation of the face images for recognition.

4. RESULTS AND DISCUSSION

We use (CK+) Dataset and put in to Weka and use K-Means Clustering [7] and apply different models shown in Fig5. And Fig6.

```
Final cluster centroids:

Attribute

=============================================================
emotion
pixels
Usage



Time taken to build model (full training data) : 0.07 seconds

=== Model and evaluation on training set ===

Clustered Instances

0       173 ( 53%)
1       154 ( 47%)
```

Fig5. Cluster Base Approach (Training set)

```
Final cluster centroids:

Attribute

=============================================================
emotion
pixels
Usage




Time taken to build model (percentage split) : 0.01 seconds

Clustered Instances

0       68 ( 61%)
1       44 ( 39%)
```

Fig6. Cluster Base Approach (Percentage Split)

Two classes predefined in cluster base approach 0 and 1 we have 147 instance, 68 instance on training set and 44 instance are test set. Results shown in fig 5 and fig 6.

Facial expression creates hesitation gives only observations, not authentic result.

Several factors limit the effectiveness of facial expression recognition.

- Image quality
- Image size
- Face angle
- Processing and storage

5. CONCLUSION AND FUTURE WORK

Facial expression recognition are highly defined in other research paper. This paper is a high-level research based project overview of automatic facial expression recognition. Due to length restrictions, only a small sample of recognition techniques is explicitly referred to. Further details can be found in the cited references.

The next section gives an overview of facial expression recognition systems. Thereafter, some outstanding research problems are pointed out, and a summary of the overview is given. Facial expressions convey non-verbal actions [8], which play an important role in interpersonal relations. Automatic recognition of facial expressions can be an important component of natural human-machine interfaces .Some robots can also benefit from the ability to recognize expressions. Automated

analysis of facial expression for behavioral science, clinical practice and security purpose.

REFERENCES

[1]" Facial expression recognition: A clustering-based approach *Author links open overlay panel Xue-wenChenThomasHuang* https://doi.org/10.1016/S0167-8655(02)00371-9

[2]" Facial Expression Recognition for HCI Applications Fadi Dornaika Institut Géographique National, France Bogdan Raducanu Computer Vision Center, Spain

[3] "Expression Recognition Based on Image Feature Mahesh Kumbhar, Ashish Jadhav, and Manasi Patil

[4] "Automatic Facial Feature Extraction and Facial Expression Recognition S´everine Dubuisson, Franck Davoine, and Jean-Pierre Cocquerez Universit´e de Technologie de Compi`egne, BP 20529, 60205 Compi`egne, France {sdubuiss,fdavoine,cocquerez}@hds.utc.fr

[5] "A MATLAB based Face Recognition System using Image Processing and Neural Networks Jawad Nagi, Syed Khaleel Ahmed Farrukh Nagi Department of Electrical and Electronics Engineering Department of Mechanical Engineering Universiti Tenaga Nasional Universiti Tenaga Nasional Km7, Jalan Kajang-Puchong, 43009 Kajang, Malaysia Km7, Jalan Kajang-Puchong, 43009 Kajang, Malaysia jawad@uniten.edu.my, syedkhaleel@uniten.edu.my farrukh@uniten.edu.my

[6] " The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression Patrick Lucey1,2 , Jeffrey F. Cohn1,2 , Takeo Kanade1 , Jason Saragih1 , Zara Ambadar2 Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, 152131 Department of Psychology, University of Pittsburgh, Pittsburgh, PA, 152602 patlucey@andrew.cmu.edu, jeffcohn@cs.cmu.edu, tk@cs.cmu.edu, jsaragih@andrew.cmu.edu ambadar@pitt.edu Iain Matthews Disney Research, 4615 Forbes Ave, Pittsburgh, PA 15213 iainm@disneyresearch.com.

[7] " International Journal of Advances In Computer Science and Cloud Computing, ISSN: 2321-4058 Volume- 1, Issue- 1,May-2013 Enhanced K-Means Based Facial Expressions Recognition System 39 ENHANCED K-MEANS BASED FACIAL EXPRESSIONS RECOGNITION SYSTEM 1TANVI SHEIKH, 2 SHIKHA AGRAWAL CSE Department CSIT Durg1 , Assistant Professor CSE Department CSIT Durg Email:tanvisheikh1990@gmail.com1 , Shikha.Agrawal030@gmail.com.

[8] "Recent advancements in facial expression recognition systems: A survey Publisher: IEEE Uzair Asad ; Nirbhay Kashyap ; Shailendra Narayan Singh

COMPARING DIFFERENT ALGORITHM OF CLASSIFICATION DATASET USING WEKA TOOL

Muhammad Arshad SarFraz Student of MSCS , Department of Computer science at Riphah international University Lahore, Pakistan.

Abstract. *Classification is on the most common issue in the filed of data extraction . I am using two diiferent classification algorithms random forest and simple logistic on weka tool . The performance of random forest and simple logistic is tested on different dataset of classification . I choose these dataset from different sites.*

Keywords. Classification, data mining, random forest, simple logistic.

1. INTRODUCTION

Begin In data mining the samples are taken from large datasets and by the method of combining statistical and artificial intelligence with management of data .In a lot of disciplines it is nearby field in today's world.It is acquire as the extension of technology and the demand of coherent Data evaluation .The goals of data mining is not grant stric law by examining the data set but it is used for determined predictions during examining a small set of the dataset.

In this paper I am going to use different algorithm on different datasets to compare the performance of these algorithms . As I choose Random forest and Simple logistic to perform some practicals using Weka tools. As It is easy to use and it has a friendly user interface.

Recently Data mining is noticed in the information of knowledge industry due to huge amount of data and the future demand is the conversion of raw data into meaningful info. Data extraction is the proper and complete application of IT based of DB technology and it has proov to be a great business value and step by step it become driving force in the profession of telecom and other power industries .

2. RELATED WORK

In recent time some studies have been managed on the base of different performance of the Random forest and simple Logstic[1].

As for Classification, the most famous measure of performance is the classification error , or equivalently , the accuracy of classification [2]. How ever some question are raised that the use of accuracy as an appropriate measure of algorithm performance .Ramdom forest is a algorithm of classification that contain trees algorithm like decision tree , Random forest etc . Use the insertion and random ness of feature during he construction of each individual tree try to creat an non related tree forest whoes predictions are more accurate than single tree.

As I have studies some paper about comparison of weka algorithm in which they compare decision tree and MLP algorithm and they found that the performance of MLP is better then decision tree . The Simple logistic work best wen the numbers

of noise variables is less then or equal to the numbers of explanatory variables and random forest experiences a higher true and false positive rate like the numbers of explanatory Data set.

3. RESEARCH QUESTIONS

Q1. Which algorithm of machine learning give us better accuracy or performance using five diffifent datasets.?

Q2. which algorithm give better result either Random Forest or Simple logistic on the basis of different datasets?

4. METHODOLOGY

4.1 Dataset

Table 1. Information of Datasets

Datasets	Instance	Attributes	NOC	Type
Australian	690	15	2	Numeric
hypothyroid	3772	30	4	Numeric
Segment	2310	20	2	Numeric
Spam base	4601	58	2	Numeric
ionosphere	351	35	2	Numeric

The 1st data is use is Asutrailian this data set have 690 instances and 15 attributes its have 2 numbers of classes and its type is numerical .

The second Dataset is Hhypothyroid and this data set have 3772 instances and 30 attributes ,Its have 4 numbers of classes and its type is numeric .

The third Data set is use is segment and this data set contains 2310 instances and 20 attributes , Its contain 2 classes and it's a numerical type dataset.

The Fourth Data set is use is Spambase and this data set contains 4601 instances and 58 attributes , Its contain 2 classes and it's a numerical type dataset [2].

The Fifth Data set is use is ionosphere and this data set contains 351 instancesc and 35 attributes , Its contain 2 classes and it's a numerical type dataset.

4.2. Weka Interface

Weka is a famous tool of machine learning algorithm and its is written in java develeopment some its cause a lot of troubles with jdk. Weka contain algorithms of classification, clustring, and it has some visualization tools and some predictive model with a friendly interface and it is easy to use [1].

A non-java version is accessible which include TLC or TK as a front end software that is used to apply algorithm on different programming languages and it is a C utilizes and Make file based system for performing ML observations.

I am using weka (3.9) that is based on java version. It is easy to use and used in many industrial area and mostly it is used in educational area students use this tool for perform different algorithms and visualization purpose .

Weka hold up quality of data extraction activities like Data processing , Data grouping , Classification, clustering, visualization and some other features.All methods of Weka tool is depend on hypothesis that data can be accessed through flat files or different relation ships .

4.3. Classification Algorithm Random Forest

Classification is the biggest part of machine learning . The capacity of accurately classifying experiments is particularly important for different business apps such as assumption of weather, that a specific user will purchase a product or it can assume that the loan will be not paid .

Data science give us a lot of algorithms for classification like Logistic Regression , SMV , MLP decision tree[1]. etc but you will see Random Forest at the top of classification ranking.This algorithm depend on decision making algorithm . It has assorted organized solution trees otherwise it's a forest contiain trees nothing else.

4.4. Classification function Simple Logistic

Logical regression is a method of "statistical training" that is recorded as a method of (supervised)machine learning (ML) derived from the work of "classification" [3]. In the last 20 years, she has won a very high award, especially for her self-help skills that make her a record holder in the financial industry. Logistic regression and other plans of the common line layout.

Logic bost with simple regression function as based learner is used to fit this model. logistic algorithm if 5 different dataset and I compare the result of Random forest with Simple Logistic and conclude that either Random forest is best or Simple Logistic is best on the base of there performance.

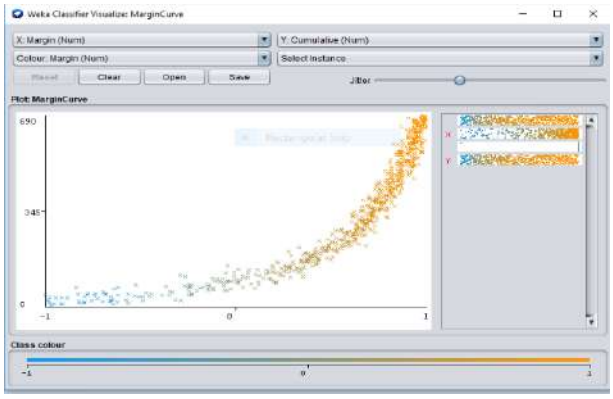
5. RESULTS

In thi section I will share my result and I will compare both the algorithms and check which on is best for give better accuracy. I Performed both algorithms of classification on 5 ifferent data set and got different result.

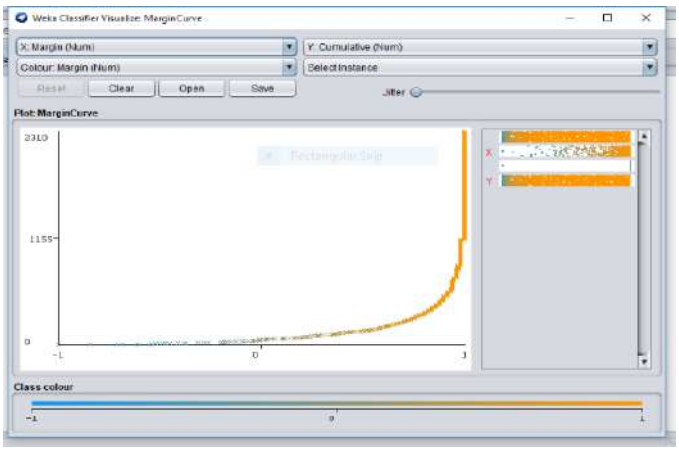
Table 2 . Accuracy on Australian dataset

S#	Parameters	Random Forest	Simple logistic
1	TP rate	0.862	0.852
2	Fp rate	0.143	0.142
3	Precision	0.862	0.856
4	Recall	0.862	0.852
5	F-Measure	0.862	0.853
6	MCC	0.721	0.706
7	ROC area	0.922	0.914
8	PRC area	0.911	0.903

In this table 2, I am comparing different parameters like TP rate , Fp rate, MCC , ROC area and PRC area of two different algorithm. First one is Random Forest and second is Simple logistic algorithms on the data set of Australian. As you can see in table accuracy of Random forest is better then Simple logistic.



In fig 1, I am represent the margin curve of Random forest on Australian Data set.



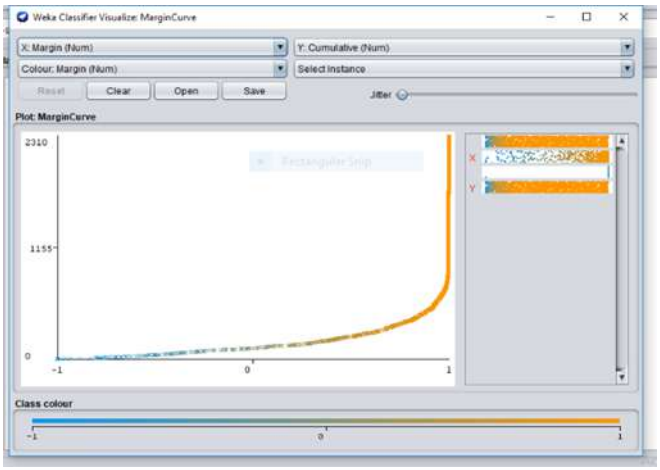
In fig 2, I am represent the margin curve of Simple logistic on Australian Data set

Table 3. Accuracy on Segment dataset

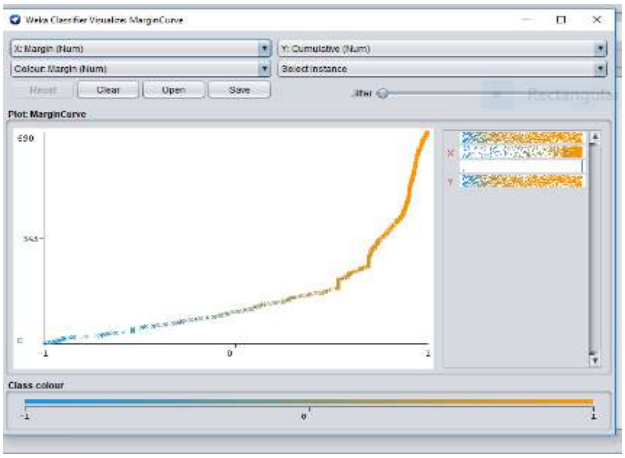
S#	Parameters	Random Forest	Simple lpgistic
1	TP rate	0.955	0.926
2	Fp rate	0.054	0.088
3	Precision	0.955	0.926
4	Recall	0.955	0.926
5	F-Measure	0.955	0.926
6	MCC	0.906	0.844
7	ROC area	0.987	0.973
8	PRC area	0.987	0.969

In this table 3, I am comparing different parameters like TP rate , Fp rate, MCC , ROC area and PRC area of two different algorithm. First one is Random Forest and second is Simple logistic algorithms on the

data set of Australian. As you can see in table accuracy of Random forest is better then Simple logistic.



In fig 3 , I am represent the margin curve of Simple logistic on Segment Data set

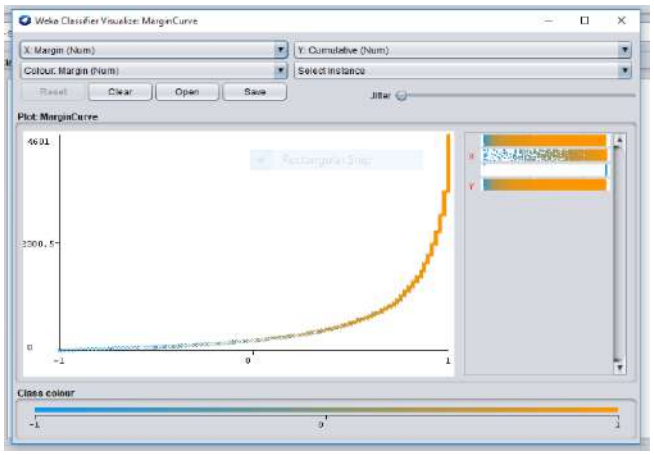


In fig 4 , I am represent the margin curve of Random forest on Segment Data set

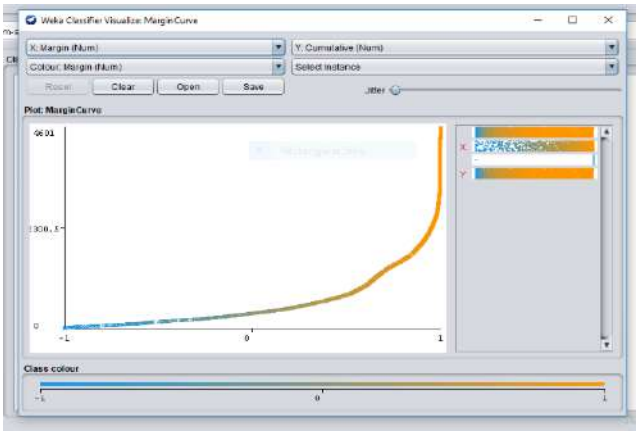
Table 4. Accuracy on spambase dataset

S#	Parameters	Random Forest	Simple lpgistic
1	TP rate	0.955	0.926
2	Fp rate	0.054	0.088
3	Precision	0.955	0.926
4	Recall	0.955	0.926
5	F-Measure	0.955	0.926
6	MCC	0.906	0.844
7	ROC area	0.987	0.973
8	PRC area	0.987	0.969

In this table 4, I am comparing different parameters like TP rate , Fp rate, MCC , ROC area and PRC area of two different algorithm. First one is Random Forest and second is Simple logistic algorithms on the data set of Australian. As you can see in table accuracy of Random forest is better then Simple logistic



In fig 5 , I am represent the margin curve of Simple logistic on Spambase Data set.

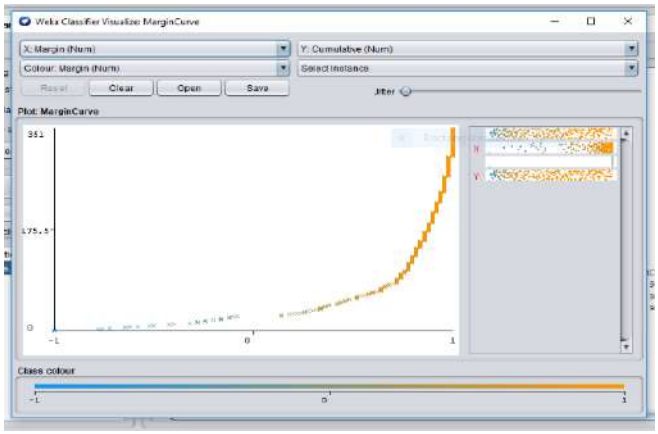


In fig. 6, I am represent the margin curve of Random Forest on Spambase Data set.

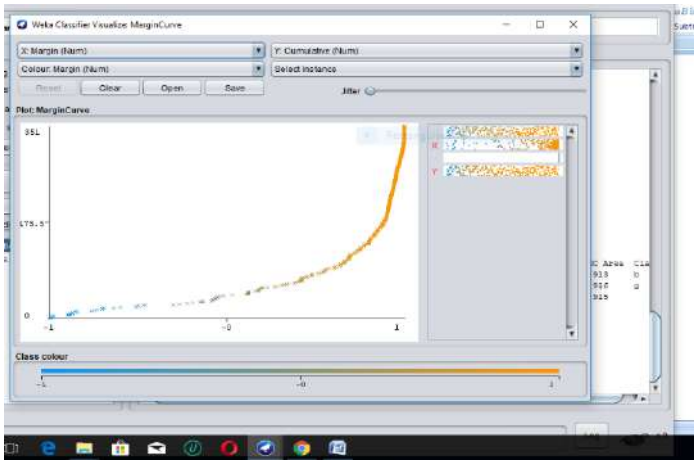
Table 5. Accuracy on Ionosphere dataset

S#	Parameters	Random Forest	Simple lpgisti
1	TP rate	0.979	0.951
2	Fp rate	0.004	0.008
3	Precision	0.979	0.952
4	Recall	0.979	0.951
5	F-Measure	0.979	0.951
6	MCC	0.975	0.943
7	ROC area	0.999	0.996
8	PRC area	0.997	0.981

In this table 5, I am comparing different parameters like TP rate , Fp rate, MCC , ROC area and PRC area of two different algorithm. First one is Random Forest and second is Simple logistic algorithms on the data set of Australian. As you can see in table accuracy of Random forest is better then Simple logistic.



In fig 7, I am represent the margin curve of Random forest on Ionosphere Data set

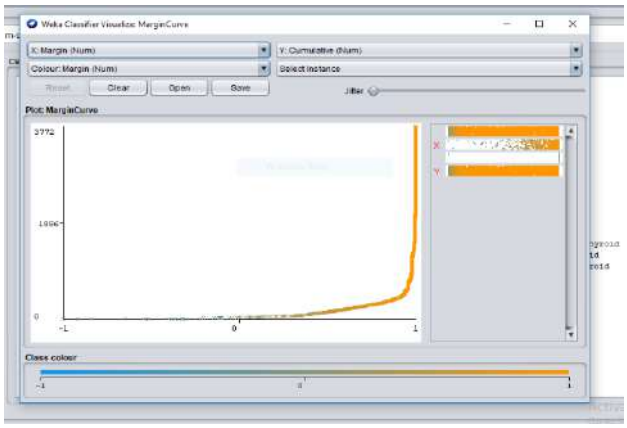


In fig 8 , I am represent the margin curve of SL on Ionosphere Data set

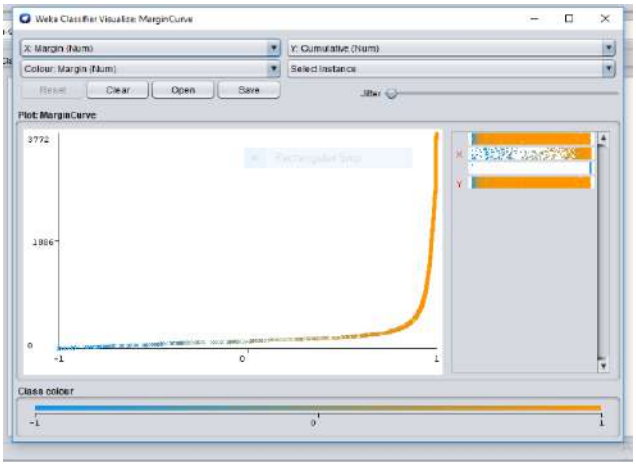
Table 6. Accuracy on hypothyroid dataset

S#	Parameters	Random Fores	Simple lpgisti
1	TP rate	0.993	0.967
2	Fp rate	0.038	0.295
3	Precision	0.997	0.864
4	Recall	0.993	0.967
5	F-Measure	0.997	0.985
6	MCC	0.963	0.827
7	ROC area	0.999	0.984
8	PRC area	0.999	0.983

In this table 6, I am comparing different parameters like TP rate , Fp rate, MCC , ROC area and PRC area of two different algorithm. First one is Random Forest and second is Simple logistic algorithms on the data set of Australian. As you can see in table accuracy of Random forest is better then Simple logistic.



In fig 9, I am represent the margin curve of Random Forest on hypothyroid Data set.



In fig 10, I am represent the margin curve of SL on hypothyroid Data set

Table 7. Accuracy measure of Random forest and Simple logistic

Sr #	Datasets	Random Forest	Simple logistic
1	Austrailian	86.239%	85.2174%
2	hypothyroid	99.3107 %	96.6861 %
3	Segment	97.8788 %	95.1082 %
4	Spam base	95.501 %	92.5886 %
5	ionosphere	92.8775 %	88.3191 %

In table num 7, I am comparing the result of two different algorithm which I used on five different dataset using in weka tool. Here you can see the performance of Random Forest is better then simple logistic and Random forest give better accuracy then simple logistic.

6. CONCLUSION

In this paper I use different parameters for measuring accuracy of Random Forest and Simple logistic algorithms. The size of dataset is vary from others that's why we got variation in out results. By review Australian, Hypothyroid and other datasets it is clear that Random forest is better algorithm. We see that mostly Random forest is best algorithm . On other hand if we see simple logistic algorithm its accuracy is not better than Random forest but its algorithm have high value of rather then others.

7. FUTURE SCOPE

Many algorithms can be selected and compare different attributes , if some how use real dataset from industries and use different algorithms to get different result improve the accuracy or you can apply clustering algorithms on it .

REFRENCES

[1] R. Arora, "Comparative analysis of classification algorithms on different datasets using WEKA," International Journal of Computer Applications, vol. 54, no. 13, 2012.

[2] P. Singh and A. Surya, "Performance analysis of clustering algorithms in data mining in weka," International Journal of advances in engineering & technology, vol. 7, no. 6, p. 1866, 2015.

[3] R. Duriqi, V. Raca, and B. Cico, "Comparative analysis of classification algorithms on three different datasets using WEKA," in 2016 5th Mediterranean Conference on Embedded Computing (MECO), 2016: IEEE, pp. 335-338.

LINKED IN PROFILES AND JOBS DATASET PREDICTION AND CLUSTERING WITH RAPIDMINER

Salma Yousaf, Department of Computer Science, Riphah International University Lahore Campus, Pakistan.

Abstract. *In this paper I am going to use the Rapid Miner tool which is one of the best data mining tool. It shows us the prediction of our data in a precise and easily readable form. I will use the linked in profiles and jobs dataset and apply different machine learning algorithms to check the results and accuracy in Data Mining and Predictive tool Rapid Miner. In this paper I will analyze the Linked in profile and jobs dataset. Data preprocessing, data prediction, clustering and outliers will be performed on this dataset. Visual data analysis, numerical analysis and hidden relations of Linked in profiles and jobs will be shown here.*

Keywords. *Data science, Data mining, LinkedIn Profiles and Jobs, RapiMiner, Prediction, K-means Clustering, X-means Clustering, Generalized Linear Model, Deep Learning, Decision Tree.*

1. INTRODUCTION

Rapid miner is one of the best tool which is used for data mining. Data mining techniques are used now days in many fields of technology.

Data mining is the repetitive technique in which different methods results are processed automatically or manually. Data mining is use in many fields. It explores different and unique ideas and gives outstanding results. [1]

In this paper dataset of Linked in profiles have been taken. Different data mining techniques are applied using Rapid miner. Prediction of data and clustering is applied on this dataset. We see the generalized linear model prediction, generalized linear model charts and simulation prediction. Then we also see the Deep learning prediction, charts, Decision tree model. Then we also see different clustering algorithms of clustering on our dataset. Here two types of Clustering algorithms are applied on the dataset: K-means clustering and X-means clustering.

2. PROBLEM STATEMENT

To apply the data mining techniques to predict the different columns values, attributes and apply the clustering algorithms on the dataset of Linked in profiles and jobs dataset.

3. LITERATURE REVIEW

1. RapidMiner is used to analyze a Pima Diabetes Data Set, which gather the information of the patients which are suffering from diabetes and also gather information of the patients which are not suffering from diabetes. Preprocessing of this dataset is performed in order to construct the diabetes prediction model. [1]

2. In this paper, we discuss how the Web of Linked Data can be mined using the full functionality of the state of the art data mining environment RapidMiner. [2]

3. Rapid-I's RapidMiner is software package supports all steps of data mining process. [3]

4. Number of datasets collected in data mining without any goal and aim. In these sets many dataset have some type of errors or dirt and missing values. These type of datasets need to be cleaned and missing values should be removed to get accurate results. [4]

5. The use of Linked Open Data in data mining has been proposed before, and implementations as RapidMiner extensions as well as proprietary toolkits exist. [5]

4. METHODOLOGY

Predict

In Predict process I selected the column of age Estimation in task and then move towards the next step where target is prepared. Here visualization of age estimation is shown in the form of graph which is more precise and easy to read. Then in next step Input will be selected. Then model type selection is occurred here. Generalized linear model, Deep Learning and Decision Tree is applied on this dataset. Now here we will discuss the all applied methods and models and the names of these methods and models are:

1. Generalized Linear Model
2. Deep Learning
3. Decision Tree

Generalized Linear Model

The generalized linear model is a malleable generalization of usual linear regression which permit for response variables that have error distribution models other than a normal distribution.

Generalized Linear Model Prediction /chart

Generalized linear model prediction chart is shown in the fig 1.4. It shows the prediction of estimated age of employees in colored format. So one can easily understand the prediction value.

Fig 1.4

Generalized Linear Model Simulation Prediction

Generalized linear model simulation prediction value is shown for the required dataset. Fig 1.5 shows the second largest predictive value.

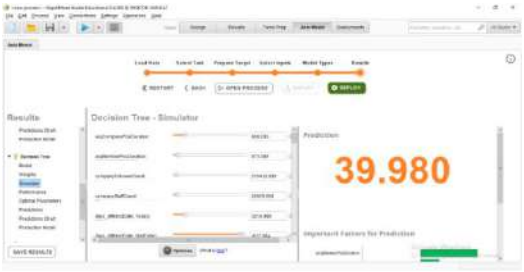


Fig 1.5

Deep Learning Simulator Prediction

Deep learning simulator prediction values is largest than generalize linear model and decision tree. Fig 1.7 shows the prediction value of the deep learning simulator.

Fig 1.6

Deep Learning Prediction Chart

Deep learning prediction chart is shown in the fig 1.9. It shows different estimated values in colored form.

Fig 1.7

Decision Tree Model

Now here is the decision tree model of the dataset of linked in profiles and jobs. Its too lengthy tree and cannot be shown in one frame. Fig 2.0 shows this decision tree.

Fig 2.0

Decision Tree Simulator Prediction

Fig 2.1 shows the predictive value of decision tree simulator. This predictive value is smaller than generalized linear model and deep learning model predictive values.

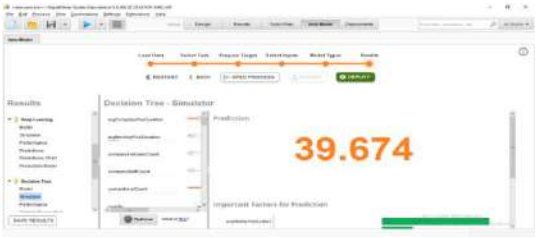


Fig 2.1

Clustering

I selected the country and age estimation as input task in clustering. Two types of clustering algorithms are applied: k means and x means. Two clusters are made namely cluster 0 and cluster 1.

K means Cluster tree

K means Cluster tree is shown in the fig 2.5 for the dataset of linked in profiles and jobs. It's a huge cluster tree and difficult to shown in one frame clearly.

Fig 2.5

X means cluster tree

X means Cluster tree is shown in the fig 2.8 for the dataset of linked in profiles and jobs. It's a huge cluster tree and difficult to shown in one frame clearly.

Fig 2.8

4. RESULTS

We see the different results in the prediction and in the clustering. In prediction deep learning model shows the largest values than generalized linear model and decision tree. In clustering K means and X means algorithms are used. They show the different results. The more precise one is X means in the case of this dataset.

5. CONCLUSION/FUTURE WORK

To get the more precise and accurately predicted value of the linked in profile and jobs dataset other models and clustering algorithms should be applied. More accurate and better results can be produce when applied the different models which are heavier and take long time to process. We can also change some attributes in the dataset and can increase the number of entries for more accurate results. Results can also be more accurate when there is no distortion in the dataset. All missing values and dirt should be removed from the dataset for better output.

REFERENCES

[1] Hofmann, M., & Klinkenberg, R. (Eds.). (2016). RapidMiner: Data mining use cases and business analytics applications. CRC Press.

[2] Rapid-I (2008), Interactive Design. Products: RapidMiner, http://rapid.com/content/view/13/69/lang.en/

[3] Larose, D. T., & Larose, D. T. (2006). Data mining methods and models (Vol. 2). Hoboken (NJ): Wiley-Interscience.

[4] Paulheim, H., & Fümkranz, J. (2012, June). Unsupervised generation of data mining features from linked open data. In Proceedings of the 2nd international conference on web intelligence, mining and semantics (pp. 1-12).

[5] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter*, *11*(1), 10-18.

СОЗДАНИЕ ИНФОРМАЦИОННОЙ МОДЕЛИ ДЛЯ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ ПРОЦЕССА ТРАНСПОРТА ЗАГРЯЗНЯЮЩИХ ВЕЩЕСТВ В АТМОСФЕРЕ

Хамдамов Р.Х., Равшанов З.Н., Научно-инновационный центр информационно-коммуникационных технологий при Ташкентском университете информационных технологий имени Мухаммада ал-Хоразмий, Узбекистан

Аннотация. *С целью проведения комплексного исследования, мониторинга и прогнозирования экологического состояния воздушных бассейнов в промышленных регионах нашей республики, а также для разработки адекватных математических моделей объекта исследования в работе приведены основные факторы и их отклики, влияющие на процесс переноса и диффузии загрязняющих веществ в атмосфере. Так как, один из основных факторов, влияющих на изменение концентрации вредных веществ в атмосфере промышленных регионов являются свойства аэрозольных соле-пылевых частиц, то необходимо учитывать их физико-механические характеристики и определить основные силы, воздействующие в ходе процесса переноса и диффузия. В статье приведены основные выводы по проведенным численным расчетам на ЭВМ, в которых учитывались эрозия почвы в зависимости от скоростей воздушной массы атмосферы, размеры и плотности частиц, а также силы, действующих на них. Анализ численных расчетов показал, что основной движущей силой ветровой эрозии почвы является вихревая подъемная сила, которая на порядок превышает касательное напряжение и плотность атмосферы, но не играет большой роли в изменении подъемной силы частиц. Она растет по экспоненциальному закону в зависимости от изменения скорости воздушной массы атмосферы.*

Ключевые слова: *математическая модель, перенос и диффузия вредных веществ, погодно-климатические факторы, численный алгоритм.*

Abstract. *In order to conduct a comprehensive study, monitoring and forecasting of air basins and the ecological state of industrial regions of the republic and protect them from technogenic factors, as well as to develop adequate mathematical models of the object of study, the main factors and their responses to the transport and diffusion of pollutants in atmosphere. In the work, since one of the main factors influencing the change in the concentration of harmful substances in the atmosphere of industrial regions is aerosol salt-dust particles, it is necessary to take into account their physical and mechanical properties and determine the main forces acting on them during mathematical modeling and conducting a comprehensive study of the transport process and diffusion of these particles. The main conclusions of the performed numerical calculations on a computer are carried out in the article, where soil erosion is taken into account depending on the air velocity of the atmosphere, the size and density of particles, as well as the*

forces acting on them. An analysis of numerical calculations showed that the main driving force of wind soil erosion is the vortex lifting force. Which is an order of magnitude greater than the tangential stress and density of the atmosphere, but does not have a large role in changing the lifting force of particles, it will grow exponentially depending on changes in the speed of the air mass of the atmosphere.

Keywords: *mathematical model, transport and diffusion of harmful substances, weather and climate factor, numerical algorithm*

1. ВВЕДЕНИЕ. Многолетний анализ статистических данных показывает, что нарушения сложившегося экологического баланса в регионах с высокой антропогенной нагрузкой и, в первую очередь, устойчивое загрязнение атмосферы, становятся причиной массы проблем – рост онкологических, аллергических, астматических и других заболеваний среди населения, ухудшение плодородности сельскохозяйственных угодий, состояния флоры и фауны.

Поэтому, проведение комплексных исследований, мониторинг, прогнозирование и анализ процесса транспорта и диффузии загрязняющих веществ в приземном слое атмосферы относятся к наиболее актуальным вопросам в проблеме охраны окружающей среды, в особенности в промышленных регионах.

Одним из эффективных инструментов мониторинга и прогнозирования экологического состояния промышленных регионов выступает математический аппарат, реализуемый в виде математической модели, вычислительного алгоритма и программного средства для проведения комплексных исследований с последующей поддержкой принятия управленческих решений.

Следует подчеркнуть, что разработки адекватного математического инструмента обязательно предваряется созданием информационной модели, в которой учитываются основные параметры, существенно влияющие на исследуемый объект.

Проблемы математического моделирования процессов переноса и диффузии вредных веществ в атмосфере во всем мире активно разрабатываются многими учеными. К настоящему времени, ими получены значительные результаты теоретического и прикладного характера. Научные исследования, направленные на разработку математических моделей и вычислительных алгоритмов для систем мониторинга и прогнозирования загрязнения атмосферы промышленных регионов, проводятся в передовых научных центрах и высших учебных заведениях мира, в число которых входят Karlsruhe Institute of Technology, Institute of Applied Geosciences (Германия), Centre for Ecology and Hydrology, Centre for Water Science (Великобритания), Finnish Meteorological Institute (Финляндия), National Institute of Hydrology, Waterloo Hydro Geologic (Индия), Институт атмосферы физики им. А.М.Обузова, Институт вычислительной математики и математической геофизики (Россия), Научно-исследовательский и проектно-

технологический институт «Атмосфера», Национальный университет Узбекистана, Ташкентский университет информационных технологий и др.

2. ПОСТАНОВКА ЗАДАЧИ. Исходя из сказанного выше, для исследования основных параметров и факторов, которые существенно воздействуют на процесс изменения концентрации аэрозольных мелкодисперсных частиц в атмосфере, рассмотрим математическую модель, описываемую с помощью уравнения в частных производных и соответствующих краевых и внутренних условий [1]:

$$\frac{\partial\theta}{\partial t}+u\cdot\frac{\partial\theta}{\partial x}+v\cdot\frac{\partial\theta}{\partial y}+\left(w-w_g\right)\cdot\frac{\partial\theta}{\partial z}+\sigma\cdot\theta= \\ =\mu\cdot\Delta\theta+\frac{\partial}{\partial z}\left(\chi\cdot\frac{\partial\theta}{\partial z}\right)+Q\cdot\delta(x,y,z); \tag{1}$$

$$\theta(x,y,z,0)=\overset{0}{\theta}(x,y,z); \tag{2}$$

$$-\mu\cdot\left.\frac{\partial\theta}{\partial x}\right|_{x=0}=\gamma(\theta-\theta_a);\ \ \mu\cdot\left.\frac{\partial\theta}{\partial x}\right|_{x=L_x}=\gamma(\theta-\theta_a); \tag{3}$$

$$-\mu\cdot\left.\frac{\partial\theta}{\partial y}\right|_{y=0}=\gamma(\theta-\theta_a);\mu\cdot\left.\frac{\partial\theta}{\partial y}\right|_{y=L_y}=\gamma(\theta-\theta_a); \tag{4}$$

$$-\chi\cdot\frac{\partial\theta}{\partial z}=\gamma(\beta\cdot\theta-F_0)\text{ при }z=0; \tag{5}$$

$$\chi\cdot\frac{\partial\theta}{\partial z}=\gamma\left(\theta\ -\theta_a\right)\text{ при }z=H. \tag{6}$$

Здесь θ - количество распространяющегося вещества; θ_0 - первичная концентрация вредных веществ в атмосфере; σ - коэффициент поглощения вредных веществ в атмосфере; δ - функция Дирака; t - время; x,y,z - координаты; μ- коэффициент диффузии; β - коэффициент взаимодействия с подстилающей поверхности; Q - мощность источников; F_0 - количество аэрозольных частиц оторвавшихся от шероховатости земной поверхности; χ - коэффициент турбулентности; γ - коэффициент для проведения граничного условия к размерному виду; θ_a - концентрация взвешенных веществ в соседних областях решаемых задач.

Во многих источниках величина F_0 является функцией от x,y,z,t, и определяется по экспериментальным данным в зависимости от метеорологических условий, свойств подстилающей поверхности, размера и плотности пылевых частиц. В данной статье значения F_0 будем вычислять в зависимости от скорости ветра у подстилающей поверхности земли, физико-механических свойств почвы и баланса сил действующих на частицы [1].

3. МЕТОД РЕШЕНИЯ И РЕЗУЛЬТАТЫ ВЫЧИСЛИТЕЛЬНЫХ ЭКСПЕРИМЕНТОВ. Из постановки задачи (1)-(6) видно, что для вычисления

объема уносимых с поверхности почвы частиц за счет эрозии, необходимо определить вихревую подъемную силу F_n, которая зависит от нескольких факторов: коэффициента подъемной силы, радиуса почвенной частицы, плотности воздуха, скорости потока за пределами пограничного слоя.

Как было рассмотрено в работе [2] определение подъемной силы и коэффициента подъемной силы основаны на экспериментальных определениях критической скорости потока, и его можно определить, если критической скорости соответствует именно вертикальный взлет частиц. Когда мелкодисперсная частица находится в состоянии неустойчивого равновесия или выступает под общей поверхностью, она не поднимается по вертикали, а начинает двигаться по горизонтали, и причиной ее движения будет лобовое давление, а не подъемная сила.

С учетом сказанного, и полагая, что частицы различаются между собой лишь крупностью, используя уравнение баланса сил в проекциях на вертикальную ось в момент отрыва потоком, нами была получена функциональный зависимость для вычисления объема уносимых с поверхности земли аэрозольных мелкодисперсных частиц.

Проведенные численные расчеты на ЭВМ показали, что этот процесс существенно зависит от коэффициента подъемной силы, шероховатости подстилающей поверхности земли, орографии местности и физико-механических свойств почвогрунта [3-4].

Численные расчеты были проведены для различных плотностей (от 1400 кг/м^3 до 1135 кг/м^3) и размеров частиц. Из анализа полученных расчетов видно, что при увеличении плотности частиц линейно растет сила их сцепления и при U равном критической скорости движение частиц начинается именно под действием подъемной силы, которая на порядок превосходит лобовую силу сопротивления.

Проведенные вычислительные эксперименты на ЭВМ показали, что один из основных факторов который существенно действует на изменение силы сцепления частиц – это скорость воздушной массы атмосферы. С ростом скорости воздушной массы атмосферы экспоненциально растет сила сцепления частиц. Особенно этот рост заметен при $U > 4$ м/с и аналогично, сила лобового сопротивления будет расти с ростом скорости воздушной массы атмосферы.

Вычислительными экспериментами установлено, что подъемная и лобовая сила сопротивления частиц будут расти в зависимости от роста скорости воздушной массы атмосферы. Экспоненциальный рост подъемной силы частиц особенно заметен, когда скорость ветра превышает критические значения. Как следует из численных расчетов и кривых, с ростом диаметра частиц экспоненциально растет их подъемная сила, а лобовая сила сопротивления незаметно растет с ростом диаметра частиц.

Еще один из существенных факторов, воздействующих на процесс отрыва аэрозольных частиц от подстилающей поверхности земли – это

плотность атмосферы, влияющая на силу лобового сопротивления. С ростом плотности атмосферы сила лобового сопротивления растет.

Из анализа расчетов следует, что по достижению скорости потока критического значения и при выровненной поверхности, частица начинается двигаться именно под действием подъёмной силы, так как она примерно на порядок превосходит силу лобового сопротивления. Численными расчетами установлено, что в случае для аэрозольных мелкодисперсных частиц, стесненных со всех сторон себе подобными, по достижении критической скорости ветра, именно подъёмная сила является единственным способом начала движения таких частиц по вертикали.

Анализируя результаты проведенных численных экспериментов можно прийти к выводу, что для процесса сальтации и подъема частиц с поверхности является вихревая подъемная сила. Тогда, можно прийти к выводу, что уносимый объем частиц с поверхности почвы за счет эрозии, непосредственно связано с силами, действующими на соле-пылевые частицы в результате турбулентного движения воздушной массы атмосферы в приземном слое. К этим силам относятся лобовая сила сопротивления частиц, подъемная сила частиц и сила сцепления частиц.

Численные расчеты так же были проведены при различных критических скоростях ветра у поверхности земли, влажности почвы, коэффициента шероховатости и других физико-механических параметров рассматриваемого процесса. Из анализа результатов численных расчетов следует, что с ростом влажности почвы по экспоненциальному закону растет сила коэффициента сцепления частиц, за счет чего существенно уменьшается объем выноса мелкодисперсных частиц в атмосферу.

Проведенными численными расчетами установлено, что изменение скорости ветра по вертикали зависит от скорости динамического трения. С ростом скорости динамического трения пропорционально растет вертикальная скорость ветра на поверхности земли, а в свободном слое атмосферы она остается неизменяемой по толщине.

Как следует из проведенных численных расчетов на ЭВМ с ростом коэффициента шероховатости земли, горизонтальная составляющая скорости ветра пропорционально уменьшается. Этот факт можно заметить при изменении толщины слоя $0 < h < 10$ м.

Одним из существенных параметров, воздействующих на процесс переноса и диффузии аэрозольных частиц в пограничном слое атмосфере, также выступают шероховатость и неоднородность поверхности земли, в следствии чего изменяются скорости воздушного потока атмосферы в зависимости от их значений.

Изменение скорости воздушной массы атмосферы в зависимости от шероховатости поверхности земли можно определить с помощью формулы:

$$V_1 = V_0 \left(\frac{H_1}{H_0} \right)^k ; \qquad (7)$$

где V_1 - скорость ветра на заданной высоте; V_0 - скорость ветра на известной высоте (H_0 = 5 м) ; H_0- заданная высота; H_1- высота измерения; k - эмпирический показатель шероховатости подстилающей поверхности.

На основе приведенного соотношения (7) проведены численные расчеты на ЭВМ.

Из анализа проведенных численных расчетов видно, что существенным параметром воздействующим на процесс изменения скорости ветра по вертикали являются коэффициент шероховатости, который изменяется в зависимости от высоты местности (орографии), шероховатости поверхности земли, а так же от наличия высотных объектов и жилых комплексов, находящихся в рассматриваемом регионе.

Коэффициент шероховатости изменяется от 0 до 0,45, на поверхности воды он равняется нулю, на полностью открытытом ландшафте с мягкой поверхностью типа взлётно-посадочных полос в аэропортах или со скошенной травой – 0,12; на открытых сельскохозяйственных землях с одиночными зданиями – 0,245; на сельскохозяйственных землях с отдельными зданиями и 8-ми метровыми оградами на расстоянии более 1250 м – 0,275; на сельскохозяйственных землях с отдельными зданиями и 8-ми метровыми оградами на расстоянии более 500 м – 0,30; на сельскохозяйственных землях с группами зданий и 8-ми метровыми оградами на расстоянии более 250 м – 0,335; в деревнях, малых городах, на сельскохозяйственных землях с отдельными зданиями и высокими оградами, лесом и резко пересечённой местностью – 0,370; в больших городах с высокими зданиями – 0,405; в мегаполисах с высокими зданиями и небоскрёбами – 0,440.

Вычислительными экспериментами установлено, что с ростом коэффициента шероховатости скорость ветра по вертикали растет по логарифмическому закону до определенной высоты 600-750 м, а далее она незначительно растет по высоте.

На основе разработанного математического инструмента (модель, численный алгоритм и программное средство) проведены вычислительные эксперименты на ЭВМ. Анализ полученных численных расчетов показал, что изменение и распространение концентрации взвешенных выбросов в рассматриваемом регионе существенно зависит от горизонтальной скорости воздушной массы атмосферы. С ростом горизонтальной скорости ветра область воздействия концентрации взвешенных веществ пропорционально расширяется со временем, а сама концентрация аэрозольных частиц уменьшается по мере удаления от источника выброса вредных веществ.

Численные расчёты показывают, что перенос и диффузия осуществляются в зависимости от метеорологических условий, от свойств частиц и существенно зависят от распределения коэффициента турбулентности по вертикали.

Из проведённых численных расчетов видно, что на объем уноса мелкодисперсных частиц с поверхности осушённой части Аральского моря существенно воздействуют:

а) составляющие скорости ветра на поверхности земли, с ростом этого параметра логарифмически растёт объем уносимых аэрозольных частиц в атмосферу. Кроме того, имеется зависимость от шероховатости поверхности земли и высоты объектов, расположенных в регионе и плотности зданий, находящих в районе;

б) растительность и влажность почвы рассматриваемого региона, с ростом этого показателя резко уменьшается количество выброса вредных веществ в атмосферу и их рассеивание в атмосфере;

с) коэффициент шероховатости земли, который является основным параметром для изменения скорости ветра по высоте слоя атмосферы.

Численными расчетами установлено, что существенными факторами, воздействующими на процесс эрозии почвы и транспорт соле-пылевых частиц из осушённой части моря в Приаральском регионе, являются: критическое значение скорости ветра у поверхности земли, влажность почвы, коэффициент шероховатости земли и физико-механические свойства породы.

Установлено, что с ростом влажности почвы по экспоненциальному закону растет сила коэффициента сцепления частиц в следствии чего существенно уменьшается объем выноса мелкодисперсных частиц в атмосферу.

Литература

1. *Равшанов Н., Тоштемирова Н., Мурадов Ф. Атмосферадаги аэрозол заррачаларининг кучиши ва диффузияси масаласини ечимини мавжудлиги ва ягоналигини тадкик килиш // Ҳисоблаш ва амалий математика муаммолари. – Тошкент, 2017. - №1 54-68 б. (05.00.00; №23)*

2. *Ravshanov N., Shertaev M., Toshtemirova N. A Mathematical Model for the Study and Forecast of the Concentration of Harmful Substances in the Atmosphere // American Journal of Modeling and Optimization. – 2015.- Vol. 3. No. 2. – Pp.35-39*

3. *Равшанов Н., Таштемирова Н., Шафиев Т. Нелинейная математическая модель для мониторинга и прогнозирования процесса распространения аэрозольных частиц в атмосфере// Научный журнал Вестник ТУИТ- Ташкент, 2019. - №2(50)*

Равшанов Н., Нарзуллаева Н., Моделирование процесса переноса и диффузии мелкодисперсных частиц в атмосфере с учетом эрозии почвы // Проблемы вычислительной и прикладной математики. – Ташкент, 2017. – № 4 (10). – С.19-35.

OMNITY: TADQIQOTCHILAR UCHUN SAMARALI KO'P TILLI QIDIRUV TIZIMI

Azamov T.N., Sultonov Y.O., Toshkent axborot texnologiyalari universiteti huzuridagi Axborot kommunikatsiya texnologiyalari ilmiy-innovatsion markazi, O'zbekiston

Annotatsiya: *"Omnity" deb nomlangan ko'p tilli hujjatlarni izlash tizimi fanlararo tadqiqotlarni soddalashtirishga xizmat qiladi. Tizim semantik xaritalashga asoslanadi, bu foydalanuvchilarga kalit so'zlardan emas, balki kontekst bo'yicha ma'lumot izlashga imkon beradi. Yangi platforma qidiruv so'rovi sifatida to'liq hujjatlardagi nodir so'zlardan foydalanadi va ularni yanada samarali va ijodiy izlanishni va'da qilgan millionlab boshqa hujjatlar bilan taqqoslaydi.*

Kalit so'zlar: *hujjatlarni qidirish, ko'p tilli tizim, ma'lumot, kontekst, so'rov.*

Abstract: *A multilingual document retrieval system called "Omnity" can simplify interdisciplinary research. The system is based on semantic mapping, which allows users to search for information in context, rather than in keywords. The new platform uses rare words in complete documents as a search query and compares them with millions of other documents, promising a more efficient and creative search.*

Key words: *document search, multilingual system, information, context, request.*

Аннотация: *Многоязычная система поиска документов под названием "Omnity" может упростить междисциплинарные исследования. Система основан на семантическом отображении, что позволяет пользователям искать информацию в контексте, а не в ключевых словах. Новая платформа использует редкие слова в полных документах в качестве поискового запроса и сравнивает их с миллионами других документов, обещая более эффективный и творческий поиск.*

Ключевые слова: *поиска документов, многоязычная система, информация, контекст, запрос.*

Omnity deb nomlangan ko'p tilli hujjatlarni izlash tizimi fanlararo tadqiqotlarni soddalashtirish uchun xizmat qiladi. Ushbu vosita 2016 yil may oyida ishga tushirilgan va foydalanuvchilarga kalit so'zlardan emas, balki kontekst bo'yicha ma'lumot izlashga imkon beradigan semantik xaritalashga asoslangan. Yangi platforma qidiruv so'rovi sifatida to'liq hujjatlardagi nodir so'zlardan foydalanadi va ularni yanada samarali va ijodiy izlanishni va'da qilgan millionlab boshqa hujjatlar bilan taqqoslaydi.

Bilim doimiy ravishda o'sib borishi bilan, endi biron bir sohada bilimda qolish imkonsiz. Mualliflar endi hamma narsani o'qiy olishmaydi va ular o'zlari bilmagan narsalarni iqtibos qila olmaydilar. Natijada, iqtiboslar orqali bog'langan

hujjatlar tarmog'i tobora kam uchraydi (aslida bir-biriga bog'lanishi kerak bo'lgan hujjatlarning 1 foizdan kamrog'i buni amalga oshiradi).

Tabiiy tilga ishlov berish, mashinani o'rganish, grafik matematikani va lingvistik blokchainni birlashtirgan Omnity oddiy nodir so'zlarning o'xshash naqshlarini topib, hujjatlar orasidagi umumiy fikrni topadi - bu havolaga asoslangan hujjatlarni odatiy kashf qilishdan tashqariga chiqadigan matematik yondashuvdir. Omnity 100 dan ortiq tillarda ishlaydi, shu jumladan belgilar asosida.

Omnity olimlar uchun ayniqsa foydalidir. Bu qidiruv tizimi olimlarga, muhandislarga, tibbiyot mutaxassislariga, moliyaviy mutaxassislarga va yuristlarga turli manbalarda, shu jumladan ilmiy jurnallarda, patentlarda, grant buyurtmalarida, klinik sinovlarda, huquqiy yozuvlarda, moliyaviy hisobotlarda va boshqalarda ma'lumotlarni qidirishga imkon beradi.

Ushbu vosita "bilim sohalari va ular o'rtasidagi munosabatlarning qimmat bo'lmagan yashirin modellarini kashf qilish" ni tezlashtirish uchun ishlab chiqilgan. Omnity-ga ko'ra, semantik qidirish tizimi bir soniya ichida 100,000 hujjatlarni taqqoslashi mumkin. Xuddi shu ishni bajarish uchun bitta odam 10000 yilni oladi.

2017 yilning dekabr oyida platforma yangilandi va hozirda yuzdan ortiq tillarni, shu jumladan xitoy, yapon, koreys, fors, rus, nemis, frantsuz, ispan va arab tillarini o'z ichiga oladi. Shuning uchun tadqiqotchilar hujjatni ushbu tillarning istalgan tilida yuklab olishlari mumkin. Hujjatlar bir-biriga bog'lanmasa yoki bir-biriga murojaat qilmasa ham, tegishli natijalar bilan grafik oladi.

Mavjud ma'lumotlarning miqdorini hisobga olib, ma'lum bir sohadagi so'nggi o'zgarishlarni kuzatish tobora qiyinlashmoqda va tadqiqotning boshqa sohalarida nima sodir bo'layotganini kuzatib borish yanada qiyinlashadi. Omnity-ning maqsadi - buni osonlashtirish. Lekin bu vosita qanday ishlaydi degan savol tug’ilishi mumkin.

Foydalanuvchi avval hisobni ro'yxatdan o'tkazishi kerak, so'ngra qidiruv satriga so'rovni kiritishi yoki hujjatni qidiruv maydonchasiga sudrab borishi kerak. Tizim matnni skanerlaydi va eng noodatiy so'zlar va iboralarni qidiradi, so'ngra ushbu imzoni matematik tasvirga tarjima qiladi va so'rov beradi. Keyin u o'z kutubxonalarini - tadqiqotning turli sohalarida skanerlaydi va foydalanuvchiga eng mos keladigan natijalarni taqdim etadi.

Natijalar birlamchi manbalarga (so'rov bilan bevosita bog'liq bo'lganlar) va ikkilamchi manbalarga (birlamchi hujjatlar bilan umumiy tushunchalarga ega bo'lganlarga) bo'linadi. Hujjatlar o'rtasidagi aloqalar juda batafsil jadvallarda, grafikalarda va xaritalarda tasvirlanadi. Foydalanuvchilar ushbu so'zlarni o'z ichiga olgan hujjatlarni ko'rish uchun "bulut" so'zini bosishlari mumkin.

Omnity olimlarga turli tillarda va tadqiqotning turli sohalarida hujjatlarni tezda skanerlash imkonini beradi, bu esa global fanlararo tadqiqotlarni osonlashtiradi. Asosiysi xizmatni bepul ishlatish mumkin, ammo korxona darajasidagi mijozlar uchun pulli versiya (ilg'or xususiyatlarga ega) ham mavjud. Hozirgi vaqtda ma'lumotlar bazasi ingliz tilidagi repazitoriyalarga qaratilgan,

ammo vaqt o'tishi bilan kompaniya ingliz bo'lmagan hujjatlarni qo'shishni rejalashtirmoqda.

Adabiyotlar

1. *Джордж Анадиотис / Языково-независимая обработка документов: нахождение связей с помощью статистики, машинного обучения и графов| 18.04.2017*
2. *https://steemit.com/*

Путри Ханум / Omnity - умная база актуальных знаний о блокчейне! http://putrihanum14.blogspot.com/

BARRIERS FOR PROMOTING THE E-COMMERCE BUSINESS IN PAKISTAN

Muhammad Umair Nawaz, Riphah International University Islamabad (Lahore Campus), Pakistan

Abstract*. Due to the vast development in the internet and mobile phones globally the world is moving towards the technological devices (Smart Phone) uses. The ratio of users is growing day after day in all over the worlds. This ratio is also encouraging the business organizations either a company or individual providers, to move their business in smart environment. The business that performed on the internet electronically buy and sells things is known as E-Commerce business. As the whole world is moving toward the E-commerce business, there the Pakistan is still facing some issues to promote the E-commerce business. In this report we discuss the E-commerce business and also mentioned the issues in Pakistan that causes of limitations for spread it.*

Keyword: e-commerce, methodology

1. INTRODUCTION

E-Commerce business appeared due to the web what's more, has altered the past exchange designs. It plays a significant job in the dissemination of product, capital and data. The organizations that receive e-commerce based business innovation have as up priority on different and lessen their expense of creation. Hence, customary business can likewise expand their benefits and expand their effectiveness. In forthcoming online business won't simply transformation of technique for cooperating anyway it will furthermore affect every piece of human culture similar activities relating to creation, business, preparing and government [1]. In actuality, E-exchange is the paperless exchange of information among clients, associations and countries using an extent of advancements or procedures together with Records Transaction, online cash exchange-mail, and further Internet center daps and it essentially implies the improvement of business on internet based business focus e-business structures oversee buying and selling of things, organizations what's more, and information over the web [2].

Subsequently, we can say that web based business is a more extensive zone and web based shopping is a limited zone in web based business that empowers shoppers and organizations to legitimately purchase items and administrations from a dealer (customer/organization) via the web without the inclusion of a mediator. A customer buys items online from a shopper through the web it’s known as shopper to customer web based shopping. Dell is a case of b2c web based shopping; which sells work area PCs, printers, scanners and so forth to shoppers just for example organizations .When a business buys items and resources by different organizations at that point known as business to business internet shopping. Cisco is an model, it gives organizing arrangements comprising of

equipment and programming to different organizations .Consumer to purchaser web based shopping is that wherein a customer purchases from another buyer over the web through the association of an outsider site that encourages accomplishment of offer and buy exchanges and charges its commission. EBay is a case of such site [3, 4].

The Pakistan is also moving towards the E-commerce business, that we are discussing in this paper with the challenges faced by the users and business providers.

2. RESEARCH OBJECTIVE AND STUDY SIGNIFICANCE

The main aims of the research are to draw a complete picture of e-Commerce business in Pakistan with the following points [14, 15]:

- To highlight and explain the issues for the adoption of e-commerce business in Pakistan. Also explain the issues with the perspective of population in Pakistan.
- Find out the future of E-commerce business in current scenario, additionally elaborate the future work of business in Pakistan.

This study will help in the business provider (Companies, individuals to target the exact users, find the mentality of users in Pakistan in terms of E-commerce, education level of users in Pakistan to use the websites, and how to promote the business. This study will also help the research community to explore this area further.

3. RESEARCH METHODOLOGY

From this line of research issues in the Pakistan related to E-commerce, we collected papers relevant to this topic from the reputed search engines such as IEEE, Springer, Taylor and Francis, Wiley & Sons, ELSEVIER, and MDPI. We searched papers with the following keywords [16, 17]:

- E-commerce Business Introduction.
- Internet use among population in the Pakistan.
- Number of current E-commerce business on Pakistan.
- Data analysis of E-commerce business in Pakistan.
- E-commerce issues in Pakistan.
- E-commerce trends in Pakistan.
- E-commerce limitations in adoption and solutions in Pakistan.
- Barriers for E-commerce business in Pakistan.
- Prevention measures to enhance the E-commerce business in Pakistan.

Our adopted research methodology is presented in the Figure.2

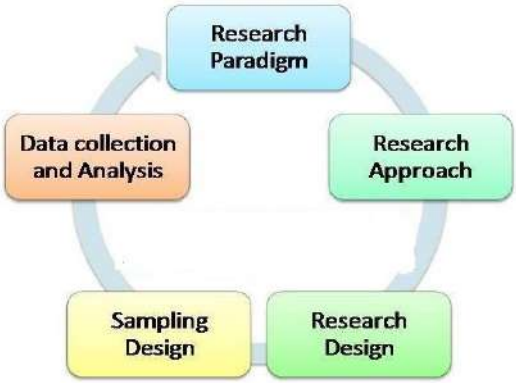


Figure.1 Adopted research methodology.

Research Paradigm:

The research paradigm is the adopted model of research proposed by research community that we followed for our study.

Research Approach:

The research approach defines the parameters to collect data, analysis and explanations. We adopted survey data collection in our study.

Research Design:

It is the arrangement of strategies and methods utilized in gathering and investigating proportions of the factors determined in the difficult research.

Sampling Design:

The samplings include the people in your study as interview or survey results but in our study we don't have any people.

Data Collection and Analysis:

It includes the data collection and analysis using the pre-defined principle methods. We adopted these methods to find the survey papers related to our study.

4. LITERATURE WORK

Pakistan has for quite some time been in a mission to change its economy into a dynamic and mechanically boosted nation having an exchange overflow, abundant remote trade saves, where its residents have more prominent openings for work with better earnings and better expectations of living. In the event that there is one such part that could be created to accomplish the national vision of flourishing, it could be securely said that E-trade is the one, as it will in general accept practically all different areas. Web based business indicates to the buying and selling of product or organizations using the web. By utilizing Web based business, associations can loosen up their compass to national and widespread markets with by and large lower capital endeavor. Pakistan's E-exchange industry has been assessed at PKR 88.2Bn in FY17 (FY17: PKR 50.6Bn) with YoY advancement of 79%. With a masses of ~217Mn and different money related fuse plans in regards to branchless banking, Pakistan offers one of the greatest unfamiliar markets on the planet for E-exchange. Despite the way that, the country has about 69.6Mn broadband supporters what's progressively, 78.5Mn 3G/4G

participations, E-exchange has not reached to a level where electronic money is overwhelmingly used rather than cash for such trades [5, 6, 8].

Pakistan is the 6th most crowded nation on the planet (around 200 million individuals) and the early Ecommerce industry in Pakistan has exponential space for development. There have been a gathering of participants in the business as an ever increasing number of individuals are making vigorous raids in the Web based business industry.

The endless utility of E-trade has at long last gotten up with the Pakistani business people and financial specialists who currently clearly perceive the gigantic potential it holds. The monetary and budgetary markers pre dictanimmense growth in speculation and business potential in the said division in Pakistan. There is a variety of elements that strikingly show how this undiscovered industry is bound to move and further stir support for business people also, little and medium ventures to usher forward in the practical financial market field. There is an immediate connection between the web clients and E-trade industry. As the web supporters increment in Pakistan, the possibility of more E-business action increments. Figure 2 shows the quantity of individuals using internet providers in Pakistan and that lone 14% of the populace is connected to the web [11, 12, 13].

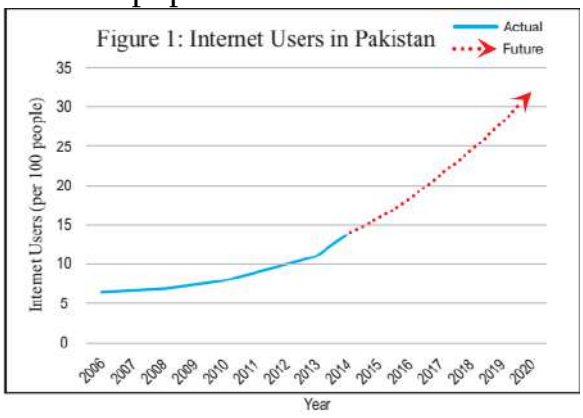


Figure.2Total Population using Internet by Yearly [11].

Approximation 28 million individuals were utilizing the web till 2014; notwithstanding, development pace of web clients has been exponential. The quantity of clients using web administrations expanded by 26% from 2013 to 2014. In view of the normal of the most recent four years as indicated by Figure 1. It can be anticipated that constantly 2020 the web clients will increment to 32% of the entire populace.

5. ISSUES WITH E-COMMERCE BUSINESS IN PAKISTAN

The many studies conducted on the barriers of the adoption of E-Commerce business in the different countries. These studies also highlight some factors such as geographic location, commonality to web, chance discernment, comfort, plan of site, item assortment, item/administration quality, item warranted, item/administration cost, and social standards and so on. Further the factors mentioned by different studies as well are: Gender, financial status, living condition, and information all were evaluated as powerful advertisers of web based shopping [11, 10 ,9].

Pakistan is also promoting the ecommerce business among the people. The Pakistan has a strong infrastructure between the developing countries. In the last years it renovated the high speed communication internet infrastructure. But still Pakistan is unsuccessful to spread the E-commerce business in all over the country. This is due to major barriers in the Pakistan as mentioned follows [4, 8, 10]:

1. Most important factor in the Pakistan is the literacy rate that cause of less adoption of E-commerce business in Pakistan. People have no proper education use these inline sites. The Pakistan needs some education about the E-commerce business both business men and users about how to use the sites and how to promote the business online. As current scenario duetocovid-19 virus whole world is moving towards the online business but still in Pakistan it is not implemented due to less awareness about the E-commerce business.
2. Financial constraints are also major drawback for the spread of E-commerce business in Pakistan. As most of the transactions are done by the PayPal in the whole world but unfortunately still in Pakistan there is no PayPal. People don't have knowledge about PayPal and how to use it. At the government level is not promoted and starts any campaign to aware the people about the PayPal. However, the PayPal takes into consideration minimal cost worldwide exchange of assets for both E-trade deals furthermore, outsourcing continues. Further, nearness of worldwide B2B and B2C entryways like Amazon and eBay is likewise ailing in Pakistan.
3. The main issue with the E-commerce business is in Pakistan is the gender. As the most of the men in the Pakistan are more using the E-commerce business sites as compare to women. The major reason of this is that men in the Pakistan were using the IT below the age of 17 and at the age of 18-30, but the women in the Pakistan started IT after the 18-30.
4. Education level is most important part for the growing of business in not even in Pakistan but all over the world. In the Pakistan statistical data shows that people holding bachelor degree are using the Internet or IT 3 hours on daily basis, whereas people with the higher education are using the IT less or more 5 hours daily. The low education holders in Pakistan are women as compare to men. Therefore the education level in the Pakistan is most important to enhance the electronic business.
5. Other limitation in the Pakistan is the access of internet. Study reveals that the most of the people in the Pakistan have limited internet access. Only 62% male population in 24% have un-limited access remaining percentages in the men and women population have limited access of internet.
6. Most of people in the Pakistan are using unlimited internet access on their laptops and on mobile phone are using the mobile data (limited access) that's the reason of less use of E-commerce sites in Pakistan.
7. 40% female members and 36% male members were found to have general information about the online business. 12% of male residents know nothing about online business.

8. The members of this examination referenced that powerful Digital laws ought to be executed. They were uncertain about the unwavering quality of new innovation. The equivalent number of guys and females were likewise worried about the nature of the items they may buy.

9. Other reason of less adoption of business in Pakistan is the not friendly user sites of business in Pakistan. Mostly the sites have less attractive graphics on the sites to attract the customers of visit and buy things.

10. There exists a solid potential that online business will be progressively mainstream among females, since they were having generally progressively agreeable experience of internet shopping. Moreover, females confided in more that the arranged items were equivalent to requested, while, the vast majority of the male members were some place progressively suspicious to this guarantee.

11. The 18% population in the Pakistan has trust on online sites for buy anything. Trust is major component for internet business that customer needs by the business provider. All people don't use or buy things online as they have doubt on the provider for their money lost.

12. 39% of the participants were sure for more brilliant eventual fate of the web based business. Besides, Gender insightful 79% men and 29% of ladies were hopeful for more promising time to come of web based business in the Pakistan.

Therefore these are some major barriers which don't promote the E-commerce business in Pakistan. All factors are important to target and solve but education, unlimited internet access, trust of customers on the providers, and educating about the E-commerce business both clients and business providers should be targeted and solved.

6. CONCLUSION

Existing study examine researched and analyzed that conceivable IT-and human associated boundaries are constraining the utilization of internet business among Pakistani residents. Nevertheless, the present examination is giving a more profound and more extensive correlation of the gender orientation based components that are upsetting men and women from the utilization of web based business in the nation. This study also highlight some factors that needs to be address existing work found that notwithstanding high instructive degrees of most of the clients, they are not rehearsing online business as it ought to be. As it isn't obvious from this investigation assuming less taught residents don't have the financial assets in request to utilize the innovation or they need something more information about the most recent patterns. This will assist with showing signs of improvement review of how internet business is seen among less taught or uneducated residents and what are the potential components which may restrain their web based business appropriation and how these impediments can be expelled.

REFERENCES

[1] Tarek, T. A. (2011). An empirical-based model for examing e-purchasing intention in electronic commerce at developing counties. ***International Journal of Academic Research, Vol. 3, No. 3, pp. 683-688 (2011)***

[2] Bukhari, F. Y., & Ramzan, M. (2013). Gender discrimination: A myth or truth women status in Pakistan. Journal of Business and Management, 8(2), 88-97.

[3] Chen, L., & Pu, P. (2014). Experiments on user experiences with recommender interfaces. Behavior & Information Technology, 33(4), 372-394.

[4] Agren, E. S., & Barbutiu, S. M. Barriers In The Adoption Of E-Commerce In Pakistan With The Focus On Gender.

[5] Imtiaz, S., Ali, S. H., & Kim, D. J. (2020). E-Commerce Growth in Pakistan: Privacy, Security, and Trust as Potential Issues. Culinary Science & Hospitality Research, 26(2), 10-18.

[6] Saeed, S. (2016). Investigating the factors affecting B2B E-Commerce Adaptation in Pakistan. International Journal of Technology and Research, 4(3), 65-68.

[7] Ghouri, M. W. A., Hussain, M. A., Kanwal, T., Afzal, B., & Rasheed, M. (2018). Influence of E-Commerce on Business Growth in Pakistan: An Approach to Analyze Privacy, Threats and Consumer Trust in Business. International Journal of Computer Science Issues (IJCSI), 15(2), 27-35.

[8] Dar, M. S., Ahmed, S., & Raziq, A. (2017). Small and medium-size enterprises in Pakistan: Definition and critical issues. Pakistan Business Review, 19(1), 46-70.

[9] Zafar, A., & Mustafa, S. (2017). SMEs and its role in economic and socio-economic development of Pakistan. International Journal of Academic Research in Accounting, Finance and Management Sciences, 6(4).

[10] Ahmad, I., & Ahmad, S. (2018). Multiple Skills and Medium Enterprises' Performance in Punjab Pakistan: A Pilot Study. The Journal of Social Sciences Research, Special, (4), 44-49.

[11] Mehmood, N., Shah, F., Azhar, M., & Rasheed, A. (2014). The factors effecting e-banking usage in Pakistan. Journal of Management Information System and E-commerce, 1(1), 57-94.

[12] Farhan Zafar, E-Commerce Industry in Pakistan, SMEDA.

[13] Yates, S. R. (2020). Factors Associated With Electronic Banking Adoption. Journal of Financial Counseling and Planning.

[14] Quinlan, C., Babin, B., Carr, J., & Griffin, M. (2019). Business research methods. South Western Cengage.

[15] Quinlan, C., Babin, B., Carr, J., & Griffin, M. (2019). Business research methods. South Western Cengage.

[16] Booth, A. (2016). Searching for qualitative research for inclusion in systematic reviews: a structured methodological review. Systematic reviews, 5(1), 74.

[17] Cooper, C., Booth, A., Varley-Campbell, J., Britten, N., & Garside, R. (2018). Defining the process to literature searching in systematic reviews: a

literature review of guidance and supporting studies. BMC medical research methodology, 18(1).

PREDICTION OF HEART DISEASE VIA MACHINE LEARNING

Bilawal Ali Jaan, Riphah international University Lahore, Pakistan

Abstract. *Depending on their properties Data can be categorized .We are categorized into existing or past records (data). One of the main purposes of classification is that we can obtain good results and increase the reliability of the data obtained. In the health care industry, clinical diagnosis is usually based on physician information. And computer-generated decision-making plays a major role in the field of medicine. The growing research in the prediction of heart diseases is important to separate the research findings and provide information to the student about the possible strategies for predicting heart disease at each stage. Medical data prediction, clustering, and classification are the best algorithms that can be used. From the research, it was observed that hybrid algorithms improve the accuracy of prediction algorithms for heart diseases. In this paper, we summarize the most common techniques used to predict heart disease and its severity.*

Keyword: *Naive Bayes Classifier,* ***Decision Trees, Random Forest, weka,*** *heart disease, prediction, machine learning, classification.*

1. INTRODUCTION

The disease of heart is the big cause of death worldwide. Cardiac dysfunction usually increases from ischemia due to insufficiency blood supply to the heart muscle. Long-term shortage of oxygen and nutrients in the uterus leads damage to the heart muscle that eventually reaches the heart attack and death. Early diagnosis of cardiac ischemia is important to help reduce the number of deaths. Heart Diseases are the leading cause of death for the previous some years. There are many computer software algorithms and techniques for assets the doctors to predict heart diseases in the early stages. The cardiovascular prediction system can help medical professionals predict heart disease based on previous clinical data of patients.

The medical industry collects huge amounts of medical data and we use this data and apply different algorithms for prediction. According to common properties, data can be classified in different sets. The Method of classification is a sub-Type of cluster analysis. The method of Classification used by developing a classify the existing model that can previous records using sample data. The researchers introduce many types of classification algorithms in previous past years. But these are the most normally used classification algorithms, AC2, New ID CAL5, C4.5, CN2, Bayes theorem, clustering recursion algorithm IT rule, k neighbor nearest, radial basis function, Naive Bayes, Kohonen Poly trees, net, LVQ self-organizing, Hard K means and C-Means but we apply some of these for finding the classification and predict the heart disease. The reason is to use this standard algorithm to present similar

structures and learn from previous data and then use experimental data and train the data and predict the results. Clustering and classification combine together to form a prediction analysis in which clustering categories the instances of the behalf of similarities and it's also unsupervised learning and the classification is a process of the label to the data which are unlabeled it is also supervised learning.

2. LITERATURE REVIEW

Classification is one of the most important tools for decision making in the field of medical science. When referring to related publications it can be seen that more research is being done on heart diseases most comprehensive among them is the (Kahramanli, 2008) many algorithms ware used to classification by their study the accuracy rate was too good more than 89% accuracy rats. The training data set and testing data set size were 313 are respectively, in this study, the 2nd best article I read (Seenivasagam2, JULY 2013) this is one of the interesting research I ever read, better accuracy and prediction and the Third article I read for my research paper is (Tanawut Tantimongcolwata, 24 April 2008) this paper is little bit complex and technical but I learn many thing technically from this paper thy use more than 13 algorithms define all algorithms specifically and apply all and as well as show all results so these are so articles I review before this research these paper are to useful and help me for my research paper.

3. METHODOLOGY

3.1 Quantitative Research

Quantitative research is characterized as an orderly examination of marvels by social affair quantifiable information and performing measurable, scientific, or computational procedures. Quantitative research gathers data from existing and potential clients utilizing testing strategies and conveying on the web overviews, online surveys, polls, and so forth. The consequences of which can be portrayed as numerical. After cautious comprehension of these numbers to anticipate the eventual fate of an item or administration and make changes in like manner.

3.2 Tools and techniques

3.2.1 Naive Bayes Classifier

It is a classification method based on Bayes' Hypothesis with the assumption of autonomy between researchers. In other words, the organizers of the Naive Bayes group think that the occurrence of a specific feature in the classroom has nothing to do with the fact that something else or that all of these components have an independent role is possible. As well as being fair, Naive Bayes is known for the highest quality of classification

3.2.2 Decision Trees

The decision tree establishes a model of classification or classification as a tree structure. Breaks data set into smaller and smaller ones, while

simultaneously improving the decision tree. The result is a tree with Decision Nodes and Leaf Nodes. A decision node has two or more branches and a leaf node represents a category or decision.

3.2.3 Random Forest

Random forests are the study of Classification, register, and other functions that work to create a large number of decision trees during training and the class extraction which is a class method (classification) or prediction that means predicting (rearranging) specific trees.

3.2.4 WEKA

Weka is attempted and tried open-source AI programming that can be gotten to through a graphical UI, standard terminal applications, or a Java API. It is generally utilized for educating, explore, and mechanical applications, contains a plenty of implicit apparatuses for standard AI undertakings, and also gives straightforward access to notable tool compartments, for example, scikit-learn, R, and Deeplearning4j.

4. STUDY ENVIRONMENT AND BACKGROUND THEORIES

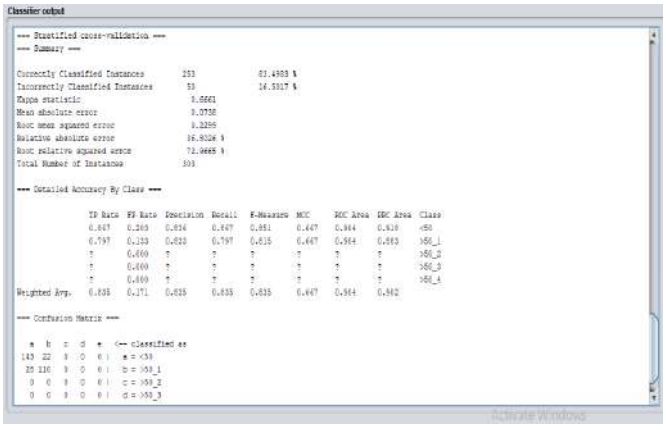
Clustering and classification were created in this investigation. The techniques that were utilized will be clarified quickly in this segment. Now we start apply different algorithms on our data set of heart diseases first of all we start with classification.

5. CLASSIFICATION

In classification we need our data set with different data we use this data for classification we put our dataset into our tool WEKA and apply different classification algorithms class In AI and measurements, Classification is a learning process managed when the PC system receives information and then uses this to find a way to integrate new ideas.

5.1 Naive Bayes Classifier

We apply naive Bayes classifier on our data we use 66% training data from data set and 44% testing data it is a technique of classification dependent on Bayes' classifier with the presumption of self-governance between the indicators.



When we apply nave Bayes theorem we see the results

Summary	Weka Result
Correctly Classified Instances	253
Incorrectly Classified Instances	50
Kappa statistic	0.6661
Mean absolute error	0.0738
Root mean squared error	0.0738
Relative absolute error	0.2299
Root relative squared error	36.8026 %
Total Number of Instances	72.9665 %

Table: 4

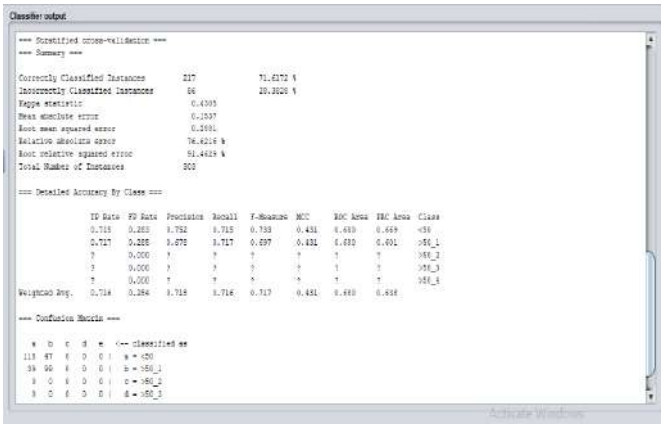
=== Confusion Matrix ===

a	b	c	d	e	classified as
143	22	0	0	0	a = <50
28	110	0	0	0	b = >50_1
0	0	0	0	0	c = >50_2
0	0	0	0	0	d = >50_3
0	0	0	0	0	e = >50_4

These are the results of algorithm which we apply when we use 66% training data and 44% testing data the results are fantastic it's near about 85% accurate we use 303 instances in our data set so result is good

5.2 Decision Trees

We incorporate the classification decisions into our data using 66% training data from the data set and 44% test data the decision tree performs a classification or regression model as a tree structure.



When we apply decision Tree classifier we see the outcomes

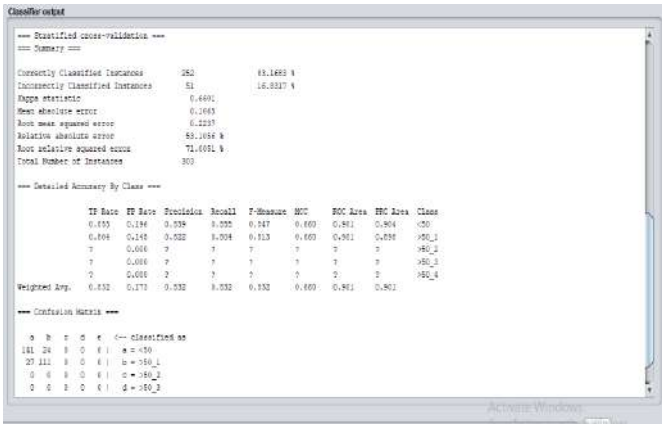
Summary	Weka Result
Correctly Classified Instances	217
Incorrectly Classified Instances	86
Kappa statistic	0.4305
Mean absolute error	0.1537
Root mean squared error	0.2881
Relative absolute error	76.62%
Root relative squared error	91.46%
Total Number of Instances	303

Table: 2

These are the results of algorithm which we apply when we use 66% training data and 44% testing data the results are fantastic it's near about 72% accurate we use 303 instances in our data set so result is good but not better then naive Bayes classifier

5.3 Random Forest

We incorporate the Random Forest classifier into our data using 66% of training data from the data set and 44% of the Random Forest test data for unique classification.



When we apply Random Forest classifier we see the results

Summary	Weka Result
Correctly Classified Instances	252
Incorrectly Classified Instances	51
Kappa statistic	0.6601
Mean absolute error	0.1065
Root mean squared error	0.2237
Relative absolute error	53.10%
Root relative squared error	71.00%
Total Number of Instances	303

Table: 3

6. DISCUSSION

any techniques of the heart rate prediction system are presented in this paper. We use three classification methods or algorithms in this paper first the Bayes classifier (classification procedure) is analyzed. The second type of decision trees used to predict heart disease is analyzed. Final Type Random Forest classifier

In these methods the data of heart diseases hospitals contains the medical data of heart patients used for the detection of cardiovascular diseases. The classification method is used for all proposed work.

7. CONCLUSION

Heart disease is one of the leading causes of death worldwide and early diagnosis of cardiovascular disease is important. The computer-aided heart rate system assists the doctor in the form of a diagnostic tool that targets heart disease. Another technique for the classification of cardiovascular diseases is now being studied in this research paper. After examination, it is established that data science plays a major role in the collection of cardiovascular diseases. The Naive Bayes class for offline training is suitable for disease prediction at the beginning of the phase because 84% of Bayesian students' accuracy and program performance can be determined by exposed and standard data. The clarification of the sections will be improved by dropping the markers.

REFERANCES

[1] Kahramanli, H. (2008). Design of a hybrid system for the diabetes and heart diseases. Science Direct, 35, 82-89.

[2] Seenivasagam2, R. C. (JULY 2013). REVIEW OF HEART DISEASE PREDICTION SYSTEM USING DATA MINING AND HYBRID INTELLIGENT TECHNIQUES. ICTACT JOURNAL ON SOFT COMPUTING,, 3, 605-609.

[3] Tanawut Tantimongcolwata, T. N.-N.-A. (24 April 2008). Identification of ischemic heart disease via machine learning. Computers in Biology and Medicine, 817 – 825.

СИГНАЛЛАРГА РАҚАМЛИ ИШЛОВ БЕРИШДА ЭНГ КИЧИК КВАДРАТЛАР УСУЛИ

Азимов Б.Р
Андижон давлат университети, Ўзбекистан.

Аннотация. *Ушбу ишда сигналларга рақамли ишлов беришда муҳим ҳисобланган энг кичик квадратлар усули ёритилган. Шу билан бирга функцияни интерполяциялаш йўли билан кўпҳадлар синф орқали тақрибий алмаштириш масаласининг камчиликлари келтирилган.*

Таянч иборалар: *интерполяциялаш, ўрта квадратик оғиш, энг кичик квадратлар, ўрта квадратик хато.*

Аннотация. *Эта статья кратко описан метод наименьших квадратов, который считается важным при цифровой обработке сигналов. Однако представлены недостатки проблемы приближенной замены полиномиальным классом путем интерполяции функции.*

Ключевые слова: *интерполяция, среднеквадратичное отклонение, наименьшие квадраты, среднеквадратичная ошибка.*

Abstract. *This article briefly describes the least squares method, which is considered important in digital signal processing. However, the disadvantages of the problem of approximate replacement by a polynomial class by interpolating a function are presented.*

Key words: *interpolation, standard deviation, least squares, standard error.*

Биламизки сигналларга рақамли ишлов беришда асосий масала сифатида сигналларни тиклаш ва қайта ишлаш масаласи ҳисобланади. Ушбу ишда ҳам сигналларга рақамли ишлов беришда қўлланиладиган кичик квадратлар усулини қурилиш кўриб чиқилган.

Фараз қилайлик, $\varphi_0(x),\varphi_1(x),\ldots,\varphi_n(x)$ етарлича силлиқ ва ҳисоблаш учун қулай бўлган чизиқли эркли функциялар системаси бўлсин. Бу функциялардан тузилган

$$P_n(x)=c_0\varphi_0(x)+c_1\varphi_1(x)+\Lambda+c_n\varphi_n(x) \qquad (1)$$

Чизиқли комбинация (c_0,c_1,Λ,c_n - доимий сонлар) умумлашган кўпҳад дейилади. Берилган $f(x)$ функцияни интерполяциялаш йўли билан $P_m(x)$ орқали тақрибий равишда алмаштириш масаласи билан танишмиз. Аммо шуни ҳам таъкидлаб ўтиш лозимки, қатор масалаларда функциянинг бундай тақрибий тасвирланиши мақсадга мувофиқ бўлавермайди. Биринчидан, тугунлар сони кўп бўлса, у ҳолда интерполяцион кўпҳадларининг ҳам даражаси ортиб боради, лекин бу яқинлашишнинг сифати ҳар доим ҳам яхши бўлмаслиги мумкин. Иккинчидан, $f(x)$ функциянинг тугун нуқталардаги қиймати бирор тажрибадан аниқланган бўлиши ҳам мумкин, у ҳолда табиий равишда бу қийматлар тажриба хатосига эга

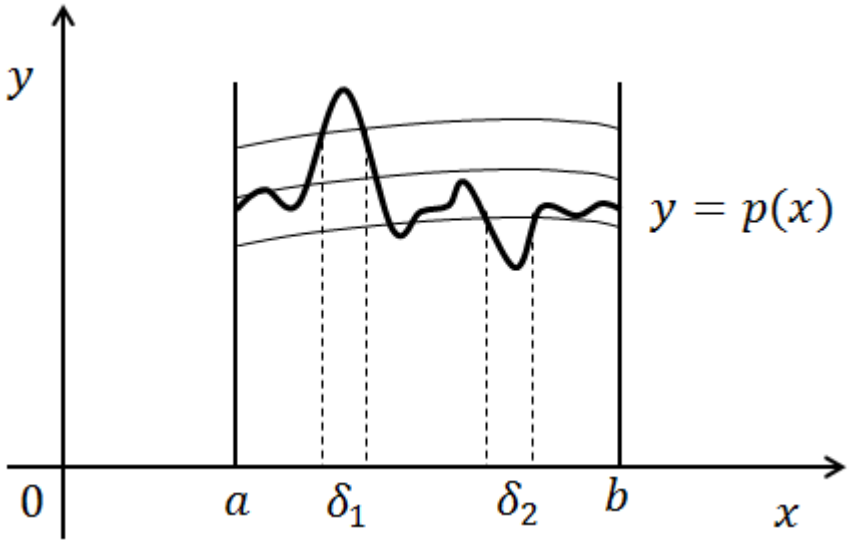
бўлиб, у интерполяцион кўпҳадга ҳам таъсир қилади ва шу билан функциянинг ҳақиқий ҳолатини ҳам бузиб кўрсатади.

Қандайдир маънода бу камчиликдан холи бўлган ўрта квадратик яқинлашувчи кўпҳадларни тузиш билан шуғулланиш мақсадга мувофиқдир. Шундай қилиб, биз функциялар учун ўрта квадратик маънода яқинлашиш масаласи қўйилишининг мақсадга мувофиқ эканлигига ишонч ҳосил қилдик. Бу масала қўйдагидан иборатдир: $[a,b]$ оралиқда аниқланган $f(x)$ функция учун (1) кўринишдаги яқинлашувчи шундай кўпҳад топилсинки,

$$\int_a^b [f(x)-P_m(x)]^2 dx \qquad (2)$$

ифода мумкин қадар энг кичик қийматни қабул қилсин.

Агар (2) интеграл кичик қиймат қабул қилса, бу шуни билдирадики, $[a,b]$ оралиқнинг кўп қисмида $f(x)$ ва $P_m(x)$ бир-бирига яқин. Шунга қарамасдан айрим нуқталар атрофида ёки бу оралиқнинг баъзи кичик қисмларида $f(x)-P_m(x)$ айирма нисбатан етарлича катта бўлиши ҳам мумкин (1-расм).



1-расм. Ўрта квадратик оғиш кўриниши.

Қуйидаги (3)

$$\delta = \sqrt{\frac{1}{b-a}\int_a^b [f(x)-P_m(x)]^2 dx} \qquad (3)$$

миқдор $P_m(x)$ нинг $f(x)$ дан ўрта квадратик оғиши дейилади ва $f(x)$ ни $P_m(x)$ билан яқинлашишда ўрта квадратик маънодаги хатони билдиради.

Агар $f(x)$ ни ўрта квадратик маънода $P_m(x)$ билан яқинлаштиришда қандайдир сабабга кўра қаралаётган оралиқнинг бирор қисмида унинг бошқа қисмига нисбатан аниқроқ яқинлаштириш керак бўлса, у ҳолда

кўпинча қўйидагича иш тутилади: вазн деб аталувчи махсус равишда танлаб олинган манфий бўлмаган $\rho(x)$ функция олиниб, (2) ўрнига ушбу

$$\int_a^b \rho(x)[f(x)-P_m(x)]^2 dx$$

интегралнинг энг кичик қийматни қабул қилиши талаб қилинади. Бу ерда $\rho(x)$ шундай танланган бўлиши керакки, агар оралиқнинг бирор нуқтаси атрофига яқинлашиш аниқлиги бошқа нуқталарга нисбатан яхшироқ бўлиши талаб қилинса $\rho(x)$ шу нуқта атрофида каттароқ қийматга эга бўлиши керак. Масалан, $[-1,1]$ - оралиқда $f(x)$ функцияни $P_m(x)$ функция билан яқинлаштиришда яқинлаштириш аниқлигини оралиқнинг четки нуқталар $x=\pm 1$ атрофида юқори бўлишини истасак, $\rho(x)=\frac{1}{\sqrt{1-x^2}}$ деб олиш мумкин.

Агар $f(x)$ функциянинг аналитик кўриниши ўрнига, унинг фақат $(n+1)$ та x_0, x_1, Λ, x_n нуқталардаги қийматларигина маълум бўлса, у ҳолда (2) интеграл ўрнига ушбу

$$\sum_{i=0}^{n} [f(x_i)-P_m(x)]^2 \qquad (4)$$

(4) - йиғиндининг мумкин қадар кичик қиймат қабул қилишлиги талаб қилинади. Бу ҳолда

$$\delta_n = \sqrt{\frac{1}{(n+1)} \sum_{i=0}^{n} [f(x_i)-P_m(x_i)]^2}$$

миқдор *ўрта квадратик оғиш* дейилади. Ўрта квадратик яқинлаштириш усули энг кичик квадратлар усули ҳам дейилади.

Адабиётлар

1. Баркалов К.А. Образовательный комплекс «Параллельные числен- ные методы». Лекционные материалы. Нижегородский государственный университет им. Н.И. Лобачевского Факультет вычислительной математики и кибернетики. Нижний Новгород 2011 г.
2.Завьялов Ю.С., Квасов Б.И., Мирошниченко В.Л., Методы сплайн-функций. Москва.: Наука, 1980. 352 с.
3. Зайнидинов Х.Н., Азимов Б.Р. Построение кубического сплайна для сигналов измеренных в неравных интервалах. Автоматика и программная инженерия. 2020, №1(31), 70-76 б.
4. Зайнидинов Х.Н., Бахромов С.А., Азимов Б.Р. Биомедисина сигналларни интерполяцион кубик сплайн моделини қуриш. Мухаммад ал-Хоразмий авлодлари, № 4 (10), декабрь 2019, 14-17 б.

5. Самарский А. А., Гулин А. В. Численные методы, Москва «Наука», главная редакция Физико-математической литературы, 1989 г.

РЯБЕНЬКО КУБИК СПЛАЙН МОДЕЛИНИ ҚУРИШ

Азимов Р.К., Абдулвохидов А.Л., Андижон давлат университети, Ўзбекистан.

Аннотация. *Ушбу ишда функцияларни сплайн функциялар билан яқинлаштириш масаласи қаралган. Шу билан бирга учинчи даражали сплайн функция моделини қуриш ёритилган.*

Таянч иборалар: *интерполяция, кубик сплайн, функция дефекти, даврий сплайн.*

Аннотация. *Эта статья рассматривается проблема аппроксимации функций сплайн-функциями. Он также охватывает построение модели функции третичного сплайна.*

Ключевые слова: *интерполяция, кубический сплайн, дефект функции, периодический сплайн.*

Abstract. *This article addresses the problem of approximating functions by spline functions. It also covers the construction of a tertiary spline function model..*

Keywords: *interpolation, cubic spline, function defect, periodic spline.*

Функцияларни сплайн функциялар билан яқинлаштириш - Силлиқлиги юқори бўлмаган функциялар учун кўпҳадлар яқинлашиш аппарати сифатида қатор ноқўлайликларга эга. Булардан энг асосийси шундан иборатки, бундай функцияларнинг бирор нуқта атрофидаги ҳолати, уларнинг тўла ҳолати билан узвий боғликдир. Охирги вақтларда шу нуқсондан ҳоли бўлган бошқа яқинлашиш аппаратларни ишлаб чиқилмоқда. Назарий тадқиқот ва татбиқларда яхши натижа берадиган аппарат – сплайн функциялар аппаратидир. Сплайннинг таърифи билан танишайлик. Ҳақиқий ўқдаги $[a,b]$ ораликда ушбу:

$$\Delta_n : a = x_0 < x_1 < \ldots < x_n = b$$

тўр берилган бўлсин. Фараз қилайлик, $H_m(P)$ даражаси m дан ортмайдиган кўпҳадлар тўплами, $C^k = C^k[a,b]$ ўзи ва k - тартибгача ҳосилалари $[a,b]$ ораликда узлуксиз бўлган функциялар тўплами бўлсин.

ТАЪРИФ: Қуйидаги иккита шартни қаноатлантирувчи ушбу

$$S_m(x) = S_m(x, \Delta_n)$$

функция дефекти 1 га тенг m- даражали полиноминал сплайн дейилади:

1. Ҳар бир $[x_i, x_{i+1}]\left(i = \overline{0,n}\right)$ ораликда $S_m(x) \in H_m(P)$
2. $S_m(x) \in C^{m-1}[a,b]$.

Бу ердаги $\{x_i\}$ нуқталар сплайн тугунлари дейилади. $S_m(x)$ сплайнинг м-ҳосиласи $[a,b]$ ораликда узилишга эга бўлиши ҳам мумкин.

Агар $k = 0,1,\ldots,m-1$ лар учун

$$S_m^{(k)}(a+0)=S_m^{(k)}(b-0)$$

тенгликлар бажарилса, $S_m(x)$ сплайн b-а даврли даврий сплайн дейилади.

$f(x)$ функциянинг Δ тўрнинг x_i тугунларидаги $f_i=f(x_i)$ қийматлари маълум бўлсин. $S_m(x)$ сплайн интерполяцион деб аталади, агарда қуйидаги шарт бажарилган бўлса

в) $S_m(x_i)=f_i, i=0,1,\ldots,N.$

Таърифни қаноатлантирувчи сплайнлар билан бир қаторда шундай сплайнлар ҳам қараладики, уларнинг силлиқлиги Δ_n тўрнинг турли қисмларида турличадир. Бундай сплайнлар $[a,b]$ оралиқнинг турли қисмларида турли силлиқликка эга бўлган функцияларни яқинлаштиришда фойдаланилади.

Сплайн ягона равишда аниқланиши учун $[a,b]$ оралиқнинг четки а ва b нуқталарида чегаравий шартлар деб аталувчи қўшимча шартлар қўйилади. Амалда учинчи даражали, яъни кубик сплайнлар кенг қўлланилади.

Сплайнларнинг хисоблаш математикасида кенг қўлланилаётганлиги сабабларидан яъна бири уларнинг қийматларини ЭҲМ ларда ҳисоблашнинг қулайлиги ва улар ёрдамида интерполяциялаш каби жараёнларнинг кенг синфдаги тўрлар учун яхши яқинлашишлигидадир.

РЯБЕНЬКО КУБИК СПЛАЙНИНИ ҚУРИШ. Қаралаётган функция етарлича силлиқ бўлса, у ҳолда бу функцияни сплайн функциялар билан яқинлаштириш мақсадга мувофиқ. Яқинлаштириш аниқлиги юқори бўлган хамда сплайн функцияни экстремал хоссасини инобатга олган холда функцияга яқинлашишда учунчи даражали Рябенько сплайнидан фойдаланиш яхши самара беради.

$f(x)$ функция $f(x)\in C^q(-\infty,\infty)$, q тартибли ҳосиласигача узлуксиз функциялар синфига тегишли бўлиб $q\le p$, бу ерда $p>0$ фиксирланган сон. $f(x)$ функцияни бир ўлчовли тенг оралиқли тугун нуқталарда қийматлари берилган бўлсин

$$f(x)=f(ih),\ (i=0,\pm1,\pm2,\ldots)$$

Рябеньконинг мақоласида $f(x)$ функциянинг $f(ih)$ қийматлари асосида берилган ораликда дефекти $p+1$ бўлган $2p+1$ даражали $S_{ph}(x)$ сплайн функция қурилган.

Бир ўлчовли ҳолатда $\left|f(x)-S_{ph}(x)\right|$ хатолик учун ($r<q$ тартибли ҳосиласигача) қуйдаги баҳо ўринли

$$\left|f^{(r)}(x)-S_{ph}^{r}(x)\right|\le h^{q-r}\ k(p)\sup_x\left|f^{(q)}(x)\right| \qquad (1)$$

бу ерда k(p) фақат p га боғлиқ бўлган ўзгармасдир.

Ихтиёрий p учун $S_{ph}(x)$ сплайнни $f(x)$ функцияни $f(ih)$ қийматлари орқали аниқ кўринишини ёзиш мумкин. Шунинг учун ҳар хил соҳаларда қўлланилаиши жуда қулай. Бу қаралган сплайн функцияни Эрмит сплайнидан афзаллиги C[a,b] да ҳам яқинлаштириш мумкин.

С.Л.Собилов ишларида $[x_i, x_{i+1}]$ оралиқда $S_{ph}(x)$ функция учун қуйидаги ифода берилган

$$S_{ph}(x) = \sum_{j=0}^{p} \sum_{k=0}^{p-j} \frac{(p+k)!}{k!\,j!\,p!} [a_{pj}(x) t^{j+k} (1-t)^{p+1} + (-1)^j a_{pj}(x_{i+1}) t^{p+1} (1-t)^{j+k}] \quad (2)$$

бу ерда $t = \frac{x - x_i}{h}$, $a_{p0}(x_i) = f(x_i)$.

(2) да берилган $S_{ph}(x)$ сплайн функция дефекти $p+1$ бўлган $2p+1$ даражали сплайн бўлиб, бу умумий ҳолда берилган p=1 бўлган ҳусусий ҳоли Рябеньконинг локал учинчи даражали интерполяцион сплайнини беради.

p=1 бўлганда локал кубик сплайн қуйдагича ҳосил қилинади

$$S_{1h}(x) = a_{10}(x_i)(1-t)^2(1+2t) + a_{10}(x_{i+1})t^2(3-2t) + a_{11}(x_i)t(1-t)^2 - a_{11}(x_{i+1})t^2(1-t)$$

$$a_{10}(x_i) = f(x_i),\ \Delta f(x_i) = f(x_{i+1}) - f(x_i)$$

$$a_{11}(x_i) = 1! \sum_{k=1}^{1} \frac{\Delta f(x_i)}{1!} S(1,1) = f(x_{i+1}) - f(x_i),\ S(1,1) = 1$$

$$a_{11}(x_{i+1}) = f(x_{i+2}) - f(x_{i+1})$$

$$S_{1h}(x) - S_3(x) = f(x_i)(1-t)^2(1+2t) + f(x_{i+1})t^2(3-2t) + (f(x_{i+1}) - f(x_i))t(1-t)^2 -$$

$$-(f(x_{i+2}) - f(x_{i+1}))t^2(1-t) = ((1-t)^2(1+2t) - t(1-t)^2) f(x_i) +$$

$$+(t^2(3-2t) + t(1-t)^2 + t^2(1-t)) f(x_{i+1}) - t^2(1-t) f(x_{i+2}) =$$

$$= \varphi_1(t) f(x_i) + \varphi_2(t) f(x_{i+1}) + \varphi_3(t) f(x_{i+2})$$

$$S_3(x) = \varphi_1(t) f(x_i) + \varphi_2(t) f(x_{i+1}) + \varphi_3(t) f(x_{i+2}). \quad (3)$$

бу ерда

$$\varphi_1(t) = (1-t)^2(1+t), \quad \varphi_2(t) = t(1+2t-2t^2), \quad \varphi_3(t) = -t^2(1-t)$$

Ушбу $[x_i, x_{i+1}]$ оралиқдаги (3) функцияни Рябенько сплайн функцияси деб аташ мумкин.

Адабиётлар

1. Завьялов Ю.С., Квасов Б.И., Мирошниченко В.Л., Методы сплайн-функций. Москва.: Наука, 1980. 352 с.
2. Самарский А. А., Гулин А. В. Численные методы, Москва «Наука», главная редакция Физико-математической литературы, 1989 г.

MODELS FOR TWO-DIMENSIONAL SIGNALS BASED ON HAAR WAVELET TRANSFORMATIONS

Ibrohimbek Yusupov
Assistant at the Department of Information Technology, Tashkent University of Information Technologies named after Muhammad al - Khwarizmi, Uzbekistan.

Annotation: *In this study, algorithm and program of digital processing processes on the basis of Haar transformations for two - dimensional signals are given. For construction models of the signals received from real objects, traditional harmonious functions are widely applied.*

Key words: *two-dimensional, Haar, Fure, Walsh fast transform, discrete transformations.*

INTRODUCTION

This thesis tells that many signals received from real objects can be easily presented by set sine and cosine waves fluctuations, which is using for the device of Fourier analysis. Result of it is transition from time to frequency functions. However, representation of time function sine and cosine wave functions is only one from many representations. Any full system of orthogonal functions can be applied for expansion in a series, which correspond with Fourier series.

1. TWO-DIMENSIONAL PIECEWISE-POLYNOMIAL HAAR'S BASIS

The technique of construction of two-dimensional integrated bilinear basic functions of Haar can be based on idea of integration of piecewise-plane orthogonal basic functions. For example, two-dimensional functions of Shauder $Shd_{ij}(x,y) = Shd_i(x) * Shd_j(y)$ as a result of operation of double integration can be constructed:

$$Shd_{ij}(x,y) = \int_0^x \int_0^y har_i(\tau) har_j(\tau') d\tau d\tau' \quad (1)$$

Factors of discrete spectral transformations in bilinear bases are calculated through so-called "diagonal" two-dimensional final differences:

$$\Delta f_{ij} = f(x_{i+1}, y_{j+1}) - f(x_i, y_j) \quad (2)$$

These differences are hypotenuses of the vertical triangles, as one of which legs of a triangle heights of functions, and other leg of a triangle – a diagonal of an

elementary platform in the size $h \times h$ on a plane (x, y). Its length is designated as Δ_{ij}.

For two-dimensional bilinear bases factors of direct discrete transformation are defined under formulas:

$$C_{kl} = \sum_i \sum_j \Delta f_{ij} har_k(x_i) har_l(x_j) \quad (3)$$

$$a_k(y) = \sum_j \Delta f_{ij} har_l(y) \quad (4)$$

Then we can write down:

$$C_{kl} = \sum_i a_k(y) har_k(x) \quad (5)$$

Return two-dimensional discrete transformation owing to material character of basic functions is carried out similarly:

$$\Delta f_{ij} = 4^{-p} \sum_k \sum_l C_{kl} har_k(x) har_l(y) \quad (6)$$

Two-dimensional discrete transformations to bases of integrated functions of Harmuta turn out similarly. Designations change only.

Function restoration in any point (x, y) can be made under the formula

$$f(x, y) = f(x_i, y_j) + h_x^{-1} h_y^{-1} \delta x \delta y \Delta f_{ij} \quad (7)$$

Where $\delta x, \delta y$ – the increments of arguments which are not surpassing accordingly h_x and h_y.

The system piecewise - plane orthogonal “Haar - like” functions can be constructed on the basis of the theory of self-similar trees in dynamic discrete space. The process beginning is crushing of an individual square on the binary-rational areas which too are squares and on these squares groups of the basic functions accepting values +1,-1 or 0 are under construction.

Any (x, y) point of area Ω $(0 \le x, y < 1)$ belongs to a binary square Q_{psr} if co-ordinates of this point belong to corresponding binary pieces $x \in h_{ps}$, $y \in h_{pr}$.

In each square Q_{psr} four equal parts which in turn are binary squares are allocated. The point (x, y) belongs to dynamically decreasing square under conditions:

$$\begin{cases} (x,y) \in Q1_{psr}, & if \quad x \in h_{ps}^{+} \ and \ \ y \in h_{pк}^{+} \\ (x,y) \in Q2_{psr}, & if \quad x \in h_{ps}^{+} \ and \ \ y \in h_{pк}^{-} \\ (x,y) \in Q3_{psr}, & if \quad x \in h_{ps}^{-} \ and \ \ y \in h_{pк}^{+} \\ (x,y) \in Q4_{psr}, & if \quad x \in h_{ps}^{-} \ and \ \ y \in h_{pк}^{-} \end{cases}$$

Thus, recursive streamlining with corresponding hierarchical numbering is made. On a square Q_{psr} three orthogonal functions hd_{psrl} of values +1 or-1 with an index l = 1, 2 and 3 are formed:

$$hq_{psr1}(x,y) = hq_{ps}(x) = \begin{cases} +1\ , x \in h_{ps}^{+} \\ -1\ , x \in h_{ps}^{-} \end{cases}$$

$$hq_{psr2}(x,y) = hq_{pr}(y) = \begin{cases} +1\ , \ y \in h_{pr}^{+} \\ -1\ , x \in h_{pr}^{-} \end{cases}$$

In group of one order p contains $3*4^{p-1}$ functions.

Requirements of high efficiency of the computing systems applied in these areas can be satisfied as due to development of new methods and algorithms of digital signals processing (DSP), and by means of multiprocessing means of in parallel-conveyor calculations. In other words, the requirements for algorithms of fast spectral transformation consists first of all, in a minimality of operating quantity, simplicity of each operation and a minimality of the demanded volume of operative memory.

2. THE COMPARATIVE ANALYSIS OF FAST TRANSFORMATIONS ALGORITHMS

Advantage of algorithms of fast transformations in bases of Haar and Harmut sharply increases with transition to two-dimensional bases. So for performance of fast transformation of Fure (FFT) 10240 arithmetic operations are required, as much arithmetic operations are demanded by algorithm of fast transformation of Walsh (WFT), but in algorithm WFT there are no operations over complex numbers in difference from FFT. And for performance of fast transformation of Harmut 3068 arithmetic operations are required that 3,34 times are less than FFT and WFT. FT Haar demands 2046 arithmetic operations. These are 5 times less than FFT and WFT and 1,5 times less than Harmut.

Let's compare also these algorithms for two-dimensional bases. So for performance two-dimensional FFT and FWT 209 715 200 arithmetic operations for performance of fast transformation of Harmut 18 874 368 arithmetic operations are required. It in 11,11 times is less than FFT and FWT. For performance of fast transformation of Haar 8 388 608 arithmetic operations that 25 times less than FFT and FWT and therefore are equal are required, 2,25 times are less than FT Harmut.

CONCLUSION

In summary we can conclude that existed one-dimensional fast transformations of Haar can be used in DSP for two dimensional signals. In that condition, it is performed in two stages: 1) by column 2) by row. Therefore it gives us much more success on performance two-dimensional FFT and FWT 209 715 200 arithmetic operations for performance of fast transformation of Harmut 18 874 368 arithmetic operations are required. It in 11,11 times is less than FFT and FWT. Thus, new types of two-dimensional fast transformations are offered. Possibilities of distribution of fast algorithms in one-dimensional piecewise-polynomial bases on two-dimensional bases, also advantages the offered bases in comparison with existing are shown. Results of the comparative analysis of the offered bases with other bases by quantity of demanded arithmetic operations, and also an example of application of the offered bases in problems of image processing are resulted.

REFERENCES

[1] Alfred Haar, "Zur Theorie der Orthogonalen Funktionensysteme, Math" Journal of Mathematische Annalen, Springer Berlin / Heidelberg Vol. 71,(1), 1911 pp.38-53.

[2] M. Vetterli, "Wavelets, approximation and compression," IEEE SP Mag., 2001, to appear.

[3] A. Cohen, I. Daubechies, O. G. Guleryuz, and M. T. Orchard, "On the importance of combining wavelet-based nonlinear approximation with coding strategies," IEEE Trans. Info. Theory, 2000, submitted.

[4] P. Prandoni and M. Vetterli, "Approximation and compression of piecewise smooth functions," Phil. Trans. Royal Society London, August 1999.

[5] P. L. Dragotti and M. Vetterli, "Wavelet transform footprints: Catching singualarities for compression and denoising," in Proc. IEEE Int. Conf. on Image Proc., Vancouver, Canada,Sep. 2000.

[6] E. J. Cand`es and D. L. Donoho, "Curvelets – a surprisingly effective nonadaptive representation for objects with edges," in Curve and Surface Fitting, A. Cohen, C. Rabut, and L. L. Schumaker, Eds., Saint-Malo, 1999, Vanderbilt University Press.

[7] S. Mallat, A Wavelet Tour of Signal Processing, Academic Press, 2nd edition, 1999.

[8] M. N. Do and M. Vetterli, "Orthonormal finite ridgelet transform for image compression," in Proc. IEEE Int. Conf. on Image Proc. (ICIP), Vancouver, Canada, Sep. 2000.

[9] E. Le Pennec and S. Mallat, "Image compression with geometric wavelets," in Proc. IEEE Int. Conf. on Image Proc., Vancouver, Canada, Sep. 2000.

[10] N. Lee and B. J. Lucier, "Wavelet methods for inverting the radon transform with noisy data," IEEE Trans. Image Proc., vol. 10, pp. 79 – 94, 2001.

[11] J. Lin, A. F. Laine, and S. R. Bergmann, "Improving pet-based physiological quantification through methods of wavelet denoising," IEEE Trans. Bio. Eng., vol. 48, pp. 202 –212, 2001.

[12] Zaynidinov H.N., Avaz Mirzaev., Hsein-Ping Kew., Do-Un Jeong. Piecewise-polynomial Bases and Specialized Processor for Digital Signal Processing. 2009 IEEE Symposium on Industrial Electronics and Applications (ISIEA 2009), October 4-6, 2009, Kuala Lumpur, Malaysia, p. 509-603.

[13] Dannanjay Singh, Zaynidinov H.N., Hoon Jae Lee. Piecewise-quadratic Harmut basis function and their application to problems in digital signal processing. International Journal of Communication Systems, John Wiley & Sons, Ltd. ,DOI: 10.1002/dac.1093, Jan. 2010. SCI-E. www.interscience.wiley.com

[14] Bohan K.A., Koroleva N.A. Estimation parameters of efficiency two-dimensional transformations Haar // Electronic modelling. 2003. №5. C.45 -56.

BUILDING A CUBIC SPLINE THAT IS IMPORTANT IN DIGITAL PROCESSING OF SIGNALS

Kuchkarov M.A., Tashkent University of Information Technology named after Muhammad al-Khorezmi, Uzbekistan.

Annotation. *This article, the construction of the cubic spline function, which is important in the digital processing of signals, is sequentially covered. The formulas for determining the coefficients of the spline function are given.*

Key words: *spline-function, interpolation cubic spline, interpolation, coefficients.*

In recent years, digital functions have been used extensively to address and analyze digital signals. An analysis of the existing literature shows that interpolation and grinding splines are used by the approximation method, and polynomial and basis splines by type of imaging. Interpolation discs are such splines that satisfy the set of boundary conditions and conditions at the points of the function definition, while grinding splines are related to the optimization of various functions.

The second spline is a natural cubic spline, and it be built as follows.

There is searched for the $S(x) = S_i(x)$ function in each $[x_i, x_{i+1}]$ $(i = \overline{1, n-1})$ interval in the form of a tertiary multiplex.

$$S_i(x) = a_i + b_i(x - x_i) + \frac{c_i}{2}(x - x_i)^2 + \frac{d_i}{6}(x - x_i)^3, \quad (1)$$

$x_{i-1} \le x \le x_i, i = 1,2,...,N,$

Here are the coefficients that a_i, b_i, c_i, d_i needs to determine.

These coefficients are calculated as follows.

$$S_i'(x) = b_i + c_i(x - x_i) + \frac{d_i}{2}(x - x_i)^2,$$

$$S_i''(x) = c_i + d_i(x - x_i), \quad S_i'''(x) = d_i,$$

the coefficients should be written based on the following

$$a_i = S_i(x_i),\ b_i = S_i'(x_i),\ c_i = S_i''(x_i),\ d_i = S_i'''(x).$$

the coefficients should be defined $a_i = f(x_i),\ \ i = 1,2,...,N$, under the conditions of interpolation. In addition, the continuity of the function should be recorded as follows.

$$S_i(x_i) = S_{i+1}(x_i),\ \ i = 1,2,\ldots,N-1.$$

Given the expressions for the $S(x)$ cubic spline, express the $i = 1,2,...,N-1$ equations as follows

$$a_i = a_{i+1} + b_{i+1}(x_i - x_{i+1}) + \frac{c_{i+1}}{2}(x_i - x_{i+1})^2 + \frac{d_{i+1}}{6}(x_i - x_{i+1})^3.$$

If we define $h_i = x_i - x_{i-1}$, in general terms the equations come to (2).

$$h_i b_i - \frac{h_i^2}{2} c_i + \frac{h_i^3}{6} d_i = f_i - f_{i-1},\ i = 1,2,...,N \quad (2)$$

Write the condition of continuity for the first order yield

$$S_i^{'}(x_i) = S^{'}{}_{i+1}(x_i),\ i = 1,2,...,N-1$$

After the above substitution, the statements appear (3)

$$h_i c_i - \frac{d_i^2}{2} h_i = b_i - b_{i-1},\ \ i = 2,3,...,N\ . \quad (3)$$

From the continuum condition of the second order derivative, the equations (4)

$$h_i d_i = c_i - c_{i-1},\ \ i = 2,3,...,N\ . \quad (4)$$

By combining (2) - (4), a system will be gained of 3N - 2 equations for an unknown

$b_i, c_i, d_i,\ \ i = 1,2,\ldots,N$ of 3N.

The two missing equations can be obtained by setting a boundary condition for cubic spline. For instance, it is said that the $f(x)$ function satisfies the $f''(a) = f''(b) = 0$ condition. From this it will be gained the following $S_1^{''}(x_0) = 0, S_N^{''}(x_N) = 0,$ it's, $c_1 - d_1 h_1 = 0,\ c_N = 0.$

Note that the condition $c_1 - d_1 h_1 = 0,$ coincides with equation (4) for $i = 1$, if it put $c_0 = 0$. Thus, a closed system is come of equations for determining the coefficients of the cubic spline:

$$h_i d_i = c_i - c_{i-1},\ \ i = 1,2,\ldots,N,\ \ c_0 = c_N = 0, \quad (5)$$

$$h_i c_i - \frac{h_i^2}{2} d_i = b_i - b_{i-1},\ \ i = 1,2,\ldots,N, \quad (6)$$

$$h_i b_i - \frac{h_i^2}{2} c_i + \frac{h_i^3}{6} d_i = f_i - f_{i-1},\ i = 1,2,\ldots,N. \quad (7)$$

to make sure that this system has a unique solution[1,2,3]. It is eliminated the variables $b_i, d_i,\ i = 1,2,...,N-1$from (5) - (7), and it will be obtain a system containing only $c_i,\ i = 1,2,...,N-1$. To do this, consider two neighboring equations (7):

$$b_i = \frac{h_i}{2} c_i - \frac{h_i^2}{6} d_i + \frac{f_i - f_{i-1}}{h_i},$$

$$b_{i-1} = \frac{h_{i-1}}{2} c_{i-1} - \frac{h_{i-1}^2}{6} d_{i-1} + \frac{f_{i-1} - f_{i-2}}{h_{i-1}}$$

and subtract the second equation from the first. Then we get

$$b_i - b_{i-1} = \frac{1}{2}(h_i c_i - h_{i-1} c_{i-1}) - \frac{1}{6}(h_i^2 d_i - h_{i-1}^2 d_{i-1}) + \frac{f_i - f_{i-1}}{h_i} - \frac{f_{i-1} - f_{i-2}}{h_{i-1}}$$

Substituting the expression found for $b_i - b_{i-1}$ the right side of equation (6), we obtain

$$h_i c_i + h_{i-1} c_{i-1} - \frac{h_{i-1}^2}{3} d_{i-1} - \frac{2h_i^2}{3} d_i = 2\left(\frac{f_i - f_{i-1}}{h_i} - \frac{f_{i-1} - f_{i-2}}{h_{i-1}}\right). \quad (8)$$

Next, from equation (5) it is obtain

$$h_i^2 d_i = h_i(c_i - c_{i-1}), \quad h_{i-1}^2 d_{i-1} = h_{i-1}(c_{i-1} - c_{i-2})$$

and substituting these expressions in (15), it is arrived at the equation

$$h_{i-1}c_{i-2} + 2(h_{i-1} + h_i)c_{i-1} + h_i c_i = 6\left(\frac{f_i - f_{i-1}}{h_i} - \frac{f_{i-1} - f_{i-2}}{h_{i-1}}\right).$$

Finally, to determine the coefficients c_i, the system is obtained of equations

$$h_i c_{i-1} + 2(h_i + h_{i+1})c_i + h_{i+1}c_{i+1} = 6\left(\frac{f_{i+1} - f_i}{h_{i+1}} - \frac{f_i - f_{i-1}}{h_i}\right). \quad (9)$$

$$i = 1,2,\ldots,N, \; c_0 = c_N = 0.$$

Due to the diagonal prevalence, system (9) has a unique solution. Since the matrix of the system is three diagonal, the solution is obtained by the progonka method. From the found coefficients c_i, (i=1,…,n) the coefficients b_i and d_i are determined using explicit formulas

$$d_i = \frac{c_i - c_{i-1}}{h_i}, \quad b_i = \frac{h_i}{2}c_i - \frac{h_i^2}{6}d_i + \frac{f_i - f_{i-1}}{h_i}, \quad (9)$$

$$i = 1,2,\ldots,N.$$

Thus, it is proved that there exists a unique cubic spline defined by conditions 1) - 2) and the boundary conditions $S''(a) = S''(b) = 0$ [4,5].

REFERENCES

1.Баркалов К.А. Образовательный комплекс «Параллельные числен- ные методы». Лекционные материалы. Нижегородский государственный университет им. Н.И. Лобачевского Факультет вычислительной математики и кибернетики. Нижний Новгород 2011 г.

2.Бахвалов Н.С., Жидков Н.П., Кобельков Г.М. Численные методы. – М.: Наука, 1987.

3.Богданов В.В., Волков Ю.С. Условия формосохранения при интерполяции кубическими сплайнами. 2019 г. https://www.researchgate.net/publication/333602092.

4.Завьялов Ю.С., Квасов Б.И., Мирошниченко В.Л., Методы сплайн-функций. Москва.: Наука, 1980. 352 с.

5,Зайнидинов Х.Н., Бахромов С.А., Азимов Б.Р. Биомедисина сигналларни интерполяцион кубик сплайн моделини қуриш. Мухаммад ал-Хоразмий авлодлари, № 4 (10), декабрь 2019, 14-17 б.

RAQAMLI SIGNALLARGA ISHLOV BERISH BOSQICHLARI

Jurayev U.S., Guliston davlat universiteti, O'zbekiston.

Annotatsiya. *Bugungi kunda jamiyatni axborotlashtirish jarayonini tezlashtirishni ta'minlAAaydigan yangi axborot texnologiyalarini yaratish va rivojlantirishdan boshqa dolzarb va ijodiy vazifalar bo'lmasa kerak. Ushbu maqolada raqamli signallara ishlov berishning rivojlanish tendentsiyalari, ishlov berish bosqichlari hamda hozirgi kundagi axamiyati to'g'risida bayon qilingan.*

Kalit so'zlar: *raqamli signal, raqamli filtr, spektral taxlil, tezkor algoritmlar, kvantlash nazariyasi, ko'p o'lchamli signallar, adaptiv, optimal, chastota, signal protsessor.*

Аннотация. *Сегодня нет более актуальной и творческой задачи, чем создание и развитие новых информационных технологий, которые ускорят процесс информатизации общества. В данной статье описываются тенденции развития, этапы обработки и текущее значение цифровой обработки сигналов.*

Ключевые слова: *цифровой сигнал, цифровой фильтр, спектральный анализ, быстрые алгоритмы, теория квантования, многомерные сигналы, адаптивный, оптимальный, частота, процессор сигналов.*

Annotation. *Today there is no moremлв urgent and creative task than the creation and development of new information technologies that will accelerate the process of informatization of society. This article describes development trends, processing steps, and the current importance of digital signal processing.*

Keywords: *digital signal, digital filter, spectral analysis, fast algorithms, quantization theory, multidimensional signals, adaptive, optimal, frequency, signal processor.*

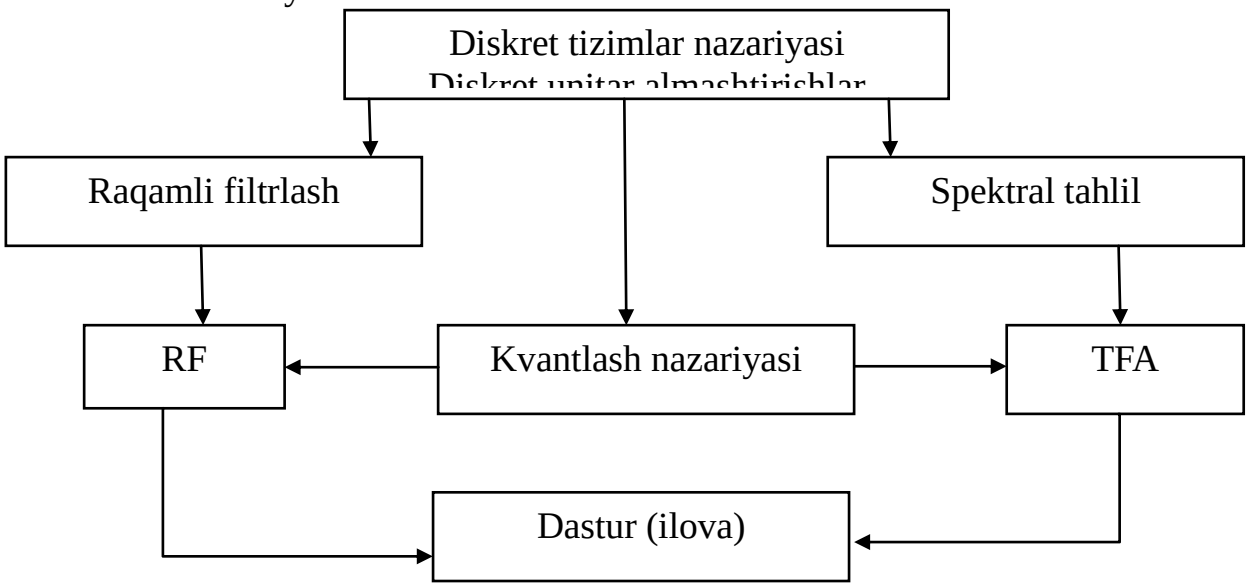
XXI-asrda axborot texnologiyalarini rivojlantirish holati va istiqbollari dunyoda telekommunikatsiya va kompyuterlashtirish dunyosidagi eng dinamik va tez rivojlanayotgan texnologiyalardan biri bo'lgan raqamli signallarni qayta ishlash texnologiyasini shakllantirish va keng qo'llash bilan tavsiflanadi. Raqamli signallarga ishlov berish (RSIB) real vaqt rejimida ma'lumotlarni qabul qilish, qayta ishlash, qisqartirish va uzatish muammolarini hal qilish uchun mo'ljallangan informatika hisoblanadi. RSIB usullari va vositalari aloqa va boshqaruv tizimlari, radiotexnika va elektronika, akustika va seysmologiya, radioeshittirish va televidenie, o'lchash uskunalari va asboblarni ishlab chiqarish kabi turli sohalarda ishlaydigan olimlar va mutaxassislarda katta qiziqish uyg'otmoqda.

Uzluksiz ma'lumotni qayta ishlash muammolarini uning diskret qiymatlarini o'zgartirishning sodda muammolari bilan almashtirishda taniqli olimlarning XX asrning 30-50-yillarida qilgan ishlari muhim ahamiyat kasb etdi. 40-yillarda birinchi kompyuterlarning paydo bo'lishi bilan hisoblash matematikasida yangi fundamental ilmiy sohaning paydo bo'lishi haqiqatga aylandi, uning bo'limlaridan

biri raqamli ma'lumotlarni qayta ishlash uchun algoritmlarni ko'rib chiqish mumkin. Biroq, o'sha yillarda ishlatilgan kompyuterlarning cheklangan hisoblash resurslari real vaqt rejimida ma'lumotlarni qayta ishlashga imkon bermadi [1]. Bu faqat real jarayonlarni modellashtirish bilan bog'liq bo'lishi mumkin. 60-yillarda kichik elektron xisoblash mashinalari sinfining paydo bo'lishi bilan vaziyat tubdan o'zgara boshladi, asosan real vaqt rejimida ma'lumotlarni boshqarish va ishlov berish muammolarini hal qilishga qaratilgan. Kichik elektron xisoblash mashinalaridan foydalangan holda raqamli signallar yordamida tabiatdagi signallarni qayta ishlash, o'zgartirish va uzatish imkoniyati, ko'plab sohalarda, birinchi navbatda aloqa, gidroakustika va nutqni qayta ishlash sohasida ishlaydigan mutaxassislarning e'tiborini tortdi.

1-BOSQICH. RAQAMLI FILTRLASH VA SPEKTRAL TAHLIL

Rivojlanishning ushbu bosqichida (1965-1975) RSIB nazariyasining asosiy predmeti raqamli filtrlash va spektral tahlil edi (1-rasm) va ikkala yo'nalish ham chastota namoyishlarining umumiy nuqtai nazaridan ko'rib chiqildi. Rivojlanayotgan hududlarning umumiy asosi chastota tanlashning raqamli filtrlarini sintez qilish edi. RSIB nazariyasining asosiy printsiplari o'sha paytda ma'lum bo'lgan mashina algoritmlari to'plamidan va birinchi navbatda tezkor Fur’ye almashtirishlari (TFA) algoritmidan foydalanib, diskret tizimlar nazariyasi va sxemalar nazariyasida amalda sinab ko'rildi.



1-rasm. RSIB nazariyasining asosiy predmeti raqamli filtrlash va spektral tahlil

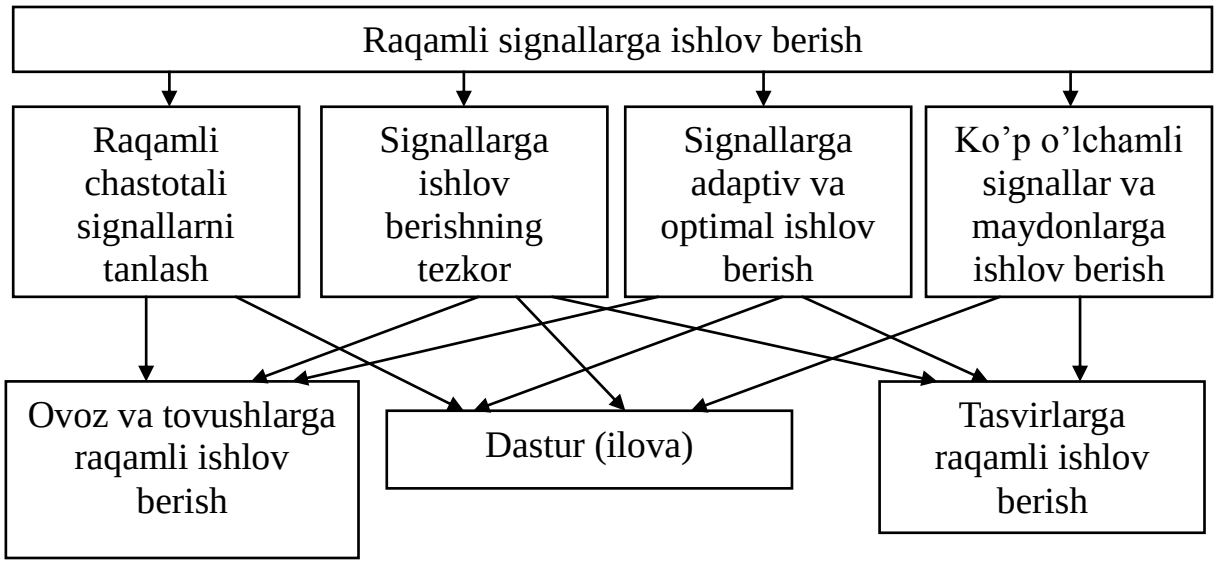
Ushbu yillar davomida samarali yechimi topilgan asosiy muammolar quyidagilardir: raqamli filtr (RF) uzatish funktsiyasini cheklangan (CHR filtrlari) va cheksiz (CHR filtrlari) impulsli reaktsiyaga ega filtrlar sinfiga mashinaning yaqinlashishi, yuqori tezlikda to'planish algoritmlarini va past shovqinli tuzilmalarni ishlab chiqish, cheksiz filtrlari, polosali filtrlari va tezkor Fur’ye almashtirishlari algoritmidan foydalanishga asoslangan raqamli spektr analizatorlarini qurish.

RSIB nazariyasiga birinchi bo’lib katta hissa qo’shgan amerikalik olimlar B. Gold va Ch. Raiderlar hisoblanishadi. Ularning monografiyasi [2] RSIB mutaxassislarining birinchi qo'llanmasi edi. RSIB nazariyasini yangi ilmiy yo'nalish sifatida rivojlantirishning birinchi bosqichini yakunlovchi fundamental ish, shubhasiz, L. Rabiner, B. Goldning kitobidir [3]. Biroz vaqt o'tgach, mualliflar tomonidan RSIB bo'yicha birinchi darsliklardan biri paydo bo'ldi: A. Oppengeym va R. Shafer [4]. Boshqa ma'lum ishlarning aksariyati, qoida tariqasida, raqamli filtrlash bilan bog'liq yoki bir qator dasturlarda RSIB usullaridan foydalanish. Bunga mahalliy olimlar katta hissa qo'shdilar.

Ushbu davrda raqamli filtrlar va spektr analizatorlarini texnik jihatdan amalga oshirish imkoniyatlarini kichik integrallashgan kompyuterlar yoki o'rta darajadagi axborot tizimlarida qurilgan maxsus qurilmalar yordamida real vaqt rejimida kompyuter simulyatsiyasi bosqichi sifatida tavsiflash mumkin.

2-BOSQICH. KO’P TEZLIKLI FILTRLASH VA SIGNALLARNI ADAVTIV QAYTA ISHLASH

70-yillarning boshlarida birinchi chipli mikroprotsessorlar (MP) paydo bo'ldi - kompyuter inqilobining yangi to'lqinining "peshqadamlari". RSIB uskunalari va kompyuter texnologiyalarini rivojlantirishda yangi bosqich boshlanadi. Yangi imkoniyatlar ochiladi va yangi muammolar yuzaga keladi. RSIB nazariyasi shartli ravishda 1975 yildan 1985 yilgacha cheklanishi mumkin bo'lgan rivojlanishning navbatdagi bosqichiga kiritilgan. Aynan shu davrda zamonaviy RSIB nazariyasining o'zaro bog'liq bo'lgan to'rtta asosiy yo'nalishi shakllangan (2-rasm).



2-rasm. RSIB nazariyasining o'zaro bog'liq bo'lgan to'rtta asosiy yo'nalishi

Birinchi yo'nalish - signallarni raqamli chastotali tanlash, raqamli polosali filtrlari va ularning to'plamlarini loyihalash sohasidagi yutuqlarni birlashtiradi va tizimlashtiradi. Ushbu yo'nalishdagi eng asl ishlar vaqt va chastotada yupqalash effektlari asosida signallarni ko'p tezlikli ishlash nazariyasini rivojlantirish bilan bog'liq edi.

Ikkinchi yo'nalish - tezkor ishlov berish algoritmlari, konversiya operatsiyalarining "ko'payishini" olib tashlash va vaqtni ko'paytiradigan

operatsiyalarni qo'shimcha va siljish operatsiyalari bilan almashtirish (Tezkor Fur’ye almashtirishlari algoritmining ko'p sonli o'zgartirilishi va son-teoretik o'zgartirish usullari) orqali yuqori tezlikdagi RSIB algoritmlarini yaratish.

Uchinchi yo'nalish- Uchinchi yo'nalish - adaptivli va optimal ishlov berish, optimal filtrlash muammolarini (Wiener, Kalman filtrlari va boshqalarni) hal qilish uchun keng usullarni qamrab oladi va o'rganilayotgan dinamik jarayonning tabiati haqida dastlabki noaniqlik sharoitida signalni qayta ishlash.

To'rtinchi yo'nalish- ko'p o'lchovli signallar va maydonlarga ishlov berish, ko'p o'lchovli raqamli tizimlar uchun bir o'lchovli signallarni qayta ishlashning tabiiy rivojlanishi.

Ko'rsatib o'tilgan yo'nalishlar bir-biri bilan uzviy bog'liq bo'lib, bu munosabatlar barcha to'rt yo'nalishni "oziqlantiruvchi" umumiy matematik asosga va boshqa yo'nalishlarda asosiy printsiplar va usullardan bevosita foydalanishga asoslangan.

Rasmlarni qayta ishlashning raqamli usullari moslashuvchanlik va samaradorlik jihatidan analoglardan ancha ustundir. Tasvirlarni chiziqli ishlov berishda turli xil ikki o'lchovli unitar almashtirishlar keng qo'llaniladi: Furye, kosinus, sinus, qiyalik, Adamar, Haar, singular va boshqalar, bu ma'lumot miqdorini bir darajaga yoki boshqasiga kamaytirishi mumkin.

Uzluksiz signallarni bog'lanmagan koeffitsientlar to'plamiga aylantirish usuli Hotelling, Karhunen va Loev tomonidan ishlab chiqilgan va chiziqli o'zgarishlarning samaradorligini baholashda keng qo'llaniladigan Karhunen-Loev transformatsiyasi deb nomlangan. Rasmlarni raqamli qayta ishlash ularni qayta tiklash va takomillashtirish uchun ishlatiladi, rasm tafsilotlarini ajratib ko'rsatish, tasvirni aniqlash, raqamli ichki va kadrlararo kodlash.

Ovoz ma'lumotlarini qayta ishlashning raqamli usullari nutqni qayta ishlashda va kodlashda, shuningdek keng polosali audio signallarni samarali konversiyalashda keng qo'llaniladi. Hozirda ishlatilayotgan vokalistlar ovoz ma'lumotlarini o'nlab yoki yuzlab marta kamaytirishi mumkin. Keng polosali audio signallari uchun kodlagichlar kamroq samaralidir, ammo ular asl nusxadan ajralib bo'lmaydigan audio sifatini ta'minlaydi.

3-BOSQICH. SIGNAL PROTSESSORLARIDA ENG OPTIMAL LOYIHALASH

80-yillarning birinchi yarmida avval NEC (Yaponiya), so'ngra Texas Instruments (AQSh) mPD7720 va TMS32010 birinchi signal protsessorlarini sanoat ishlab chiqarilishini e'lon qildi va shu bilan RSIB texnologiyasida yangi davr - signallarini qayta ishlash davri boshlandi. Mikroprosessor tizimlarining yangi sinfi aslida ichki arxitekturada klassik RSIB algoritmlarini yuqori darajada samarali dasturiy va apparat ta'minotiga yo'naltirilgan bitta chipli mikrokompyuterlar oilasi edi. Nisbatan qisqa vaqt ichida - 15 yil davomida raqamli signallarni qayta ishlash protsessorlari (RSIB) rivojlanishning bir necha bosqichlaridan o'tdi. RSIB tizimlarini maqbul avtomatlashtirilgan loyihalashning umumiy tushunchasi shakllantirilgan. Tizimni simulyatsiya qilish bosqichidan boshlab signal uzatish protsessorlariga va juda katta integral mikrosxema (JKIM)

signallarini qayta ishlashga qadar avtomatlashtirilgan dizaynni qo'llab-quvvatlash uchun kuchli dasturiy vositalar yaratilmoqda.

Bular tarkibiga The MathWorks, Inc., firmasining MATLAB Hyperception, Inc., firmasining Hypersignal singdirilgan qobiqlari va Momentum Data Systems (AQSH) firmasining QEDesign, Radis, Ltd (Rossiya) firmasining DFID va PICLOR raqamli filtr sintez paketlari kiradi.

4-BOSQICH. BIR KIRISTALLI KO'P PROTSESSORLI TIZIMLAR VA DASTURLASHTIRILADIGAN MANTIQIY INTEGRAL MIKROSXEMALARNI OPTIMAL LOYIHALASH

90-yillarning ikkinchi yarmida signallarni qayta ishlash usullari va texnikasini rivojlantirishning hozirgi bosqichi bitta chipli ko'p protsessorli larning yangi noyob imkoniyatlari (TMS320C80 oilasi) va dasturlashtiriladigan mantiqiy integral mikrosxemalar (DMIM) asosida arxitektura jihatdan qayta dasturlashtiriladigan juda katta integral mikrosxema (JKIM) RSIBdan foydalanish bilan belgilanadi.

RSIB tizimini qurishning yangi kontseptsiyasi DMIMlarning potentsial imkoniyatlaridan keng foydalanishga asoslangan bo'lib, minimal texnik xarajatlar bilan belgilangan sifat ko'rsatkichlariga erishilishini kafolatlaydigan optimal loyihalsh uslubiga asoslangan. RSIB nazariyasi va texnologiyasining eng dolzarb vazifalari quyidagilardan iborat:

- Turli yo'nalishlarda raqamli signallarni qayta ishlash usullari va algoritmlarini tizimlashtirish va RSIB tizimlarini avtomatlashtirilgan loyihalash uchun dasturiy paketlarni yaratish;
- signal protsessorlari va DMIMlarda RSIB tizimlarini maqbul dizayni uchun usul va dastur paketlarini ishlab chiqish;
- RSIB nazariyasining asosiy yo'nalishlarida yangi tushunchalarni ishlab chiqish - ko'p tezlikli ishlov berish, tezkor algoritmlar, moslashuvchan ishlov berish, spektral baholash, vaqt chastotasini qayta ishlash, to'lqinli va fraktal o'zgarishlar, chiziqli bo'lmagan filtrlash, ko'p o'lchovli signallarni qayta ishlash va boshqalar.

RSIB nazariyasining mustaqil ilmiy soha sifatida rivojlanishini tahlil qilib, biz uni boshqa bilim sohalaridan ajratib turadigan bir qator masalalarni va xarakterli xususiyatlarni ajratib olishimiz mumkin. Berilgan matematik transformatsiya operatorini amalga oshiruvchi RSIB tizimini yoki qurilmasini ishlab chiquvchisi uchun konversiya natijalaridan qanday foydalanish muhim emas, vazifa birinchi navbatda hisoblash va apparat xarajatlarini minimallashtirishdir. Bitta va bir xil turdagi transformatsiya algoritmlarining natijalari, masalan, chiziqli yig'ilish, istalgan chastota xususiyatining takrorlanishini ta'minlaydi, ikkinchisida - filtrlash, uchinchida - adaptiv tuzatish yoki eho kompensatsiya qilishi imkoniyatlarini beradi.

Adabiyotlar

1. *Витязев В.В. Цифровая обработка сигналов; ретроспектива и современное состояние //Электросвязь. -1997.- №6.*
2. *Голд Б., Рэйдер Ч. Цифровая обработка сигналов / Под ред. М. Трахтмана. - М.: Сов. радио, 1973. - 367 с.*
3. *Рабиннер Л., Голд Б. Теория и применение цифровой обработки сигналов. - М.: Мир, 1978. - 848 с.*

Оппенгейм А.В., Шафер Р.В. Цифровая обработка сигналов. - М.: Связь, 1979. - 416 с

COMPARATIVE PERFORMANCE ANALYSIS OF CLUSTERING ALGORITHMS USING KNIME

Imdad Hussain, Syed Mujtaba Hassan, Syeda Tahira Batool. Dept of Computer Science, Riphah International University, Lahore campus, Pakistan.

Abstract. *This paper discusses the most useful and popular clustering algorithms used by Machine Learning systems, especially artificial intelligent systems. It focuses on the practical clustering algorithms. Pakistani air quality monitoring dataset used in this document. It is technique that create group according to the rule of maximum same within the class and minim the same between classes. The results of the work show which algorithm is most convenient for a given data set. The common analysis of the dataset, including clustering algorithms, are implemented in the KNIME tool.*

Keywords: *Cluster, Fuzzy C Means, K Means, KNIME, Analysis, Pie Chart.*

1. INTRODUCTION

Data mining / Machine learning is a very important area of research. There are many data mining techniques which we are using in our daily life. We use these techniques in education, medicine, weather, business, banking etc. Before working on all these techniques it is important for us to understand it. Clustering is under of datasets which are observational, with the aim of discovering unsuspected relationships between data sets and to summarize [2].

Description of data possible through visualization, clustering, association and sequential analysis. Cluster analysis is also recognized technique that make the crustal of our data. Enter each point in the data record in. Just single cluster is the foundation of common grouping method in which, as a grouping algorithm, it truly divides the untagged data set into different groups based on similarity [3].

In second terms, clustering is a division of data sets into groups so that the data of every group share certain general characteristics. Typical conditions for clustering in data mining are expandability, the capability to treat distinct kinds of attributes, the findings of groups with different forms, the minimum knowledge of the field to set the input parameters, the capability to manage noisy data, incremental grouping and indifference to input of input data, large dimensionality, grouping founded on constraints, interpretation and usability[4].

We used KNIME tool to apply fuzzy c-means and K Means algorithm on dataset and compere their result to check which give better result.

2. RESEARCH QUESTION

In this research paper we discussed different clustering algorithms and apply algorithms on dataset and compere their result.

1. Comparative Performance Analysis of Clustering Algorithms and which algorithms is performing better clustering efficiency on our given dataset?

3. LITERATURE REVIEW

Clustering is a domain of active research in data mining with several informed methods. Few of the researchers improved the data clustering technique, while others implemented new ones, and few analyzed and compared the existing pooling algorithms. Most of the researcher performing different clustering technique on them apply different technique and compere their result. They apply different algorithms and check their result and compere them [5]**.** In the researchers used the agglomeration technique to construct the genogram and used some technique and make the crustal. Most of researcher used agglomerative clustering algorithms based on similarities and presented their effectiveness. Respect for the different grouping algorithms based on the following factors: size of the data set, verity of data set and tool used for grouping. Many researchers used Math Lab tool and applying different clustering algorithms on different data sets and compare their result on the base of number of iteration and number of cluster and clustering percentage etc. [6] Some researcher used waka tool and applying different clustering algorithms like as K Means EM Dentistry based algorithm and on different dataset compare their number of cluster distribution number of iterations time taken to build a model and get result [7]. A few researcher applying mass based and density based clustering technique on numeric data and compere their run time take to build a model time number of cluster and unsinged cluster etc. and fined the efficiency of different algorithms [8].Some people applying different clustering algorithms on large data source and compere their efficiency. A few researcher applying clustering technique on different types of datasets and compere their result on base of number of classes number of iteration error and saq error etc.

4. METHODOLOGY

here are many clustering techniques like EM, Canopy, Cobweb, K-Means, and Fuzzy C Means etc. We used k means and fuzzy c means on Pak weather data set and analyze their efficiency using KNIME tool.

4.1 Working

At first step, file reader node was used for access and import dataset. After importing, it show data in table. To remove missing values in data we used missing value node we connect it with file reader and configure it apply. Next step of working was partitioning of data because we used 20% training and 80% testing data. For this purpose we used partitioning node. After this we applied clustering algorithms for this cluster assigner node used that assign different cluster to data. Color manger node used to assigns the different color to cluster. At the end we used pie chart to show our result.

4.2 Data access

First step in workflow is to access the file path of Air-Quality-Monitoring-Dataset-Pakistan for further implementation.

4.3 Data partition

Partitioning node is used to divide dataset into test and train data. Here, for training 80% data used and 20% data for testing.

4.4 K-MEANS

K Means is a simple type of algorithm. It is the most popular algorithm. Basically it creates cluster and analyze the data on the base of position and gap between input data points. The division of objects into mutually exclusive groups (K) is carried out in such a way that the objects inside each group stay as close to each other as possible but as far as feasible from the objects in different groups. Partitioning objects into k non-empty subsets. Identify the group centroids (midpoint) of the current partition [9]

4.5 FUZZY C-MEANS

The Fuzzy C Means is a known unsupervised learning to reveal the underlying structure of the data. It permits every data point to own to multiple groups, with a degree of belonging to each of them. This is employ for analyses based on the gap between several data points input this calculation is utilized for investigation dependent on the separation between different information focuses input. Gatherings are shaped dependent on the separation between information focuses and bunch focuses are framed for each group.in actuality [10]

5. EXPERIMENT AND ANALYSIS

K Means and Fuzzy C Means clustering algorithm were implemented in Air-Quality-Monitoring-Dataset-Pakistan using KNIME.

We have used the best popular data mining tool for analysis open source tool KNIME. The Air-Quality-Monitoring-Dataset-Pakistan used and perform of a comprehensive set of clustering algorithms has been analyzed. The analysis has been performing on a Windows 10 operating system with Intel® Core ™ i4 CPU, 2.0 GHz Processor and 6.00 GB RAM

5.1 Dataset

In this paper I used kaggle website to download dataset of Air-Quality-Monitoring-Dataset-Pakistan. In this dataset there are 7 Attributes. Data set Date, Time, Temperature, Humidity, CO_AQI, NO2_AQI, and SO2_AQ. This data consist record 13/08/2018 - 09/08/201

5.2 K-MEANS

In this working first take file reader which is used for data reading we import data set in file reader configure if and apply it and show the data in table in this table we note some missing value to remove these missing value we used missing value node we connect it with file reader and configure it apply. Next step of our working is partitioning of our data because we use 20% training and 80% testing data for this purpose we use partitioning node. After this we use our clustering algorithms. In the first step followed by using K-Means Algorithm on K-Means Algorithm node =>Cluster Assigner node => Color Manager node => Pie Chart node, for displaying cluster grouping data.

Pie Chart

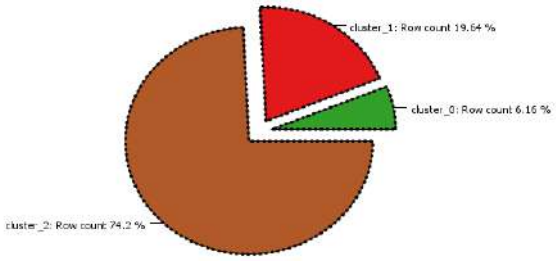


Figure 1. K Means Pie Chart

Pie chart show our result that we get by using K Means algorithms. The Pie chart show us three different part these part are called cluster it means our data divide in to three different parts on their similarity. K Means algorithms create three cluster. Cluster0 6.16% cluster1 19.64% and cluster2 74.2%.

5.3 FUZZY C-MEANS

In this working first take file reader which is used for data reading we import data set in file reader configure if and apply it and show the data in table in this table we note some missing value to remove these missing value we used missing value node we connect it with file reader and configure it apply. Next step of our working is partitioning of our data because we use 20% training and 80% testing data for this purpose we use partitioning node after this we select Fuzzy C-Means Algorithm node =>Cluster Assigner node => Color Manager Node => Pie Chart node, for displaying cluster grouping data.

Pie Chart

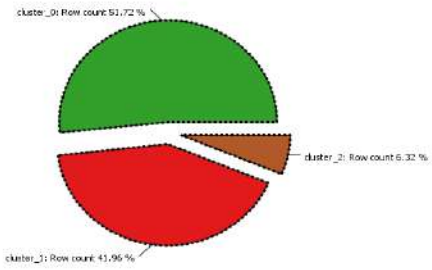


Figure 2. K Means Pie Chart

Pie chart show our result that we get by using Fuzzy C Means algorithms. The Pie chart show us three different part these part are called cluster it means our data divide in to three different parts on their similarity. K Means algorithms create three cluster. Cluster0 51.72%% cluster1 41.96% and cluster2 6.32%. When we see these clustering partitioning we compare it with K Means result at the end we find result Fuzzy C Means give us better result.

6. RESULT AND DISCUSSION

On the dataset “Air-Quality-Monitoring-Dataset-Pakistan for Clustering” we used two algorithms which are K-Means and Fuzzy c-Means on the KNIME tool for

clustering. After applying both algorithms Fuzzy c-Means provide better grouping of data instead of K-Means. In result number of cluster are same and also partitioning percentage same but at the end Fuzzy C-Means give us better cluster the data present in clusters is better average then K Means so On our given dataset Fuzzy C means is better as displayed in below table.

Table 1. Results

Algorithm	Data Set	Partitioning	No. of Cluster	Cluster Distribution
K-Means	Air-Quality-Monitoring-Dataset-Pakistan	Absolute: 100 Relative[%]: 80	3	Cluster_0 =>6.16% Cluster_1 =>19.64% Cluster_2 =>74.2%
Fuzzy c-Means	Air-Quality-Monitoring-Dataset-Pakistan	Absolute: 100 Relative[%]: 80	3	Cluster_0 =>51.72% Cluster_1 =>41.96% Cluster_2 =>6.32%

7. CONCLUSION

Recently, data mining techniques have covered all areas of our lives. There are many data mining techniques which we are used in our daily life. We used these techniques in education, medicine, weather, business, banking etc. Before working on all these techniques it is important for us to understand all these data mining techniques. In this paper we discussed one of data mining technique is clustering. It is the type of data mining technique in which make the group or cluster of our data on the behalf of their similarity and difference. There are many algorithms which are used for clustering and many tool are also used for clustering. In this paper we used only two algorithms of clustering K Means and Fuzzy C means and we use KNIME tool for this. These two algorithms are very simple and give better result on every dataset. We use these two algorithms on Air-Quality-Monitoring-Dataset-Pakistan and get result. The k Means provide best result butt Fuzzy C means give us better result. Fuzzy C means divided our data set into three crustal with better average. So at the end we conclude that fuzzy c means is batter then k means for this type of data.

8. SUGGESTIONS AND FUTURE PLAN

In this paper we discussed some mostly used clustering algorithms like K means and Fuzzy C means. For future work we can also use some other clustering algorithms on this dataset. And also use some other algorithms and compare with each other and also change dataset and apply some other algorithms.

REFERENCES

[1] https://www.kaggle.com/mahmedphdcs17seecs/air-quality-monitoring-dataset-pakistan.

[2] Soumi Ghosh and Sanjay Kumar Dubey "Comparative Analysis of K-Means and Fuzzy CMeans Algorithms," ((IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 4, No.4, 2013

[3] V. S. Rao and Dr. S. Vidyavath.: "Comparative Investigations and Performance Analysis of FCM and MFPCM Algorithms on Iris data," Indian Journal of Computer Science and Engineering, vol.1, no.2, 2010 pp. 145-151

[4] Pallavi, Sunila Godara.: "A Comparative Performance Analysis of Clustering Algorithms" Pallavi, Sunila Godara / International Journal of Engineering Research and Applications (IJERA), Vol. 1, Issue 3, pp.441-445

[5] Aastha Joshi and Rajneet Kaur.: "A Review: Comparative Study of Various Clustering Techniques in Data Mining" IJARCSSE, 2013

[6] Soumi Ghosh and Sanjay Kumar Dubey.: "Comparative Analysis of K-Means and Fuzzy CMeans Algorithms," ((IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 4, No.4, 2013

[7] Prakash Singh1, Aarohi Surya2: "PERFORMANCE ANALYSIS OF CLUSTERING ALGORITHMS IN DATA MINING IN WEKA" International Journal of Advances in Engineering & Technology, Jan., 2015.

[8] Rekha Awasthi1, Anil K Tiwari2, Seema Pathak3.: "Analysis of Mass Based and Density Based Clustering Techniques on Numerical Datasets" Journal of Information Engineering and Applications, Vol.3, No.4, 2013

[9] Han J. and Kamber M.: "Analysis of Clustering Approaches for Data Mining In Large Data Source" International Journal on Recent and Innovation Trends in Computing and Communication, Volume: 2 Issue: 9

[10] https://towardsdatascience.com/understanding-k-means- clustering-in-machine-learning-6a6e67336aa1.

STUDY AND ANALYSIS OF COVID-19 OUTBREAKS IN PAKISTAN USING LINEAR AND RANDOM FOREST REGRESSION ALGORITHMS

Syeda Tahira Batool, Syed M. Mujtaba Hassan, Imdad Hussain. Dept. Of Computer Science, Riphah International University, Lahore campus, Pakistan.

Abstract. *COVID 19 turned out to be a pandemic threat around world. It is spreading rapidly worldwide. The current study interpret the condition of this epidemic disease in Pakistan and analyze confirm, death and recover cases emerging in Pakistan. The purpose of this study is to understand the nature of virus and its rate of disperse using machine learning algorithms in order to take precise decision and firm action.COVID-19 Pakistan datasets from January 1, 2020 to May 3, 2020 has been taken for analysis. Supervised machine learning algorithm i.e. linear regression and random forest regression used to study and analyze the COVID 19 disease in Pakistan.*

Keywords: *Visualization, analysis, COVID 19, linear regression, random forest, KNIME tool, Pakistan*

INTRODUCTION

On January 30, 2020 WHO reported COVID-19 outbreak a PHEIC abbreviated public health emergency of international concern. 77 countries have been declared corona cases from the date of March 14, 2020[1].The current study interpret the condition of this pandemic disease in Pakistan and analyze the results and visualize confirm, death and recover cases raised in Pakistan.

The motivation of writing this paper is to understand the nature of virus and its rate of disperse using machine learning algorithms in order to take precise decision and firm action. In this paper Linear regression and Random forest regression techniques used to study and analyses the COVID 19 outbreaks in Pakistan. Regression analysis investigate the association among target and independent variable. It is use to determine the strength of predicators, forecasting an effect and trend forecasting [2].

The aims of this paper are as follow

- To analyze the results of both algorithm
- Graphically Visualization of confirmed, recovered and death cases of coronavirus till proposed date in Pakistan.

The remaining portion of paper is formulated as. (Section 2) Literature review. (Section3) describes methodology. (Section 4) discusses experimental and analysis. (Section 5) describes results and discussion and then last (Section 6) describes the conclusions of the paper.

Literature review

Guptaa Rajan et al [3] conducted a study of COVID-19 in India from January 30, 2020 to March 30, 2020 using SEIR model and Regression model. Their paper based on the prediction of presumed cases for coming two weeks in India can rise between 5000-6000.Sathish Kumar et al [4] proposed a paper using latest UNSW-NB15 dataset for comparing the result of two models which is linear regression and random forest using WEKA tool. Their study purposed that random forest gave better result than linear regression. Rahmatollah Moradzadeh [5] proposed a paper based on COVID-19 situation in Iran. The aimed of this study is to highlight the challenges and readiness at the start of COVID-19 pandemic in Iran cities. Samrat Kumar Dey et al [6] proposed an article for analyze and visualize coronavirus case inward and outward of the china. They presented a comparative study of different cases reported worldwide.

Methodology

For study and analysis of COVID-19 cases in Pakistan, below diagram determined the workflow from initial to finish level using KNIME analytical tool.

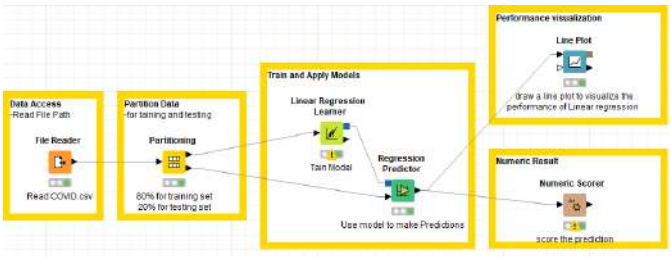


Figure 1. Work flow in KNIME

Data access

First step in workflow is to access the file path of COVID-19 dataset through file reader node for further implementation.

Data partition

Partitioning node is used to divide dataset into test and train data. Here 80% data is used for training and 20% data is used for testing.

Linear Regression technique

It is one of the easiest and powerful algorithm in machine learning. It is a statistical model that attempts to show the relation between two variables to the linear equation [4].Its accuracy and best fit is measured by loss, R squared, Adjust R squared etc.

Random Forest technique

It is a versatile supervised machine learning technique capable of performing both regression as well as classification. The main advantage of random forest algorithm is to deal with the missing values and prevent for

overfitting [8]. In the workflow Linear regression learner and Radom forest regression learner node used to train the data and regression predictor node used for prediction.

Analyze and Visualization

Numeric result have been calculated using scorer node as well as visualization done using different visualize node of KNIME tool.

Experiment and analysis

Purposed algorithms were performed in COVID 19 Pakistan dataset for choices concerned features so as to analyze the results. KNIME analytical tool version 4.1.2 used for Implementation. Computed system installation based on Windows 10, 64-bit operating system, CPU Intel ® Core i3 having 6 GB random access memory.

Dataset

COVID 19 datasets has been used for result analysis provided by John Hopkins University, USA. Dataset consists of Jan 22, 2020 to May 03, 2020 time Period. It carried confirm, death and recover cases from globally. In spite of, this work emphasized only on Pakistan's data of COVID-19 cases for study and analysis.

Line plot graph

This graph shown COVID-19 Pakistan's confirm, death and recover cases happened since March 22, 2020 to May 03, 2020.

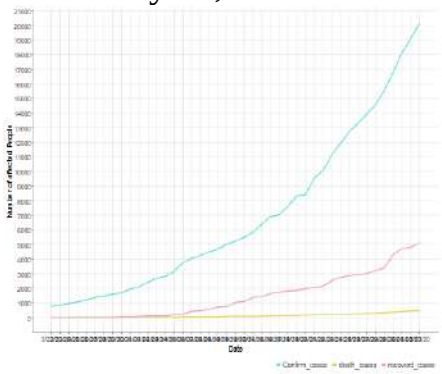


Figure 2.Line plot graph

Bar chart visualization

It displayed confirmed, death and recovered cases of viruses in Pakistan. Chart view as horizontal. After exploring the datasets, there were 20084 confirmed cases till May 03, 2020 in Pakistan. This show that cases increase rapidly since March 2020.

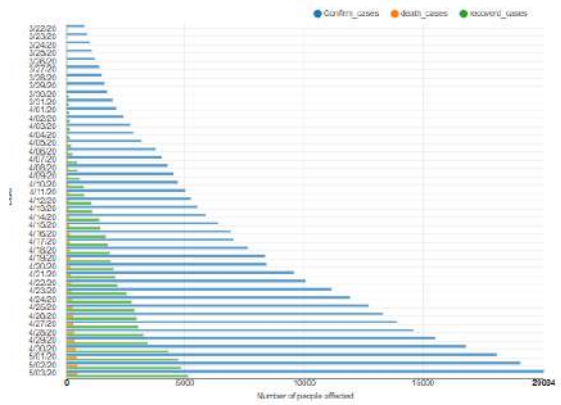


Figure 3. Bar chart view of cases in Pakistan

Scatter plot

In below presented graph, at x-axis date show till May 3, 2020 and at y-axis number of confirmed case.

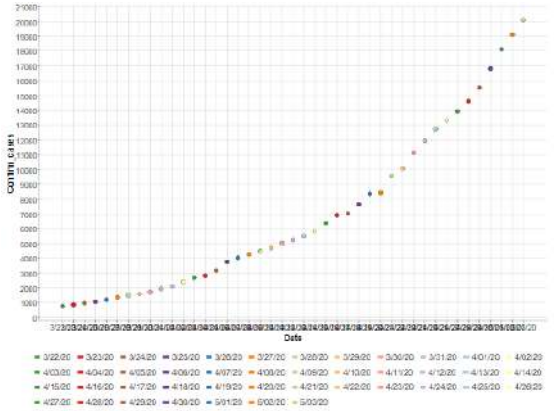


Figure 4.Scatter plot of confirmed cases in Pakistan

Below graph presented number of death cases where at x-axis date is show till May 3, 2020 and at y-axis number of death cases.

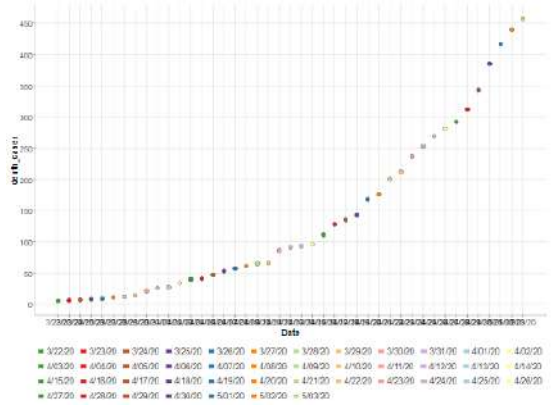


Figure 5.Scatter plot of death cases in Pakistan

Below graph presented number of recovered cases where at x-axis date is show till May 3, 2020 and at y-axis number of death cases.

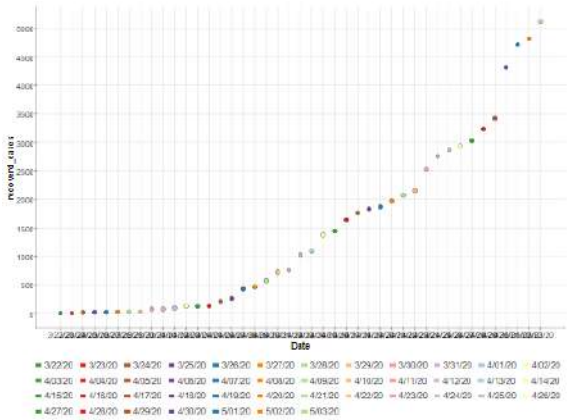


Figure 6.Scatter plot of recovered cases in Pakistan

Pia chart

Pie chart display the total numbering of confirm, death and recover cases reported till purposed date.

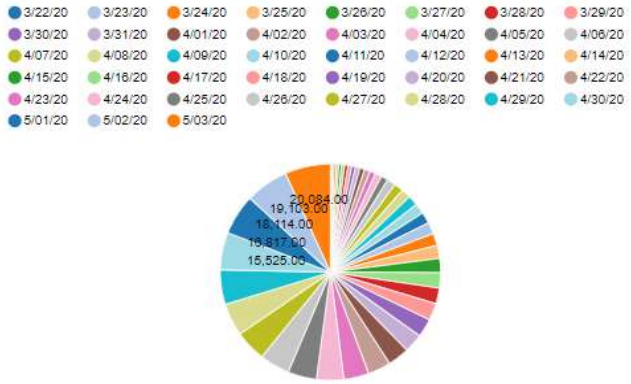


Figure 7.Pie chart of confirmed cases in Pakistan

Result and Discussion

The define algorithms Linear and Random Forest Regression were applied on COVID-19 Pakistan datasets. From dataset confirmed cases column taken as target variable whereas rest of taken as independent variable for analyze the proposed models. In order to get the best result KNIME analytical tool used to run algorithms. The below figure show the result of Linear Random regression

Table "Scores" - Rows: 6 | Spec - Column: 1 | Properties | Flow Variable

Row ID	D Prediction (Confirm_cases)
R^2	0.985
mean absolute error	429.893
mean squared error	257,851.562
root mean squared error	507.791
mean signed difference	-65.354

Figure 8.Linear regression statistics

The below figure show the result of random forest regression

Table "Scores" - Rows: 6 | Spec - Column: 1 | Properties | Flow Varia

Row ID	D Prediction (Confirm_cases)
R^2	0.998
mean absolute error	99.633
mean squared error	36,066.539
root mean squared error	189.912
mean signed difference	-82.574

Figure 9. Random forest regression statistics

By comparing the result of both algorithms we recommend that Random Forest regression carried out better results than Linear Regression.

Conclusion and future work

In this paper, linear regression and random forest regression supervised machine learning algorithms were used to study and analyze the effect of COVID-19 viruses in Pakistan. Graphically representation were shown to understand the dispersing rate of virus in order to take precise decision and firm action. Experiment results shows that random forest algorithm performed better then Linear Regression algorithm.

In future, different machine learning algorithm can be used to get results in order to understand the behavior of disease. This will help authorities to make better strategies to control the virus for spreading.

REFERENCES

[40] World Health organization (2020).Responding to community transmission of COVID-19.Retrieved from https://www.who.int/publications-detail/responding-to-community-spread-of-covid-19

[41] Halili, Festim & Rustemi, Avni. (2016). Predictive Modeling: Data Mining Regression Technique Applied in a Prototype. 5. 207-215.

[42] Rajan Gupta, Gaurav Pandey, Poonam Chaudhary, Saibal Kumar Pal ***doi:*** *https://doi.org/10.1101/2020.04.01.20049825*

[43] Kumar, P. & Raaza, A. (2018). Study and analysis of intrusion detection system using random forest and linear regression. Periodicals of Engineering and Natural Sciences. 6. 197-200. 10.21533/pen.v6i1.289.

[44] Moradzadeh, R. (2020). The challenges and considerations of community-based preparedness at the onset of COVID-19 outbreak in Iran, 2020. Epidemiology and Infection, 148, E82. doi:10.1017/S0950268820000783

[45] Dey, S. K., Rahman, M. M., Siddiqi, U. R., & Howlader, A. (2020). Analyzing the Epidemiological Outbreak of COVID-19: A Visual Exploratory Data Analysis (EDA) Approach. Journal of Medical Virology. doi:10.1002/jmv.25743

[46] Rajan Gupta, Saibal Kumar Pal, Gaurav Pandey ***doi:*** *https://doi.org/10.1101/2020.04.08.20058347*

[47] Niklas Donges (2019). A COMPLETE GUIDE TO THE RANDOM FOREST ALGORITHM. Retrieved May 04,2020,from https://builtin.com/data-science/random-forest-algorithm

ОБ ОДНОМ НОВОМ АЛГОРИТМЕ ВЫЧИСЛЕНИЯ РАССТОЯНИЙ ХЭММИНГА МЕЖДУ КОДОВЫМИ СЛОВАМИ

Абдурахманов Ж . К., Овхунов И.А., Андижанский Государственный Университет, Узбекистан.

Аннотация. *В данной работе предлагается новый алгоритм для вычисления расстояний Хэмминга между кодовыми словами двоичных кодов, контролирующих ошибки. Приведен сравнительный анализ этого алгоритма с традиционным алгоритмом рассматриваемой задачи. Алгоритмы описываются в виде функций на языке программирования C++.*

Аннотация. *Мазкур мақолада хатоларни назорат қилувчи иккилик кодларнинг код сўзлари орасидаги Хэмминг масофасини ҳисоблаш учун янги алгоритм таклиф қилинмоқда. Бу алгоритмнинг қаралаётган масаланинг ананавий алгоритми билан қиёсий таҳлили келтирилган. Алгоритмлар C++ дастурлаш тилининг функциялари шаклида берилган.*

Annotation. *In this paper, we propose a new algorithm for calculating Hamming distances between code words of binary codes that control errors. A comparative analysis of this algorithm with the traditional algorithm of the problem under consideration is given. Algorithms are described as functions in the C++ programming language.*

Ключевые слова: *алгоритм, расстояние Хэмминга, двоичные векторы, функция-подпрограмма, конечное множество, дискретное множество, подмножество.*

Калит сўзлар: *алгоритм, Хэмминг масофаси, иккилик векторлар, функция-қисмдастур, чекли тўплам, дискрет тўплам, қисмтўплам.*

Keywords: *algorithm, Hamming distance, binary vectors, subroutine function, finite set, discrete set, subset.*

Как известно [1], расстояние Хэмминга определяется как расстояние между двоичными векторами длины ***n***, или, говоря математическим языком – как особо определенное расстояние между двумя точками ***n***-мерного пространства, координаты которых состоят только из ***0*** и ***1***. Если

$\bar{a} = (a_1, a_2, \dots, a_n)$ и $\bar{b} = (b_1, b_2, \dots, b_n)$

– двоичные векторы длины *n*, то расстояние Хэмминга $h(\bar{a}, \bar{b})$ между ними можно вычислить по следующей формуле:

$$h(\bar{a}, \bar{b}) = \sum_{i=1}^{n} |a_i - b_i|$$

На языке программирования *C++* для вычисления расстояния Хэмминга можно воспользоваться, например, следующей функцией:

```
// Функция вычисляет расстояние Хэмминга по определению
int  h(int a[], int b[], int n) {
```

```
    int k, m = 0;
    for (k = 0; k < n; k++)  // тело этого цикла выполняется  n  раз
    if (a[i] – b[i])
        m++;
    return (m);
}
```

В этой функции, каковы бы ни были массивы, представляющие двоичные векторы, проверка внутри цикла осуществляется ровно ***n*** раз. Другими словами, число проверок равно числу итераций цикла ***for***.

Например, для кодовых слов *(0, 0, 0, 1, 1, 0, 1)* и *(0, 0, 0, 0, 0, 1, 0)* длины *7*, проверка происходит ровно *7* раз, хотя эти векторы в первых трех позициях совпадают. Можно ли исключить проверку в таких совпадающих позициях с нулевыми значениями?

Рассмотрим множество *M* всех подмножеств множества первых *n* натуральных чисел *{1, 2, 3, …, n}*. Каждому двоичному вектору длины *n* поставим в соответствии из этого множества подмножество, состоящее из номеров ненулевых координат (позиций) данного двоичного вектора. Очевидно, этим устанавливается взаимно-однозначное соответствие между множеством *M* и множеством всех двоичных векторов длины *n*. Пусть теперь двоичным векторам

$$\bar{a} = (a_1, a_2, \ldots, a_n) \text{ и } \bar{b} = (b_1, b_2, \ldots, b_n)$$

соответствуют подмножества *A* и *B* множества *M* соответственно. Обозначим через *|A|*, *|B|*, *|A∩B|* число элементов в подмножествах *A*, *B*, *A∩B* соответственно (т.е. мощность этих подмножеств), где *A∩B* – пересечение подмножеств *A* и *B*. Тогда имеет место следующая теорема, которая по существу предлагает иной способ (алгоритм) вычисления расстояния Хэмминга.

Теорема. $h(\bar{a}, \bar{b}) = |A| + |B| - 2|A \cap B|$.

Если *|A| = k*, *|B| = m* и *|A∩B| = p*, то следующая функция, написанная на языке программирования *C++*, вычисляет расстояние Хэмминга по этой новой формуле. В этой функции подмножества *A* и *B* реализованы в виде одномерных массивов с числом элементов *k* и *m* соответственно. Массивы предполагаются упорядоченными (по возрастанию или по убыванию – не имеет значения).

```
//Функция вычисляет расстояние Хэмминга по новому методу
int  H( int  A[],  int  B[],  int  k,  int  m ) {
    int  i = 0,  j = 0,  s = 0;
    do {       // тело этого цикла выполняется меньше чем  k + m - p раз
        if (A[i] == B[j]) { s++;  i++;  j++;  continue; }
        if(A[i] < B[j])  i++; // использование …
        else  j++;              // … упорядоченности массивов
    } while((i < m) && (j < n));
    return (k + m – 2*s);
}//int H(…)
```

По смыслу наших обозначений, если в исходных двоичных векторах нет совпадающих нулевых позиций, то всегда выполняется равенство $k + m = n$. Если же есть в исходных двоичных векторах совпадающие позиции, то выполняется строгое неравенство $k + m < n$. Поскольку тело цикла выполняется меньше чем $k + m - p$ раз, то получается, что в этой функции число повторений всегда меньше числа n.

Проверка для различных двоичных векторов этой функции в среде программирования *Borland C++ Builder* показала, число итераций гораздо меньше числа n. Случай, когда $k + m = n$ кажется на первый взгляд самым наихудшим, но в этом случае число итераций является наименьшим и равно наименьшему из чисел k и m, т.е. равно $min(k, m)$.

Разумеется, платой за это улучшение алгоритма вычисления расстояний Хэмминга является упорядоченность массивов A и B.

При составлении вышеприведенных функций (подпрограмм) мы руководствовались стандартом языка программирования C++, описанном например, в книгах [2] и [3].

Литература

1. *Р. В. Хэмминг. Теория кодирования и теория информации: Пер. с англ. – Москва, “Радио и связь”, 1983. – 176 с., ил.*
2. *James P. Cohoon, Jack W. Davidson. C++ Program Design. An Introduction to Programming and Object-Oriented Design. Irwin/McGraw-Hill, 1997, 754 pages.*
3. *Б. Страуструп. Язык программирования C++, 3-е изд./Пер. с англ. – СПб.; М: «Невский диалект» - «Издательство БИНОМ», 2000 г. – 991 с., ил.*

АҚЛЛИ КУТУБХОНА МАЖМУАСИДА ЎЗАРО БОҒЛИҚ МАЪЛУМОТЛАРНИ ТАРҚОҚ ХОЛДА БУЛУТЛИ АХБОРОТ ТИЗИМЛАРИДА САҚЛАШ

Турсунов Н.Х., Мухаммад ал-Хоразмий номли Тошкент ахборот технологиялари университети қошидаги ахборот-коммуникацион технологиялари илмий-инновацион маркази. Кичик илмий ходим,

Аннотация. *Ушбу мақолада ўзаро боғлиқ маълумотларни интеграллашган холда семантик фазода тасвирлаш масаласи қаралади. Боғланган маълумотлар турлари ва уларни қидириш сўровларини қайта ишлаш механизмлари таҳлил қилиниб, бу муаммоларни хал қилиш услубларининг архитектуравий ечимлари солиштирилади. Архитектуравий ечимларни солиштириш натижасида ҳар бир ечимнинг камчиликлари кўрсатилади. Маълумотларни қайта ишлаш маркази ресурсларини янада самарали ишлатиш мақсадида булутли тизимларга асосланган ечим таклиф қилинади.*

Калит сўзлар: *ақлли кутубхона, узаро боғлиқ маълумотлар, булутли ахборот тизими, маълумотлар интеграцияси, интернет, веб стандартлари, файл-сервер тизими, маълумотларни қайта ишлаш маркази.*

Аннотация. *В данной статье рассматривается проблема представления интегрированных связанных данных в семантическом пространстве. Анализируется виды связанных данных и механизмы обработки поисковых запросов этих данных. Сравниваются архитектурные решения реализации поисковых запросов. Путем сравнения архитектурных решений показывается недостаток каждого предложенного решения. С целью эффективного использования ресурсов центра обработки данных предлагается решение основанную на облачных систем.*

Ключевые слова: *умная библиотека, взаимосвязанные данные, облачная информационная система, интеграция данных, интернет, стандарты веба, система файл-сервер, центр обработки данных.*

Annotation. *This article discusses the problem of representation of integrated linked data in semantic space. The types of related data and mechanisms for processing search queries for this data are analyzed. Architectural solutions for implementing search queries are compared. By comparing architectural solutions, the drawback of each proposed solution is shown. For the purpose of efficient use of data center resources, a solution based on cloud systems is proposed.*

Keywords: *smart library, interconnected data, cloud information system, data integration, Internet, web standards, file server system, data center.*

Интеграллашган электрон кутубхона турли манбаларда жойлашган маълумотларни фойдаланувчига унификация қилинган ягона шаклда тақдим этиш жараёнини бажаради. Масалан, исломшуносликга оид интеграллашган

электрон кутубхона Ўзбекистон мусулмонлари идораси, ислом цивилизацияси маркази, шарқшунослик институти, Ўзбекистон халқаро ислом академияси каби муассасалар кутубхоналари манбаларида жойлашган маълумотларни бирлаштириб, маълумот қидирувчига ягона шаклда тақдим этиши мумкин. Интеграллашган электрон кутубхоналар яратишда кўп холларда ўзаро боғлиқ маълумотлар билан ишлашга тўғри келади. Боғланган маълумотлар (Linked Data) – информатикада ўзаро интернет орқали боғланган маълумотлар мажмуалари коллекциясидир[1]. Ушбу атама ўзаро боғлиқ тузилмалаштирилган маълумотларни интернетда ёритиш усулларини ҳам кузда тутади. Усуллар Вебнинг стандартлари ҳисобланмиш HTTP, RDF ва URI кабиларга асосланади ва ахборотларни компьютер ёрдамида ўқишга қулай шаклда тарқатиш имконини беради. RDF маълумотлар орасидаги муносабатларни баён қилишнинг абстракт модели хисобланади[2]. RDF триплет вебда XML конструкция орқали тасвирланади. Тим Бернерс-Ли интернетни такомиллаштириш йўналишида гиперматндан автоматлаштирилган холда маълумот чиқариб олиш, уни таҳлил қилиш, янги ахборотни генерация қилиш, маълумотлар боғлиқлигини ўрнатиш таклифини киритди. Бундай ёндошувни у семантик тур деб атади. Бундан ташқари семантик турни амалиётга жорий қилишни “боғланган маълумотлар” атамаси билан атади ва бу жараённинг асосий тамойилларини ишлаб чиқди: 1. Моҳиятлар номи сифатида универсал идентификаторни қўллаш (URI – Uniform Resurs Identifiers). 2. Ном бўйича мурожат қилиш имкониятини жорий қилишда HTTP URI ни қўллаш. Бу маълумотлар инсонлар томонидан ёки дастурий воситалар томонидан топилиши керак. 3. Стандарт форматлардан фойдаланган холда URI бўйича мурожат қилинганда моҳият ҳақидаги маълумотларни тасвирлаш. 4. Қидирувни енгиллаштириш мақсадида бошқа боғланган URI ларга ҳаволалар орқали мурожат қилиш. Бу тамойилларни жорий қилиш маълумотларни тасвирлашнинг RDF конструкцияси орқали амалга оширилиши мумкин бўлди. Шундай қилиб, интеграллашган маълумотларни тасвирлашнинг URI-RDF комбинацияси юзага келди. Бу имконият турли манбалардаги маълумотлар билан ишлашга йўл очади. Маълумотлар турли белгилар орқали боғланган бўлиши мумкин. Масалан, муаллиф номи, асар номи, тадқиқот йўналиши, нашр йили ва бошқа белгилар. RDF модель визуал кўринишда граф шаклида акс эттирилади. Боғланган маълумотлар хажми кенгайган сари граф улчами кенгайиб, ишлаш учун ноқулайлик тўғдиради.

Боғланган маълумотларни интеграциялаш жараёнида боғловчи восита сифатида Z39.50 протоколи ишлатилади[3]. Ушбу протокол турли автоматлаштирилган кутубхона ахборот тизимларида (АКАТ) маълумот қидиришни унификациялаш имконини берди, бундан ташқари ягона синтаксис тизимини ишлатган холда турли АКАТ лардаги библиографик ёзувларни қидириш мумкин бўлди. Z39.50 протоколи ёрдамида боғланган маълумотларни интеграциялаш жараёнида қўйидаги муаммоларга дуч келинди: 1. Турли тизимлар фойдаланувчиларга турли қидирув

интерфейсларини таклиф этади. 2. Компьютерда фойдаланишга мулжалланган мета маълумотларни ҳамкорликда ишлатиш имкониятининг йўқлиги.

Ушбу муаммоларни хал қилиш икки хил ёндошувни жорий қилиш орқали амалга оширилди: 1. Z39.50 протоколини ишлатган холда турли хил кутубхоналардан маълумотларни биргаликда қидириш. 2. Турли манбалардан маълумотларни марказий серверда аккумуляция қилиш.

Ҳар бир ёндошувнинг ўзига хос камчиликлар мавжуд. Биринчи усулда, яъни барча кутубхоналардан бир варакайига қидиришда манбалар сони 100 дан ошса, тизим ишламай қолади. Иккинчи ёндошувни жорий қилиш учун янги OAI-PMH протоколи яратилди[4]. Бу протокол турли АКАТ лардаги мета маълумотларни автоматик тарзда йиғади ва йиғилган маълумотларни марказий серверда аккумуляция қилади. Шу пайтгача дунё миқиёсида ушбу услуб ишлатилиб келинмокда. OAI-PMH протоколини ишлатишнинг файл-серверли схемаси ахборот технологияларининг замонавий босқичида яхши самара бермайди. Чунки файл-серверли тизимда масштаблаштириш ва виртуаллаштириш механизмлари бўлмагани учун маълумотларни қайта ишлаш маркази ресурсларини оптимал ишлатиш имконияти пасайиб кетиши мумкин.

Ушбу марузада ақлли кутубхона мажмуасидаги узаро боғлиқ маълумотларни тарқоқ холда булутли ахборот тизимларида сақлаш услуби таклиф қилинади[5]. Бунда маълумотларни қайта ишлаш маркази физик хусусиятларидан келиб чиқиб, бир неча виртуал тизимларнинг образлари яратилади. Ҳар бир виртуал тизим динамик масштаблаштириш имкониятига эга. Динамик масштаблаштириш виртуал тизим ресурсларини HTTP-сўровлар талабларига мослаштириб беради ва маълумотларни қайта ишлаш марказининг физик ресурсларини самарали ишлатиш имконини яратади. Таклиф қилинаётган услуб хозирги кунда энг инновацион ҳисобланадиган булутли тизимлар техналогиясини электрон ақлли кутубхоналар яратиш жараёнига тадбиқ қилиш имконини беради.

Адабиётлар

1. Tom Heath, Cristian Bizy. Linked Data: Evolving the Web Info. Global Data Space – Morgan & Gloypool Publishing. 2011.

2. Жлобинская, О. Н. Как представлять библиографические данные в RDF: докл. на конф. «ЛИБНЕТ-2013», г. Звенигород, 18 – 22 нояб. 2013 г. [Электронный ресурс] / О. Н. Жлобинская // Российский коммуникативный формат : офиц. сайт. — Электрон. дан. — [Москва, 2015]. — 11 с. — Режим доступа: http://www.rusmarc.ru/publish/bibdateRDF.pdf.

3. Сысойкина, М. А. Протокол Z39.50: обоснование необходимости использования [Электронный ресурс] / М. А. Сысойкина // Информационные технологии в книжном бизнесе. — Электрон. дан. — Режим доступа: http://www.bookresearch.ru/Z3950usage.htm.

Davison, S. Enhancing an OAI-PMH service using Linked Data: the case of the Sheet Music Consortium [Electronic resourse] / Stephen Davison // SWIB'13:

Semantic Web in libraries, Nov. 25 – 27, 2013, Hamburg, Germany. — Electronic data. — 27 Nov. 2013. — Abstr., slides [29 p], video. — Mode of access: http://swib.org/swib13/programme.php

DATA SCIENCE CLASSIFICATION COMPARISON: RAPID MINER, WEKA AND ORANGE

Zark Ali Jan. Riphah International University, Lahore Campus, Pakistan.

Abstract. *In this paper compare different type of classification algorithm working with several completely free available data mining and knowledge discovery programs like WEKA, Rapid miner, and also Orange. The Precision of classification algorithm such as SVM, Random Forest and also Naïve Bayes algorithm are contrasted utilizing each of 3 components. Ebola_2014_2016 [1] data-set is employed for analyzing the Classification algorithm so as to rectify the individuals without the Ebola disease.*

Keyword*: Data Science, Classification, Machine Learning, Accuracy, Naïve Bayes, SVM, WEKA, Orange*

1. INTRODUCTION

Due to the fact the info remains extremely climbing, it will become problematic for a person to manually assess the exact info to tactical decision-making. Ergo humans need to have assistance of datamining to mine intriguing information out of the data that is available. Datamining is the procedure of finding interesting information from considerable quantities of information saved in data bases, data warehouses, or any different data repositories. Some of many significant issues in data mining would be that the Classification that will involve finding regulations which partition specified info into predefined categories. From the datamining realm at which trillions of info can be properly used, the implementation timing of present calculations could eventually become time-consuming. Thus, there's a demand for automatic applications which may assist individuals in altering those big data right into Info. Now we have lots of free basis data-mining applications and tool are all offered for usage such as for instance the RapidMiner [2], Waikato Environment Knowledge Analysis (WEKA) [3], Orange [4], NLTK etc. All these programs and applications furnish a variety of algorithms and methods which support in improved diagnosis of information. These programs assist in audience investigation, information conception, reversion analysis, Conclusion bushes, Prognostic analytics, Text mining.

We ran a comparison analysis amongst classification algorithm including as SVM, Random Forest and also Naïve Bayes algorithm with WEKA, RapidMiner and also Orange instrument. The precision step; that reflects the proportion of correctly labeled examples, can be useful for estimating the operation of the classification algorithm.

2. METHODOLOGY OF LITERATURE REVIEW

In Logical Archives such as (ScienceDirect, SpringerLink, PubMed, and Google Scholar) were searched between the period April 30, 2020, to May 3, 2020. Find a Suitable title and keywords and used. After a search in detail, we have a

number of articles found multiple keywords were selected, based on their significance to target studies in 'Ebola Classified,' Science Science on Ebola. A mixture of keywords, thrown in for a detailed search for each specific data. Scientific articles on practical subjects, only for humans (in human subjects), and published in English, were considered for a systematic review.

3. CLASSIFICATION ALGORITHMS

Info Classification Algorithm can be really a method of choosing the hypothesis by a pair of choices which best suits a pair of observations. Info Classification Procedure comprises two measures:

i) Assembling the Classifier product: This the classifier is assembled by finding out exactly the practice group along with also their associated class tags. ii) Utilizing Classifier for Classification: During this measure, the classifier can be useful for classification. This evaluation data can be utilized to gauge the validity of classification regulations.

We've researched the next Classification Algorithm within our paper:

- SVM
- Naïve Bayes
- Random Forest

SVM: A support vector machine (SVM) can be a machine-learning type that works by using segmentation algorithms to find two-party split news. After receiving a set of SVM version of marketing information for every single division, they are in the process of categorizing the new text. So, you're working on the issue of text separation.

Random Forest: Random forests or designated forests of debate are a learning process for the classification, processing and other activities that work by creating a large number of finishing projects during a work and class outreach that is a form of classes or individual prediction of trees.

Naive Bayes: is a household of uncomplicated probabilistic classifiers in line with employing Bayes' theorem with solid independence assumptions amongst your qualities.

4. TOOLS GUIDE

4.1 Rapid Miner

It has always been the study of machine learning and data analysis processes. Its open-source, free trade is fully implemented in Java. It represents a high-level way of looking at the most complex of problems - modular theory that allows the style of the red chains to detect large parts of learning problems. RM uses XML to spell the trees for the user in charge of technology acquisition (KD) processes. RM incorporates flexible operators to receive installation and installation information in many ways. It contains 100 great learning strategies such as differentiation, tolerance and bonding responsibilities.

4.2 WEKA:

Has been a reliable tool for machine learning and data analysis previously developed at the University of Waikato in New Zealand. It contains a large selection of advanced machine learning and data computation published in Java. WEKA combines applications for classification, classification, rules of association and mergers, to see WEKA popularized by industry and academic research staff, and widely used in training activities.

4.3 Orange:

It has always been a learning process for learning and using databases (written with Python). It is a transparent system for applications to get visual data and analysis, and it can likewise serve as a Python library. The application is archived and processed through the Bioinformatics Laboratory of the School of Computer and Information Science at the University of Ljubljana. Orange can be a very efficient software system for data mining and machine learning and statistics. The elements can be accessed by widgets and all include direct data visualization, stand-alone selection and preparation, robust investigations of understanding calculations and forecasts. After searching the data selection and the external read-out process to be deleted from the Data and Structure after the entire page has been read and you get to Explore the phylogenetic tree.

5. EXPERIMENT

5.1 Dataset.

We have downloaded ebola_2014_2016 from the Kaggle [4]. We have total numbers of a row are 2485. In this data collection, we have No Certified Death Records and Certified Death Records, which may be suspected. The data was collected from Kaggle. Certified death records or No is a class label used to divide into groups (patient death of Ebola or Recovery). This data collection has 10 country Ebola records and 2014 to 2016 Ebola patient records. The dataset has the following attributes:

- Country
- Date
- No. of suspected cases
- No. of probable cases
- No. of confirmed cases
- No. of confirmed, probable and suspected cases
- No. of suspected deaths

5.2 Classification Algorithm using Rapid miner:

In Rapid miner we get highest accuracy in Random forest and lowest accuracy in SVM as shown in fig 1. Its a crucial point from the RapidMiner we get different accuracy from the different classification algorithms like Naïve Bayes, SVM and randomforest. These algorithms are applied on the same dataset (ebola_2014_2016).

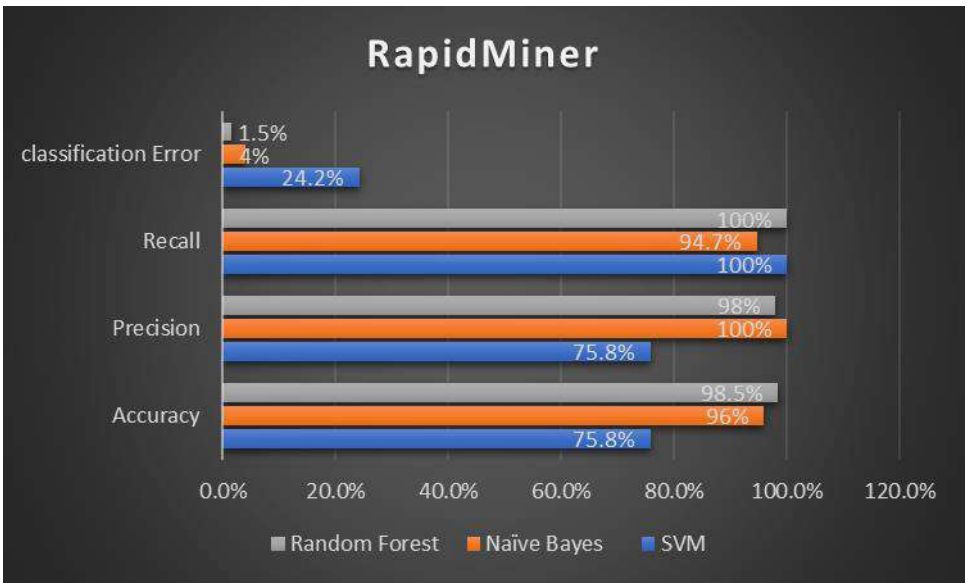


Figure 5:RapidMiner Result

5.2 Classification Algorithm using Orange:

In Orange we get highest accuracy in Random forest 92.8% and lowest accuracy in SVM 82.2% as shown in fig 2. It's a vital fact from the Orange we get different accuracy from the different classification algorithms like Naïve Bayes, SVM and randomforest. These algorithms are applied on the same dataset (ebola_2014_2016). Same as theses algorithms are applied through the tool weka for the to measure the accuracy of this dataset from the different machine learning tools.

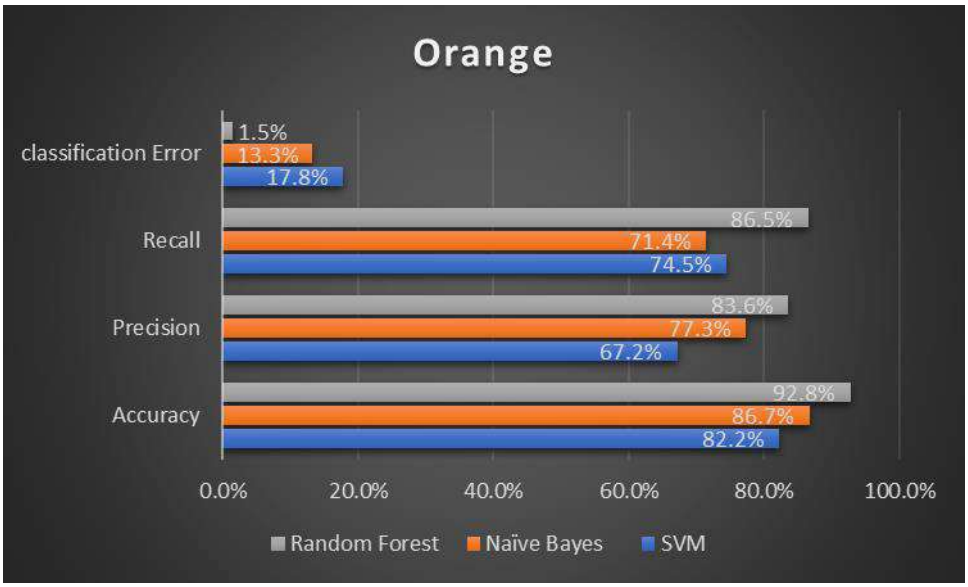


Figure 6:Orange Result

5.4 Classification Algorithm using Weka:

Weka is the most common and the crucial tool in the machine learning but there are some drawbacks in weka like weka not support the big dataset files that why in Weka we only apply Naïve Bayes algorithm, other algorithms are not supported with this dataset.

Algorithm	Accuracy	Precision	Recall	classification Error
Naïve Bayes	81.5%	79.3%	71.4%	18.5%

Table 1: Accuracy measure using weka

6. CONCLUSIONS:

In this paper Ebola, patient data collection used and get data from the ebola_2014_2016 Kaggle Data Set. It has 2485 rows with 10 attributes and 1 class variable. There was a concert of accuracy classified models result attached, which is as clear as the tracks in Figure 3.

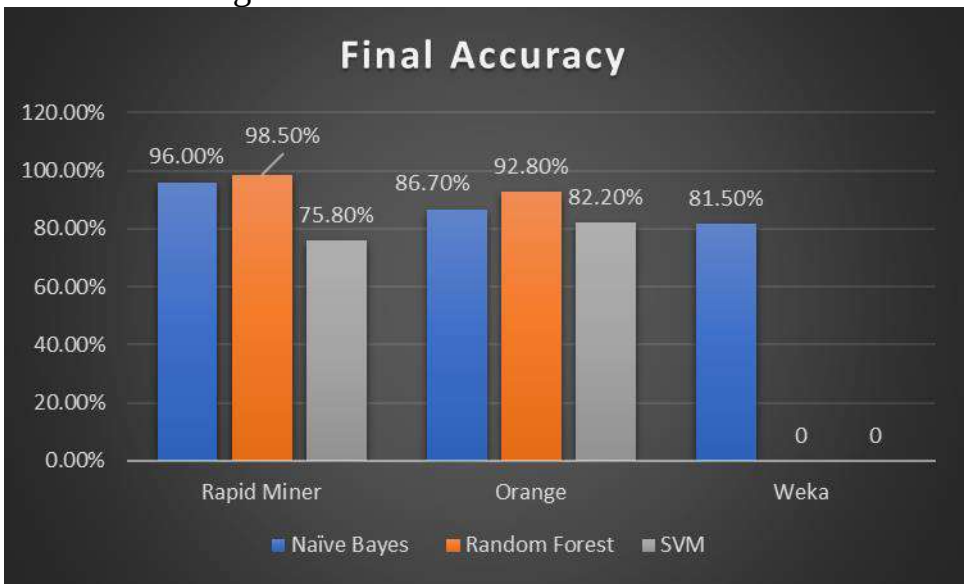


Figure 7:Final Accuracy

From the above in Fig it can be deduced that the WEKA tool measures the low accuracy of the Naive Bayes, but with the same process in the Rapid Miner application system it measures better accuracy when associated with WEKA. When calculating SVM the Orange tool showed lower accuracy and the Rapid Miner application showed better accuracy compared to the previous one. Ultimately usually the Rapid Miner application predicts advanced accuracy forms for the whole three partitioning process. Above Fig [3] also proves that the accuracy of Rapid Miner and Naïve Bayes is well correlated with SVM.

REFERENCES

[1] Devakumar Kp , "Ebola 2014-2016 Outbreak" <https://www.kaggle.com/imdevskp/ebola-outbreak-20142016-complete-dataset>

[2] <https://rapidminer.com/get-started/>

[3] <https://sourceforge.net/projects/weka/>

[4] <https://orange.biolab.si/>

[5] <https://en.wikipedia.org/wiki/Support-vector_machine>

[6]< https://en.wikipedia.org/wiki/Support-vector_machine>

[7] <https://en.wikipedia.org/wiki/Naive_Bayes_classifier>

[8] YongzhenGeaShuoZhaobXiqiangZhaoa, "A step-by-step classification algorithm of protein secondary structures based on double-layer SVM model" Sciencedirect 2020

[9] Amrita Naik,Lilavati Samant, "Correlation review of classification algorithm using data mining tool: WEKA , Rapidminer , Tanagra ,Orange and Knime" ScienceDirect 2016

PREDICTION OF HEPATITIS VIA MACHINE LEARNING

Zark Ali Jan, Arslan Raza. Riphah International University, Lahore Campus, Pakistan.

Abstract. Depending on their properties Data can be categorized. We are separated from existing or past records (data). One of the main purposes of classification is that we can obtain good results and increase the reliability of the data obtained. In the field of health care, the clinical diagnosis is usually based on medical information. And computer-based decision-making plays a major role in the field of medicine. Growing research into the prediction of hepatitis is important to differentiate the findings of the study and to provide the reader with information on the possible strategies for Hepatitis prediction at each stage. Medical data prediction, integration, and classification of the best algorithms can be used. From research, it was observed that hybrid algorithms improve the accuracy of Hepatitis prognostic algorithms. In this paper, we summarize the most common techniques used to predict Hepatitis and its severity.

Keyword: Data science, Machine Learning, Accuracy Prediction, Classification, Naïve bays, Decision Tree, Random Forest

1. INTRODUCTION

Hepatitis, an inflammation of the liver from a variety of causes, of infection and infertility. Infectious agents that cause hepatitis include viruses and viruses. Non-infectious causes include specific drugs and toxins. In some cases hepatitis is caused by an autoimmune reaction directed against the body cells. Early detection of cardiac ischemia is important to help reduce the number of deaths. Hepatitis is the leading cause of death some years ago. There are many computer software algorithms and asset strategies that doctors predict for Hepatitis in the early stages. A computer-based prediction program can help medical professionals predict Hepatitis based on past clinical data of patients. Patient details may include unwanted and subtle symptoms, especially when patients have more than one type of illness from the same stage of the disease. Doctors may not be able to diagnose you properly.

The pharmaceutical industry collects huge amounts of medical data and we use this data and use different algorithms for predictions. Depending on the general properties, the data can be divided into different sets. Method of classification of sub-cluster analysis. The classification method used to construct the current model that can record records using sample data. Researchers introduced many types of algorithms in stages over the years. but these are most commonly used in stages by algorithms, AC2, New ID CAL5, C4.5, CN2, Bayes theorem, clustering algorithm, close neighbor, radial base function, Naive Bayes, Kohonen Polytrees, net, LVQ Preparation, Hard K methods and C-Means but we use some of these to find the category and predict Hepatitis. Other widely used techniques for recording and

classifying algorithms. The reason is to use this standard algorithm to introduce similar properties and learn from previous data and then use experimental data and train the data and predict the results.

2.LITERATURE REVIEW

In Logical Records such as (ScienceDirect, SpringerLink, PubMed, and Google Scholar) were searched between the period April 30, 2020, to May 3, 2020 and find a Appropriate title and keywords. Classification is one of the most important tools for decision making in the field of medical science. Much research has been done on this subject. If you refer to the related literature it can be seen that more research was done on Hepatitis which is more complete among them (Hashem, JUNE 2018) most ware algorithms used in their studies by their accuracy rate was significantly better than 89% of mice with accuracy. The set of training and test setup size was 155 respectively, in their research, the second best article I've read (Hoang, 2015) this is some of the most interesting research I've ever read, better accuracy and predictability and one of the best things that in this article is your use of simple words so I easily understand everything and the third article I read for my research paper (Emamjomeh, 2013) this paper is not so small and technical but I read many things from this page your use of more than 13 explains all algorithms precisely.

3.METHODOLOGY

3.1. Quantitative Research

Quantitative research is seen as a systematic examination of socially undesirable data and performing measurable, scientific, or engaging processes. Quantitative research collects data from existing and potential customers using surveillance techniques and transmitting web-based visualizations, online surveys, and more.

4. TOOLS AND TECHNIQUES

4.1. Classification Algorithm

In machine learning, classification is a supervised learning method where it learns a computer program from input data and uses this command to distinguish new observations This data collection can be of two parts (such as identification, illness, or male or female, whether or not spam is being sent). Other practical examples of data segmentation problems are speech recognition, manuscript recognition, cardiovascular diagnostics, document classification, etc.

4.2. Naive Bayes Classifier

It is a classification method based on Bayes' Hypothesis on the assumption of autonomy among researchers. In other words, the organizers of the Naive Bayes group think that the occurrence of a particular feature in the classroom has nothing to do with the fact that something else or that all of these components have independent roles.

4.3. Decision Trees

The decision tree establishes a model of classification or classification as a tree structure. It breaks data set to smaller and smaller, while simultaneously improving the decision tree. The result is a tree with Decision Nodes and Leaf Nodes. A decision node has two or more branches and a leaf node represents a category or decision.

4.4. Random Forest

Random forests are the study of Classification, register, and other functions that work to create a large number of decision trees during training and the class extraction which is a class method (classification) or prediction that means predicting (rearranging) specific trees. Random forests prepare decision trees the habit of qualifying for their training.

4.5. WEKA

Weka is attempted and tried open-source AI programming that can be gotten to through a graphical UI, standard terminal applications, or a Java API. It is generally utilized for educating, explore, and mechanical applications, contains a plenty of implicit apparatuses for standard AI undertakings, and also gives straightforward access to notable tool compartments, for example, scikit-learn, R, and Deeplearning4j.

4.6. Study Environment and Background Theories

Clustering and classification were created in this investigation. The techniques that were utilized will be clarified quickly in this segment. Now we start apply different algorithms on our data set of Hepatitis first of all we start with classification

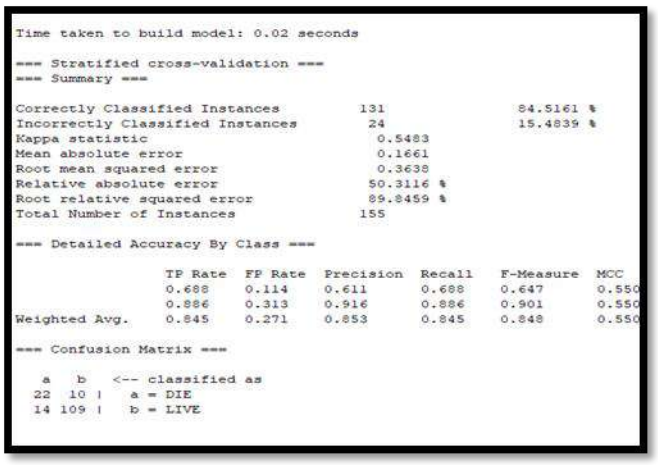
5. EXPERIMENT

5.1. Classification

In classification we need our data set with different data using this classification data we add our database to our WEKA tool and apply class-level algorithms to AI and classification, partitioning the learning process managed when the PC system receives the data and then uses this to find a way to combine ideas.

5.2. Naive Bayes Classifier

We apply naive Bayes classifier on our data we use 66% training data from data set and 44% testing data It is a technique of classification dependent on Bayes' classifier with the presumption of self-governance between the indicators. it’s near about 89% accurate we use 151 instances



```
Time taken to build model: 0.02 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         131               84.5161 %
Incorrectly Classified Instances        24               15.4839 %
Kappa statistic                          0.5483
Mean absolute error                      0.1661
Root mean squared error                  0.3638
Relative absolute error                 50.3116 %
Root relative squared error             89.8459 %
Total Number of Instances              155

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC
                 0.688    0.114    0.611      0.688    0.647      0.550
                 0.886    0.313    0.916      0.886    0.901      0.550
Weighted Avg.    0.845    0.271    0.853      0.845    0.848      0.550

=== Confusion Matrix ===

   a   b   <-- classified as
  22  10 |   a = DIE
  14 109 |   b = LIVE
```

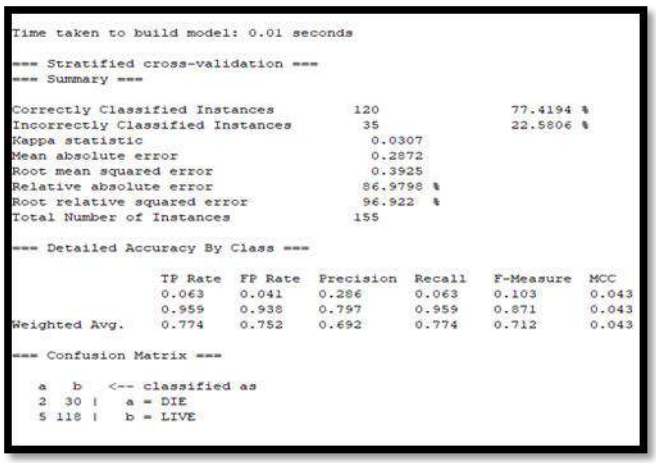
Summary	Weka Result
Correctly Classified Instances	131
Incorrectly Classified Instances	24
Kappa statistic	0.5483
Mean absolute error	0.1661
Root mean squared error	0.3638
Relative absolute error	50.3116 %
Root relative squared error	89.8459 %
Total Number of Instances	72.9665 %

Table: 8

a	b	classified as
22	10	a=DIE
14	109	b=LIVE

5.3. Decision Trees

We enter the classification decisions of our data using 66% training data from the data set and 44% of the test data The decision tree makes a classification or classification model as a tree structure. It divides the information index into smaller and smaller pieces, but at the same time creates a related judgment tree. The final product is a tree containing the Verdict Hub and the Leaf Hub. it's near about 77% accurate we use 155 instances.



```
Time taken to build model: 0.01 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         120               77.4194 %
Incorrectly Classified Instances        35               22.5806 %
Kappa statistic                          0.0307
Mean absolute error                      0.2872
Root mean squared error                  0.3925
Relative absolute error                 86.9798 %
Root relative squared error             96.922  %
Total Number of Instances              155

=== Detailed Accuracy By Class ===

                 TP Rate   FP Rate   Precision   Recall  F-Measure   MCC
                 0.063     0.041     0.286     0.063     0.103     0.043
                 0.959     0.938     0.797     0.959     0.871     0.043
Weighted Avg.    0.774     0.752     0.692     0.774     0.712     0.043

=== Confusion Matrix ===

   a   b   <-- classified as
   2  30 |   a = DIE
   5 118 |   b = LIVE
```

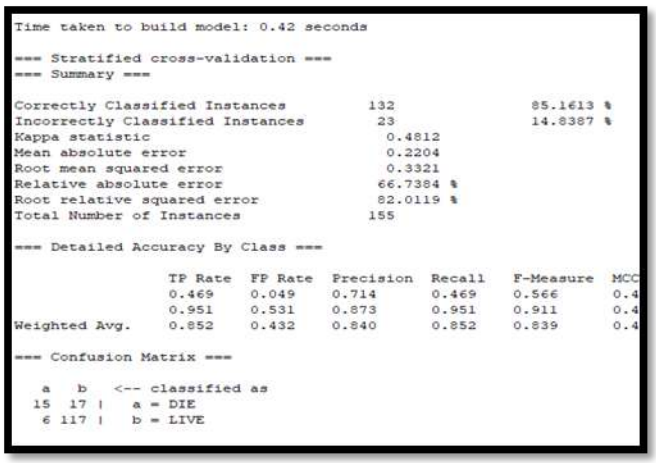
Summary	Weka Result
Correctly Classified Instances	120
Incorrectly Classified Instances	35
Kappa statistic	0.0307
Mean absolute error	0.2872
Root mean squared error	0.3925
Relative absolute error	87%
Root relative squared error	96.922 %
Total Number of Instances	155

Table: 2

a	b	classified as
2	30	a=DIE
5	118	B=LIVE

5.4. Random Forest

We incorporate the Random Forest classifier into our data using 66% of training data from the data set and 44% of the Random Forest test data for unique classification, registry, and additional functions that work to create the largest number of decision trees during the preparation and removal of a class-based class (division) or prediction which means the prediction (rearrangement) of certain trees.



```
Time taken to build model: 0.42 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         132               85.1613 %
Incorrectly Classified Instances        23               14.8387 %
Kappa statistic                          0.4812
Mean absolute error                      0.2204
Root mean squared error                  0.3321
Relative absolute error                 66.7384 %
Root relative squared error             82.0119 %
Total Number of Instances              155

=== Detailed Accuracy By Class ===

                 TP Rate   FP Rate   Precision   Recall  F-Measure  MCC
                 0.469     0.049     0.714       0.469   0.566      0.4
                 0.951     0.531     0.873       0.951   0.911      0.4
Weighted Avg.    0.852     0.432     0.840       0.852   0.839      0.4

=== Confusion Matrix ===

   a   b   <-- classified as
  15  17 |   a = DIE
   6 117 |   b = LIVE
```

When we apply Random Forest classifier we see the results

Summary	Weka Result
Correctly Classified Instances	132
Incorrectly Classified Instances	23
Kappa statistic	0.4812
Mean absolute error	0.2204
Root mean squared error	0.3321
Relative absolute error	66.7384%
Root relative squared error	82.0119%
Total Number of Instances	155

Table: 3

a	b	Classified as
15	17	a=DIE
6	111	B=LIVE

6. DISCUSSION

Many Hepatitis prognostic strategies are presented in this paper. We use three classification methods or algorithms In this paper first the Bayes classifier (classification process) is analyzed. The second type of decision trees used to predict Hepatitis is analyzed. The last type is a rare Forest type of forest

In these ways, the Hepatitis hospital database contains medical information of Hepatitis patients used for the diagnosis of cardiovascular disease. The classification method is used for all proposed work.

7. CONCLUSION

Hepatitis is one of the leading causes of death worldwide and early diagnosis of hepatitis is important. The computer-assisted Hepatitis rate system assists the physician in the form of a diagnostic tool aimed at Hepatitis. Another method for

the characterization of cardiovascular diseases is now being studied in this research paper. After examination, it is established that data science plays a major role in the collection of cardiovascular diseases. The Naive Bayes class for offline training is suitable for disease prediction at the beginning of the phase because 89% of Bayesian students' accuracy and program performance can be determined by exposed and standardized data. The specification of these components will be enhanced by discarding the markers.

REFERENCES

1. Emamjomeh, A. (2013). Molecular BioSystems. Predicting of protein–protein interactions, 8.

2. Hashem, S. (JUNE 2018). Comparison of Machine Learning Approaches. IEEE/ACM TRANSACTIONS ON COMPUTATIONAL BIOLOGY AND BIOINFORMATICS, 861-867.

3. Hoang, P. (2015). HEPATITIS C DIAGNOSIS. Supervised Learning in Baseball Pitch Prediction and Hepatitis C Diagnosis, 91

FOREST FIRE DETECTION, PREDICTION BY USING MACHINE LEARNING

Nudrat Batool, Department of Computer Science Ripha International University, Lahore, Pakistan.

Abstract. *Forest fires cause an imbalance in nature and threaten biodiversity, reducing the wealth of the animal and plant world. Traditional methods of fire prevention are ineffective, and it is currently important to raise public awareness on this issue, especially among those who live near or in forest areas. So we developed a project to forecast and fight forest fires. Existing methods, such as satellite systems, optical sensors, and digital cameras, and wireless sensor networks, have some problems.*

Key Words: *Classification, Random Forest, SMO and linear regression*

1. INTRODUCTION

Forests are the protectors of the ecological balance of the earth. Unfortunately, the forest fire is generally only observed when it has already spread over a large area, making its control and interruption sometimes difficult and even impossible. The result is devastating loss and irreparable damage to the environment and the atmosphere (30% of the carbon dioxide (CO2) in the atmosphere comes from forest fires) [2], as well as irreparable damage to ecology (large amounts smoke and carbon dioxide (CO2) in the atmosphere). Other dire effects of fores fires include long-term catastrophic consequences, such as exposure to local weather conditions, global warming, and the extinction of rare species of flora and fauna.

The problem with forest fires is that forests are generally remote, abandoned / uncontrolled areas filled with trees, dry wood and dandruff, leaves, etc., that serve as a fuel source. These elements form a flammable material and represent the ideal context for initial ignition and serve as fuel for the later stages of the fire.

Fire can be caused by human activities, such as smoking or barbecue parties, or by natural causes, such as high temperatures on a hot summer day or broken glass that works like a collective lens, focusing sunlight on a small spot during a time, therefore leading to fire. Once the ignition begins, the combustible material can be easily replenished to fuel the central location of the fire, which then grows larger and wider. The initial stage of ignition is generally called the "surface fire" stage.
This can lead to the feeding of neighboring trees, and the flame of the fire becomes higher and higher, becoming a "top fire". Basically, at this stage, the fire becomes uncontrollable, and the damage to the landscape can become excessive and can continue for a long time, depending on weather and terrain conditions.

The goal is to detect a fire as quickly as possible, and its precise location and early notification from fire departments are vital. This is a flaw that the present invention is trying to correct, by detecting a forest fire at an early stage, in order to

improve or provide an opportunity to extinguish it before it gets out of control or does significant damage.

There are a series of detection and monitoring systems used by the authorities. These include observers in the form of patrols or control towers, aerial and satellite monitoring, and the increasingly common detection and monitoring systems based on optical camera sensors, as well as various types of detection sensors or combinations thereof.

Synchronization of multispectral images from Google Earth, as well as GPS data, can be achieved using automatic image registration methods. Predicting the occurrence of a fire and identifying vulnerable areas were performed with sufficient precision using the neural network and QGIS methods. Therefore, the main objective of this system is to evaluate in real time the evolution, forecast and control of the fire front in geographic coordinates.

2. LITERATURE REVIEW

Vision-based fire detection systems have recently outperformed traditional sensor-based fire detection systems. The popularity and need for video surveillance in residential, industrial, public, and corporate locations supports the widespread use of vision-based fire detection systems. These fire detection systems are based on detecting the color of the flame and incorporate other features such as movement and frame areas. A fire detection system based on the LUV color space and the mixed transformation is proposed.

Mingyi Zhu, Jiamin Zhang., Design of an intelligent neural network alarm and fire detection system. Its principles are to measure the concentration of smoke particles in controlled cosmic air, while detectors are often installed in places where other particles are present in these granular materials, and smoke particles are sometimes difficult to distinguish, which renders the obtained system information unable to judge and distinguish.

Qiang Yang, Pei Bo and Zhao Juanjuan. Intelligent recognition of forest fire images based on a neural network. The results of the experiment show that the method is superior in recognition speed, recognition speed and anti-interference capacity, compared to the traditional fire recognition method. Therefore, the results illustrate the validity and generalization of the method.

3. METHODOLOGY AND RESULT

Our project includes the following methods. They divide the observed forest area into 'n' zones. Each zone consists of a digital camera, GSM and SENSOR (based on IR), a laptop, a microcontroller and an LCD screen. Each sensor monitors an "M" square foot. The sensors inform the microcontroller. The information is then sent to the monitoring station via GSM. The monitoring station evaluates any sharp and large temperature deviation.

- Micro controller (ATMEL)
- Digital Camera
- GSM modem
- Fire Sensor (IR based)
- Laptop

- LCD
- Stepper motor
- Power supply
- RS232
- MAT LAB 2014 A
- PROTEUS 8.0

3.1. Data and study area

Forest fire detection, prediction and processing of new outbreaks using the proposed approach require multispectral and hyperspectral imaging along with a coordinate system. The Bhutan and Indian Himalayan regions were selected as the research area due to their availability for fieldwork and the availability of high-quality data. Bhutan was chosen as the study area due to the presence of a large forest cover and the absence of such a system in the country [3]. As the number of forest fire incidents increases, over time, forest cover decreases below 60%. Currently, the area of interest for research is located in and around the CST campus with a latitude of 260 50 '51 "N to 260 51' 12" and a longitude of 890 23 '24 "to 890 23' 59" with a total area of 0.68771 sq. M. km (68,771 ha) and for detection we used the image of a specific burned area.We selected cloudless and snowy scenes from image data sets because the snow cap makes it difficult to map smoke, because clouds may be incorrectly classified as smoke due to fire. Effective study and analysis requires multiple satellite images of a specific (multispectral) research area, UAV images, altitude details (SRTM DEM). Relevant information on land, such as moisture content, soil type, rainfall, etc., is crucial for forest management, and some of them were collected through surveys.

4. METHODOLOGY

4.1. Automatic Forest Fire Detection

Fire detection, forecasting, and management can be accomplished through various methods. To detect fire outbreaks, the necessary data is satellite images that can be multispectral or hyperspectral, but hyperspectral images are preferred because they have high spectral resolution and give more accurate results. In satellite images, objects can be detected based on their reflectance and their pixel values in each band, which are red, green, and blue bands. Detection of fires in satellite images can be performed using GIS tools or image processing tools and by joining MATLAB. Classification methods, such as neural network and clustering methods, are used to detect forest fires and other parameters for fire processing. The type of tree species, the characteristics of the vegetation cover, the proximity of the drainage, the urban proximity, the moisture content in the vegetation and the NDVI are the main parameters that can be extracted from satellite / UAV images [4]

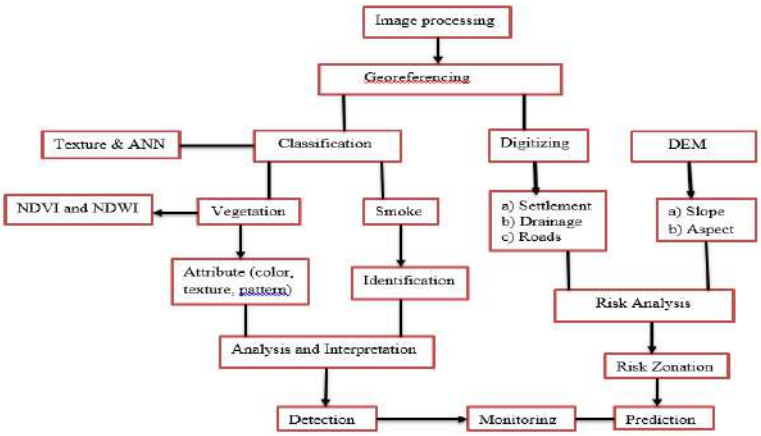


Figure 1: General flow of forest fire detection

Classification methods like random forest, SMO and linear regression using are also implemented for detection, prediction and monitoring of forest fire. Random forest classification is done to classify fire pixels from non-fire pixels and obtain fire. Random forest gives the accurate results (classified image) if the network is well trained. SMO and linear regression methods are implemented in MATLAB to distinctively extracted fire broke out zones.

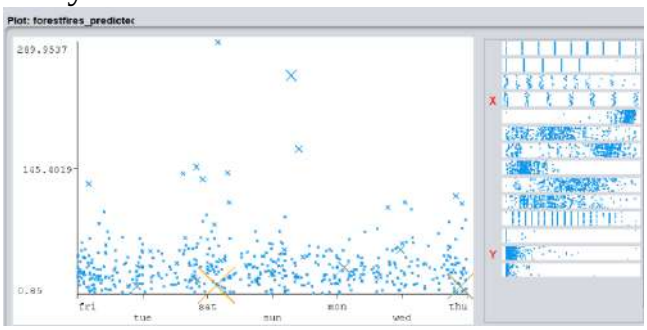
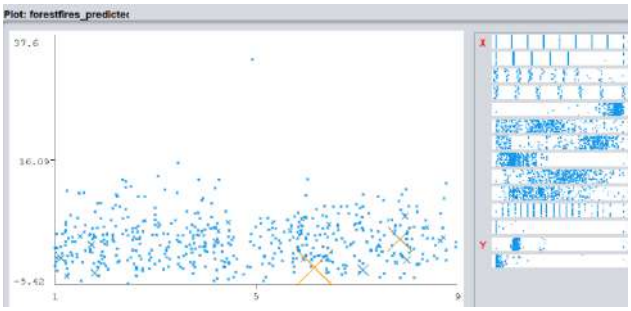


Figure: 2; Random Forest visualization graph error

The random forest algorithm subtracts the red reflectance values from the near-infrared and divides it by the sum of near-infrared and red bands.

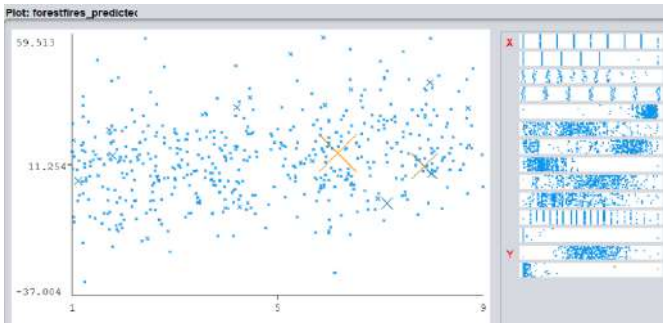
All attributes in the dataset are displayed as a string in the left half and to the right are histograms representing the distribution of various attributes for data mining. The class predicts the risk attributes for 3 different at Label of visualize the graph errors



This formulation allows us to cope with the fact that two identical patches of vegetation could have different values if one were, for example in bright sunshine, and another under a cloudy sky [6]. The bright pixels would all have larger values, and therefore a larger absolute difference between the bands. This is avoided by dividing by the sum of the reflectance. Theoretically, random forest values are represented as a ratio ranging in value from -1 to 1 but in practice extreme negative

values represent water, values around zero represent bare soil and values over 0.6 represent dense green vegetation [5].

Figure 4: Figure of forest fire mapping with NDVI. a: NIR band in grayscale b: Red channels (grayscale view); c: the NDVI values from -1 (black) to +1 (white); d: the forest fire location after threshold. Erroneously classified areas are shown with arrows.



Although a random forest is successful in determining the vegetation index, it has a limitation on free water such as a river, lakes, streams, oceans, and sea. Therefore we adapt a random forest where we can detect the presence of standing water, which helps control a new fire outbreak. Due to the different spectral information of the lakes, uncontrolled classification methods cannot be used for automatic mapping of lakes, rivers, streams, etc. random forest (Huggel et al, 2002) along with the spectral ratio of blue to SWIR has been derived from the preprocessed imagery. For lake, river and stream classification, a blue channel with maximum reflection (B_{blue}) and a NIR channel with minimum reflection (B_{NIR}) are required.

To determine the threshold, lakes that appear dark (in true color) should be investigated due to their higher values of random forests caused by less reflection in the blue channel. In this risk assessment context, it is important to detect all lakes, even if lower general accuracy needs to be accepted. This may be justified by the higher probability of manual correction of misclassifications than detection of undetected lakes (Huggel et al. 2010).

5. FOREST FIRE PREDICTION AND MANAGEMENT

Since boundary detection is the initial step in recognizing objects, it is important to know the different types of detection methods. However, gradient-based algorithms, like the Prewitt filter, have a major disadvantage because they are very sensitive to noise. The size and coefficients of the core filter are fixed and cannot be adapted to this image. An adaptive edge detection algorithm is needed to provide a reliable solution.that is adaptable to the varying noise levels of these images to help in distinguishing valid image contents from visual artifacts introduced by

	Causes	Forest Types	Percentage %
1	Intentional Burning for new grass for cattle	Bamboo, Pine, Coniferous	

2	Agriculture Debris Burning	Brushes and forest, blue and chir pines, mixed coniferous	60
3	Lemon grass Harvesters	Coniferous	
4	Smokers	Chir Pine, Corniferous	20
5	Children playing with ignition sources	Pine	
6	Road side workers	Chir Pine	
7	Picnickers	Bamboo, Pine, blue and chir pines, mixed coniferous	
8	Camp fires	Uncontrolled camps, cooking fire, warming fire and road	
9	Accidental cases	Uncontrolled camps, cooking fire, warming fire and road	15
10	Unknown reasons	Uncontrolled camps, cooking fire, warming fire and road	5

noise. The performance of the Canny algorithm depends heavily on the adjustable parameters. Canny's edge detection algorithm is computationally more expensive compared to Sobel, Prewitt and Robert's operator. However, the Canny's edge detection algorithm performs better than all these operators under almost all scenarios [1]. Therefore canny's edge detection techniques also are adapted to extract fire or objects more clearly.

6. RESULT

6.1. Parameters Associated with Forest Fire

In the annual report on the evaluation and compilation of forest fires, the main problem of the fire outbreak was related to the topographic conditions combined with the fuel load and the unstable wind direction, as well as the lack of trained main energy. It was also discussed that the fire is the result of the practice of burning garbage, lemon grass, Chir pines, and development activities such as road workers, power lines and children playing with matches. It concludes that Blue Pine (Pinus Wallichiana) are highly susceptible to the fire [3].

There, fires generally occur when the vegetation is dry, and during the winter months there is a large amount of fuel in the ground, and it differs depending on geographic location. A fire generally occurs between January and June in eastern Bhutan, while it occurs between November and May in western Bhutan.

6.2. Manually Inventory

In manual inventory, the forest fire are digitized using SMO

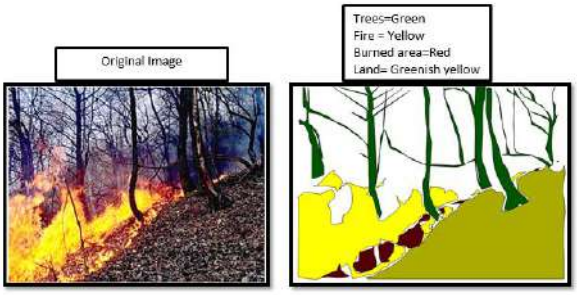


Figure 7: Digitizing the SMO image to obtain the forest fire.

Different classification methods in SMO and linear regression are used for calculating accuracy of each method. Table 1 shows the accuracy of different methods obtained by different classification methods in SMO and linear regression for the image of our study. Maximum likelihood has the highest accuracy with accuracy of 97.7699% and 0.9590. Binary Encoding is the least accurate method with zero value for both accuracy

7. CONCLUSION

The automatic fire detection program using remote sensing techniques could help to detect fire and predict different places of fire zone and also can monitor those fire from further spread provided the multispectral images of study area, and methods are available. It is impossible for anyone to do an on field study on forest fire, thus remote sensing is the only option to conduct study and monitor the fire prevention. Remote Sensing is basically defined as retrieving information about an object without being in contact with that object.

This article implements a simple algorithm and can withstand high temperatures. Robust, economical and easy installation process. Easily retrieve data from an image using DIP. No special knowledge is required to understand the camera data. All components, such as a temperature sensor and GSM, are easy to use and inexpensive. This project can be used to more effectively monitor and plan a forest fire.

8. FUTURE WORK

The development of a multifunctional sensor is mandatory. Use of small satellites such as MICROSATELLITE OR PICOSATELLITIS. A group of small satellites is installed so that each small satellite is used for various purposes. If installing a laptop in a forest becomes a problem, forest fire detection can only be done using GPS and TEMPERATURE SENSORS as well as SMOKE SENSORS

REFERENCE

[1] M. Raman and A. Himanshu "Study and Comparation of Various Image Edge Detection Techniques". International Journal of Image Processing (IJIP), Vol. III., Issue (1), India

[2] An Inconvenient Truth, "Directed by Davis Guggenheim about former United States Vice President Al Gore's campaign [Documentary]," Los Angeles, NY, USA.

[3] D. Tandin, "Fire Situation in Bhutan". International Forest Fire News(IFFN) No.34 (January – June 2006). Annual Forest Fire Evaluation and Compilation Report

[4] P.V. Arun, "An investigation over UAV based framework for forest fire monitoring and prediction in Bhutan,"2014.

[5] Holme, A.McR., Burnside, D.G. and Mitchell, A.A. (1987). The development of a system for monitoring trend in range condition in the arid shrublands of Western Australia.

Australian Rangeland Journal 9:14-20.

[6] Rouse, J. W., R. H. Haas, J. A. Schell, and D. W. Deering (1973). Monitoring vegetation systems in the Great Plains with ERTS, Third ERTS Symposium, NASA SP-351 I, 309-317.

[7] Christian Huggel, W. H. (n.d.). An assessment procedure for glacial hazards in the Swiss Alps. 2004.

[8] H.Frey, C. (2010). Automated detection of glacier lakes based on remote sensing. 10th International Symposium on High Mountain Remote Sensing Cartography, 12.

A COMPARATIVE ANALYSIS BETWEEN DATA MINING TOOLS USING CLASSIFICATION ALGORITHMS

Hafiz Muneeb Ahmad, Muhammad Sohail. Department of Computer Science Riphah International University Islamabad (Lahore Campus), Pakistan.

Abstract. Currently, there is a large amount of data and knowledge accessible to all, and it might be stored different types of storage such as database and knowledge mines, in addition to widley available over the Internet or in hard form. With this amount of data, powerful techniques are required to better interpret these data, which exhaust the capacity of the human to better understand and make decisions. The idea behind this paper we used the data mining tools and comparison of these tools using some methods of classification.Two tools are used in this paper, 1st one is WEKA and 2nd one is the ORANGE tool and apply the three different algorithems of classification and take the one data set from kaggle.Orange tool succeeded the highest improvement in the classification performance.

Keywords: Data mining tools, classification algorithems, Weka , Orange

INTRODUCTION

Begin With the advancement of IT and online sources there is huge amount of data from all over the world and there is a huge need of storage medium. In order to fetch useful information from this data and how to utilize for further proceedings and judgements raw data is collected from different sources and sites and further, to get useful information, it is passesd through a process which provide us the data in a shape from which we can take decesions and our judegements. As it is a broad filed and it is widely used in many fileds of science like economics, mareting etc. WEKA, Orange and many other tools which are available over the internet and they provide the facility of classifications and many others algorithms as well.

The accuracy indicator, provides the information about the correctly classified and incorrectly classified instances and the final perfomance of these tools were analyzed on the based of correctly classified instances.

LITERATURE RIVIEW

Data Mining is the field of data science and the interconnects domain of the computing. That provides the different platforms of the data mining tools that are used for the different classification and clustring algorithms. In this paper classification had perform a diifernt task of algorithms that learn the different data sets of classes. After perform the algorithms that we predict the target results are achieved. In this paper also used the different machine learning methods and techniques are used to complete the classification methods [1].

This paper provides an overview of the different methodologies used to determine the relative importance of input variables in data description models.Two types are data used in this paper,the one is synthetic data with assume algorithms and second one is real data that take the foundry industry. The over all information about data mining evaluation by the different algorithms methods which used in this paper and also used the machine learning methods and they were study the different techniques of machine learning algorithms[2].

In this paper author described the comparison of different data mining tools.They were used the different data sets,atleast eight to nine data sets are taken from the Kaggle. In this paper I read the some algorithms that were used the classification method.The comparison of tools using different clasification algorithms and use the different eight datasets. The author conclude that the data mining tools like knime,weka,orange are slightly different to the each other. But after the comparison of that tools the author said that the weka are more efficient then the other data mining tools [3].

In this paper author explain a fuzzy clustering technique which are find the amount of appropriate clusters preserved the pattern spirit. In this paper data was taken from the kaggle repository. Different experiments for algorithm assessment were done which showed a better routine related to the average performance that broadly used K-means clustering algorithm [4].

METHODOLOGY

The study methodology consists of collecting a collection of free data mining and knowledge discovery tools to be tested, defining the data sets to be used, and choosing a collection of classification algorithms to check the efficiency of the tools.

Tools discription

In order to perform comaparative analysis on the selected dataset two tools were selected. These tools were Weka and Orange. Both of these are famous for data mining. These tools are easy to use and freely available over the internet.

WEKA: Weka is a free and open source software which is available over the internet. This data mining tool used to implement machine learning algorithms. In weka the raw data is selected and preprocessing done on it. After this classifications and many other algorithms applied on it to see their results according to one's need. The pioneers of this tool belong to New Zealand and it was developed at university. This tool provides the facility of GUI. It is easy to use and easy to learn [5].

ORANGE: It also a data mining tool and easily available over the internet. It has an feature of open source and it provides the best visulization than weka tool. It also provides the feature such as drag and drop. This tool is very useful for the executions of machine learning algorithms. It provides extraordinary GUI and add-ons facility. It has python scripting as well. This tool is very helpful for researchers and many researches had already done in it [3].

Dataset explanation

The dataset used in this research was taken form the kaggle. The name of the dataset was breast cancer. This dataset contained 10 attributes and 286 instances. The type of the dataset was nominal. The dataset are used in two different tools for classification.Furhter,it contained two classes. One is recurrence events and second is non- recurrence events.

Table 1: Data set explanation

Datasets	Instance	Attributes	NOC	Type
Breast- Cancer	**286**	**10**	**2**	**Nominal**

On the behalf of this data set,we campare of two tools weka and orange using classification algorthmes in next terms.

Classification algorithm description

After selecting a data set, we choose the classification agorithmes then test the 2 different tool of data mining. After selection the tool,then we choose three different classification algorithms.

1-Random Forest algorithm.

2-Naïve Bayes algorithm.

3-K Nearest Neighbor algorithm.

RESULTS & DISCUSSIONS

In weka tool. In this section I will sahre the reults in weka tool using different algorithems. Different algorithms have have different accuracy using percentile split.

For performance issue Table 2 shows the different three algorithms accuracy usin Weka tool. In Random Forest (correctly classified instances) is 41 and percentage split 71.9298 %. 2nd one (Incorrectly classified instances) is 16 and percentage are 28.0702 %. The running time of this model on test split was 0.01 seconds. In Detailed Accuracy By Class are two classes.These classes are the target variables.One class is no-recurrence-events and second class is recurrence-events.

In **KNN Algorithm** (correctly classified instances) is 38 and percentage split 68.9298 %. 2nd one (Incorrectly classified instances) is 19 and percentage are 29.0702 %. The running time of this model on spilit was 0.02 seconds. In Detailed Accuracy By Class are two classes.These classes are the target variables.One class is **no-recurrence-events** and second class is **recurrence-events.**

In 3rd algorithm **Navive Bayes** (correctly classified instances) is 40 and percentage split 70.1754%. 2nd one (Incorrectly classified instances) is 17 and percentage are 70.1754%. The running time of test model on test split: 0.01 seconds. In Detailed Accuracy By Class are two classes.These classes are the target variables.One class is **no-recurrence-events** and second class is **recurrence-events.**

Table 2: Accuracy provided by weka

Dataset	KNN	NB	RF	Type
Breast cancer	68.929%	70.1754%	71.92%	**Nominal**

In orange tool. In this section I will sahre the reults in orange tool using different algorithems same as weka algos. Different algorithmes have have different accuracy using test and score.

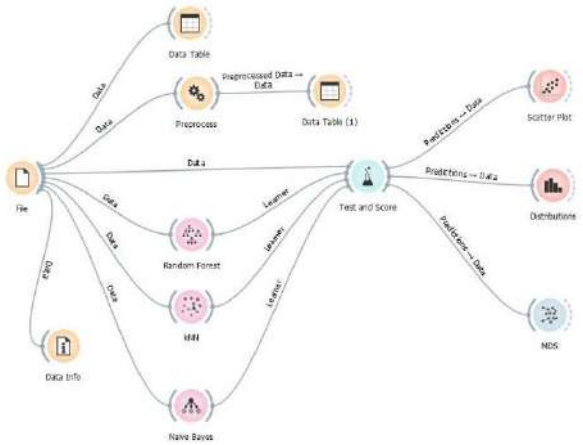


Fig 1: Work flow of orange tool

After shows the work flow,I will apply the three algorithms of classification and shows the test and score in Table 3.

Evaluation Results

Model	AUC	CA	F1	Precision	Recall
kNN	0.670	0.724	0.701	0.704	0.724
Random Forest	0.736	0.734	0.715	0.717	0.734
Naive Bayes	0.736	0.729	0.726	0.724	0.729

Fig 2: Classification accuracy in orange tool

In KNN Algorithm the classification accuracy is 72%, which slightly more efficient from weka tool percentage.

In Random Forest Algo the classification accuracy is 73%, which again slightly more efficient from weka tool percentage.

In Naïve Base Algorithm the classification accuracy is 72%, which again slightly more efficient from weka tool percentage.

So we can say that the percentage of orange tool is more efficient and good in front of Weka Tool.

In orange tool I apply dataset and made the two classses of Heart Disease. One class is **no-recurrence-events** and second class is **recurrence-events.** Show the graph of the two classes of data set.

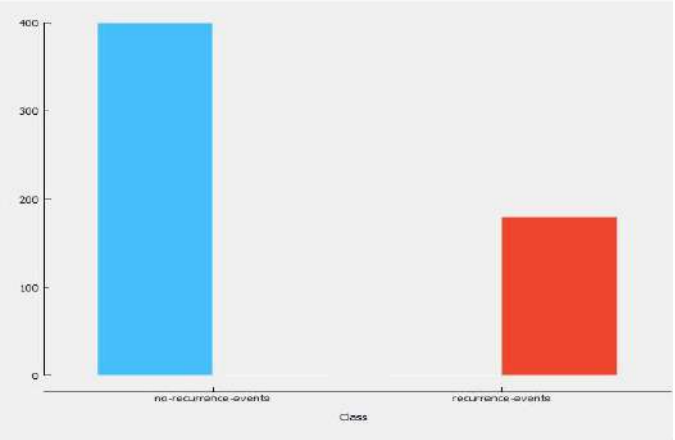


Fig3: Two classes of data set in orange
No-recurrence-events and **Recurrence-events.**

The more efficient results in orange then WEKA, that we use the **satter plot for the better visulization** of algorithms results.

Scatter plots are surely one of the best method visualizations in Orange. And in weka tool no sucs as vizulization with scater plots. This is also true for all other point-based visualizations in Orange, namely t-SNE, MDS, Radviz, Freeviz, and Linear Projection.

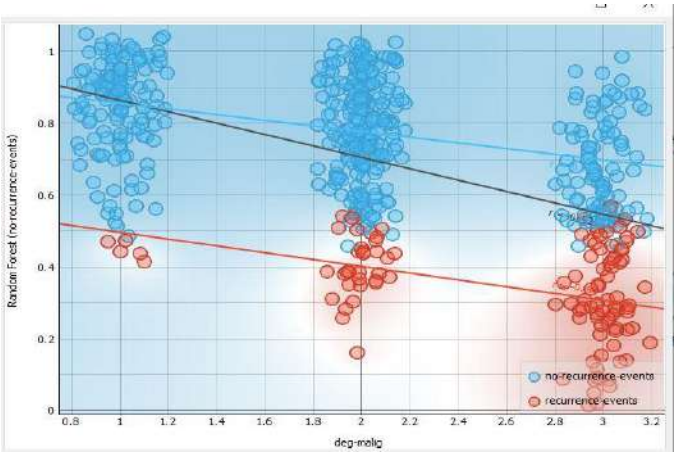


Fig 4: Scator plot

Table 3:Tools Comparison

TOOLS	ALGORITHMS		
	KNN	RF	NB
ORANGE	72%	73%	72%
WEKA	68	71%	70%

These results shows that the Orange tool percentage is more efficient then Weka tool using some classification Algorithms.

So we can say that orange tool is best because it provides better visualization than weka. It provides best accuracy and provides different charts, scatter plot facilities. It is easy to use and easy to learn. It provides the facility of drag and drop features. Interface of orange is best than weka. Orange tool provide better algorithms functionalities than weka that is why it provides better accuracy. It has strong graphics and while using classification algorithms it provides classes with best visualizations.

CONCLUSION

This research has perform a coparison of two different data mining tools for the purpose of classification, one dataset was used to judge the two different tools using three different Algoritms. One is KNN, second is Random Forest and 3rd is Naïve Bayes. It is concluded from this research paper Orange is a best tool than Weka while using classification algorithms. Finally Orange tool is achieved the highest performance use the percentage split mode followed by Weka repectively.

FUTURE WORK

In furture research, we use the different tool like Knime, Tanagra, Rapid Miner in classification algorithms and compare the different kinds of results for more efficient work. This tecchnique will also be helpful on different datasets.

REFERENCES

[1] L. C. Borges, V. M. Marques, and J. Bernardino, "Comparison of data mining techniques and tools for data classification," in *Proceedings of the International C* Conference on Computer Science and Software Engineering*, 2013, pp. 113-116.

[2] M. Perzyk, A. Kochanski, J. Kozlowski, A. Soroczynski, and R. Biernacki, "Comparison of data mining tools for significance analysis of process parameters in applications to process fault diagnosis," *Information Sciences*, vol. 259, pp. 380-392, 2014.

[3] A. H. Wahbeh, Q. A. Al-Radaideh, M. N. Al-Kabi, and E. M. Al-Shawakfa, "A comparison study between data mining tools over some classification methods," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 2, pp. 18-26, 2011.

[4] R. M. Rahman and F. Afroz, "Comparison of various classification techniques using different data mining tools for diabetes diagnosis," *Journal of Software Engineering and Applications*, vol. 6, no. 3, pp. 85-97, 2013.

[5] R. JANOŠCOVÁ, "Mining Big Data in WEKA," *11th IWKM, Bratislava, Slovakia*, 2016.

ВОЗМОЖНОСТИ ВЕЙВЛЕТ ПРЕОБРАЗОВАНИЯ В РАСПОЗНАВАНИИ ЧЕЛОВЕЧЕСКОЙ РЕЧИ

Абдуллаева М.И. ТУИТ, Узбекистан.

Аннотация: *В данной работе рассмотрены основные методы преобразования речевых сигналов. Целью работы является оценить возможности известных методов распознавания речевых данных. Оценены возможности преобразований и выделены преимущества Вейвлет-преобразования над всеми другими.*

Abstract*: The main conversions of speech signal recognition are considered in this paper. The aim of this work is to evaluate the possibilities of well-known speech data recognition methods. The possibilities of transformations are evaluated and the advantages of wavelet transformation over all other ones are highlighted.*

Ключевые слова: *Вейвлет преобразование, речевой сигнал, преобразование Фурье, непрерывное вейвлет преобразование, распознавание речи.*

Key words: *Wavelet transform, speech signal, Fourier transform, continuous wavelet transform, speech recognition.*

Легко заметить повсеместное внедрение систем автоматизированной обработки информации и несмотря на это человеческая речь остается главным в коммуникативном взаимодействии. На то подтверждением служит мировая статистика, указывающая что большей частью информации, передаваемой по каналам, является аудио(речевые)-сообщения. Естественно, что данная картина сохраниться и в будущем, так как ни что не сможет заменить человеческое общение со всеми его уникальными инструментами и эффектами человеческого присутствия, таких как эмоциональные окраски, аутентификация и другие. Тяжело найти заменяющую систему, которая смогла бы быть достойной альтернативой данному речевому процессу. Поэтому задачи обработки речевых данных является одним из важный направлений в обработке информации.

Также, на сегодняшний день возрастают требования к производительности речевых технологий. Несмотря на огромный ассортимент программных решений по распознаванию речи, данная отрасль остается актуальной по причине, что человеческая речь сложна и многогранна. Вместе с тем, можно твёрдо утверждать, что нет программного продукта (ПП) позволяющего распознать узбекскую речь с высокой точностью.

При больших объемах аудио данных ставиться важным обеспечение высокой эффективности передачи и хранения речевых сообщений. Для повышения качества и скорости обработки сигналов важно правильно

очищать его от пауз не потеряв нужную информацию и не повредив исходные данные. Задача удаления пауз в речевом сигнале осуществляется с применением общедоступного и известного метода *k*-средних. Но когда вопрос касается распознавания данных все оказывается не настолько прозрачным и ясным как с удалением пауз.

Для распознавания речевых сигналов имеются ряд преобразований и алгоритмов среди, которых часто используемым является Фурье-преобразование. Широко используемое преобразование Фурье для анализа сигналов, как непрерывных, так и дискретных, оказывается недостаточно эффективным при обработке сложных сигналов. Речевой сигнал является примером нестационарного процесса, в котором информативным является сам факт изменения его частотно-временных характеристик.

Для анализа таких процессов требуются базисные функции, способные выявлять в исследуемом сигнале как частотные, так и его временные характеристики, т.е. функции со свойствами частотно-временной локализации. Такие возможности предоставляют вейвлеты, являющиеся обобщением спектрального анализа. Вейвлет из английского языка переводится как «маленькая волна» и на протяжении своей истории данный термин менялся несколько раз, но смысл самого преобразования практически не менялся, а только дополнялся новыми свойствами. Большой вклад в Вейвлет Преобразование (ВП) внес мировой ученый Альфред Хаар, сформировав уникальный вейвлет (Вейвлет Хаара) обладающий такими свойствами как масштабирование функции (вейвлета) и ее сдвига по оси времени.

Произвольный сигнал $x(t)$ можно представить в умножения коэффициентов разложения C_k на базисную функцию $\psi_k(t)$, то есть $x(t) = \sum_{n=0}^{k}\left(C_k\,\psi_k(t)\right)$. Если в преобразовании Фурье вместо базисной функции $\psi_k(t)$ использовалась функция из косинусов и синусов, то в случаи

ВП вместо базисной функции применяются вейвлеты из набора материнских вейвлетов обладающих рядами свойств. В отличие от традиционного преобразования Фурье, вейвлеты позволяют выявлять особенности сигналов, одновременно локализуя их на временной шкале. Иными словами, вейвлет-анализ можно охарактеризовать как спектральный анализ локальных возмущений.

ВП имеет два вида Непрервыное Вейвлет Преобразование (НВП) и Дискретное Вейвлет преобразование (ДВП). Каждый из них имеют группы задачь, где являются наиболее применимыми.

ВП можно представить в следующем виде:

$$T(a,b) = \int_{-\infty}^{\infty} x(t)\psi_{a,b}{}^{*}\ dt$$

где $x(t)$ – исходный сигнал, $\psi_{a,b}{}^{*}$ - вейвлет с масштабом a и со свдигом b.

При решении задачи $\psi_{a,b}$ будет вқбриатья из набора материнских вейвлетов

таких как, вейвлет мексиканская шляпа, вейвлет Хаара, Вейвлет Добеши, Вейвлет Морла, Вейвлет Гауса и др.

Легко убедиться в том, что Фурье преобразование несмотря на ряд своих достоинств уступает Вейвлет Преобразованию в обработки и распознавании сигналов. В связи с чем становится актуальным разработка программных продуктов по распознаванию узбекской речи с помощью ВП.

Речевой сигнал является примером нестационарного процесса, в котором информативным является сам факт изменения его частотно-временных характеристик. Таким образом, разработка численных алгоритмов вейвлет-преобразования сигналов является актуальной задачей, решение которой будет способствовать повышению эффективности обработки информации в широком спектре приложений, в частности в речевых технологиях.

Литература

1. *Горшков Ю.Г. исследовательский комплекс частотно-временного анализа речевого сигнала с использованием вейвлет-технологии. Вестник МГТУ им.Баумана. Серия Приборостроение, - 2011, №4, с.78-87.*
2. *Новиков Л.В. Основы вейвлет-анализа сигналов. Учебное пособие. 1999. 152 с.: ил.*
3. *Алгоритмы сегментации с применением быстрого вейвлет-преобразования, Т. Ермоленко, В. Шевчук*
4. *Осипов Л. А. Обработка сигналов на цифровых процессорах. Линейно-аппроксимирующий метод : Справочное пособие/ Л. А. Осипов. -М.: Горячая линия -Телеком, 2001.-110 с.. -14146 р. 1 экз.*

Разработка и моделирование алгоритмов быстрого непрерывного вейвлет-преобразования с применением к обработке речевых сигналов, Семинов В.И.

MS EXCELDA SANOQ SISTEMALARI BILAN ISHLASH

Saydaliyev Y. talaba, Alimqulov N. Andijon davlat universiteti, O'zbekiston.

***Annotatsiya**: Ushbu maqolada MS Excel dasturida sanoq sistemalari bilan ishlash keng yoritilgan.*

***Kalit so'zlar:** sanoq sistemalari, standart funksiyalar.*

***Аннотация**: В этой статье подробно описывается, как работать с системами счисления в MS Excel.*

***Ключевые слова**: системы счисления, стандартные функции.*

***Annotation**: This article describes in detail how to work with number systems in MS Excel.*

***Keywords**: number systems, standard functions.*

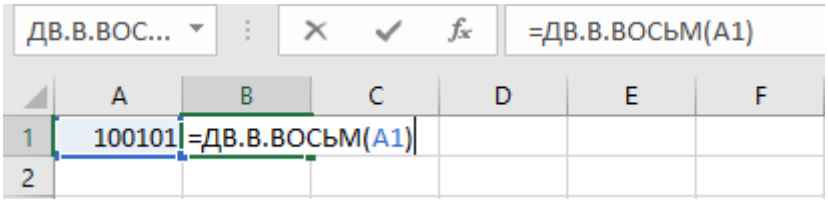
Microsoft Excel dasturida sanoq sistemaari ustida ishlash imkoni ham mavjud. Standart holatda MS Excelda 2 lik, 8lik, 10lik va 16 lik sanoq sistemalari bilan ishlash imkoniyati mavjud. Buning uchun birinchi navbatda MS Excelimizning tizim tilini bilishimiz kerak. Hozirgi kunda bizning MS Excel dasturlarimiz rus tilida bo'lganligi uchun biz rus tilida maxsus funksiyalarni bilib olamiz.

Biz quyidagi 4 ta sanoq sistemalarining biridan ikkinchisiga o'tkazishni ko'rib chiqamiz, so'ngra sanoq sistemalari ustida amallarni bajarish orqali misollarni ishlashni ko'rib chiqamiz.

Bir sanoq sistemasidan ikkinchi sanoq sistemasiga o'tkazish

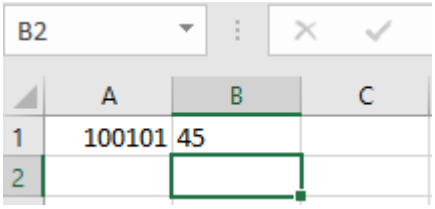
1. Ikkilik sanoq sistemasidan 8, 10, 16 lik sanoq sistemasiga o'tkazish.

- 2 lik sanoq sistemasida 8 lik sanoq sistemasiga o'tkazish uchun ДВ.В.ВОСЬМ() funksiyasidan foydalaniladi. Buni excelda quyidagicha foydalanamiz:



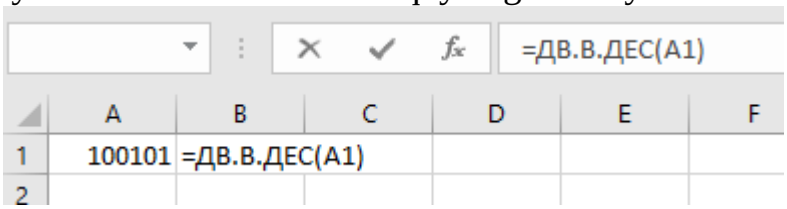
1-rasm.

Bu yerda A1 yacheykaga 100101 2 likda son kiritilgan. B1 yacheykaga esa =ДВ.В.ВОСЬМ(A1) funksiyasi yozilganidan so'ng Enterni bossak xosil bo'ladi.



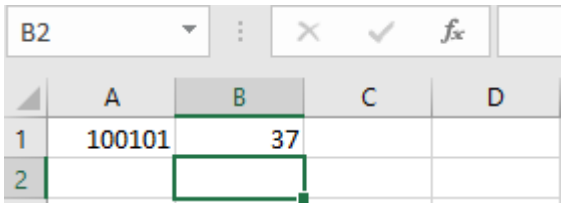
2-rasm.

- 2 lik sanoq sistemasida 10 lik sanoq sistemasiga o‘tkazish uchun ДВ.В.ДЕС() funksiyasidan foydalaniladi. Buni excelda quyidagicha foydalanamiz:



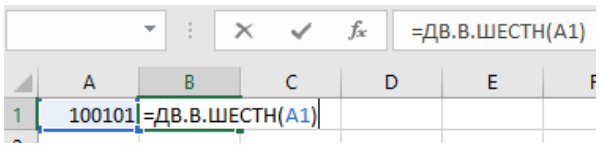
3-rasm.

Bu yerda A1 yacheykaga 100101 2 likda son kiritilgan. B1 yacheykaga esa =ДВ.В.ДЕС(A1) funksiyasi yozilganidan so‘ng Enterni bossak quyidagi rasm xosil bo‘ladi.



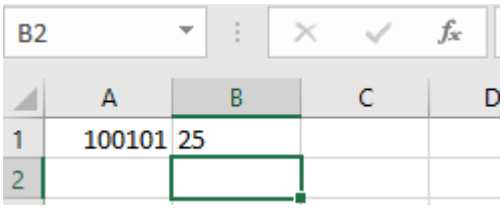
4-rasm.

- 2 lik sanoq sistemasida 16 lik sanoq sistemasiga o‘tkazish uchun ДВ.В.ШЕСТН( ) funksiyasidan foydalaniladi. Buni excelda quyidagicha foydalanamiz:



5-rasm.

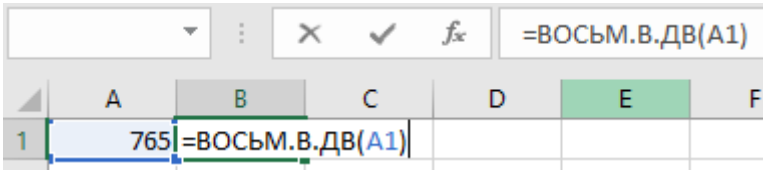
Bu yerda A1 yacheykaga 100101 2 likda son kiritilgan. B1 yacheykaga esa =ДВ.В.ШЕСТН(A1) funksiyasi yozilganidan so‘ng Enterni bossak quyidagi natija xosil bo‘ladi:



6-rasm.

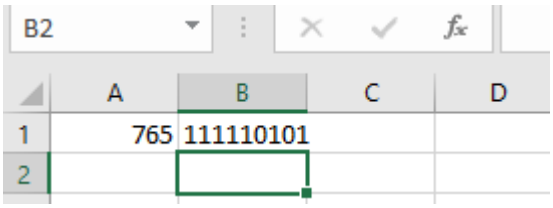
2. Sakkizlik sanoq sistemasidan 2, 10, 16 lik sanoq sistemasiga o‘tkazish.

- 8 lik sanoq sistemasida 2 lik sanoq sistemasiga o‘tkazish uchun ВОСЬМ.В.ДВ() funksiyasidan foydalaniladi. Buni excelda quyidagicha foydalanamiz:



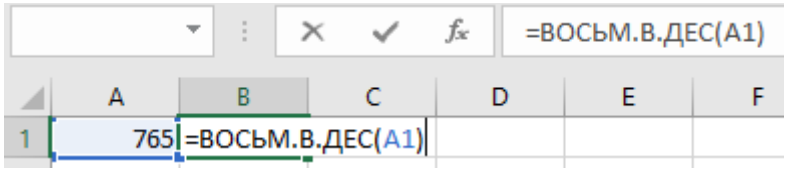
7-rasm.

Bu yerda A1 yacheykaga 765 - 8 likda son kiritilgan. B1 yacheykaga esa = ВОСЬМ.В.ДВ(A1) funksiyasi yozilganidan so‘ng Enterni bossak quyidagi natija xosil bo‘ladi.



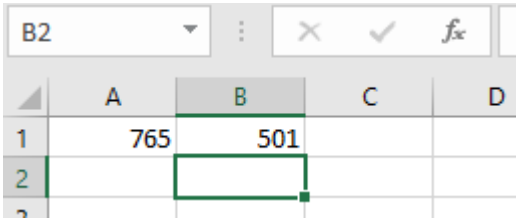
8-rasm.

-8 lik sanoq sistemasida 10 lik sanoq sistemasiga o‘tkazish uchun ВОСЬМ.В.ДЕС() funksiyasidan foydalaniladi. Buni excelda quyidagicha foydalanamiz:



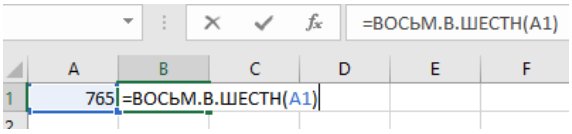
9-rasm.

Bu yerda A1 yacheykaga 765 - 8 likda son kiritilgan. B1 yacheykaga esa = ВОСЬМ.В.ДЕВ(A1) funksiyasi yozilganidan so‘ng Enterni bossak quyidagi xosil bo‘ladi.



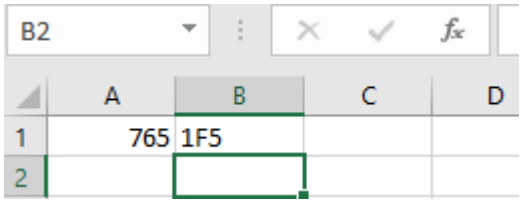
10-rasm.

-8 lik sanoq sistemasida 16 lik sanoq sistemasiga o‘tkazish uchun ВОСЬМ.В.ШЕСТН() funksiyasidan foydalaniladi. Buni excelda quyidagicha foydalanamiz:



11-rasm.

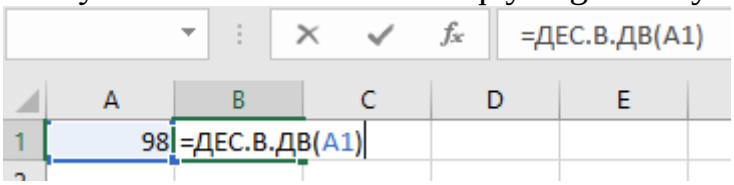
Bu yerda A1 yacheykaga 765 - 8 likda son kiritilgan. B1 yacheykaga esa = ВОСЬМ.В.ШЕСТН (A1) funksiyasi yozilganidan so‘ng Enterni bossak quyidagi natija xosil bo‘ladi.



12-rasm.

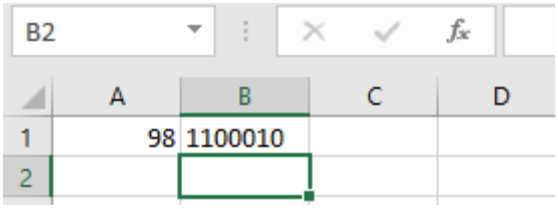
3. O‘nlik sanoq sistemasidan 2, 8, 16 lik sanoq sistemasiga o‘tkazish.

- 10 lik sanoq sistemasida 2 lik sanoq sistemasiga o‘tkazish uchun ДЕС.В.ДВ() funksiyasidan foydalaniladi. Buni excelda quyidagicha foydalanamiz:



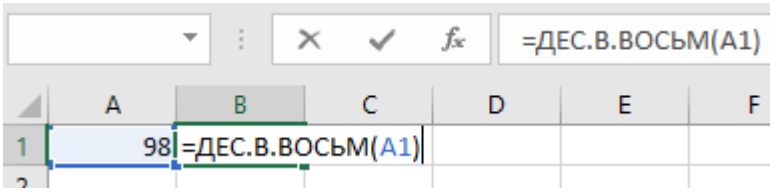
13-rasm.

Bu yerda A1 yacheykaga 98 - 10 likda son kiritilgan. B1 yacheykaga esa = ДЕС.В.ДВ(A1) funksiyasi yozilganidan so‘ng Enterni bossak quyidagi natija xosil bo‘ladi.



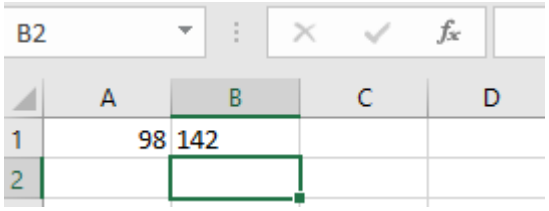
14-rasm.

- 10 lik sanoq sistemasida 8 lik sanoq sistemasiga o‘tkazish uchun ДЕС.В.ВОСЬМ() funksiyasidan foydalaniladi. Buni excelda quyidagicha foydalanamiz:



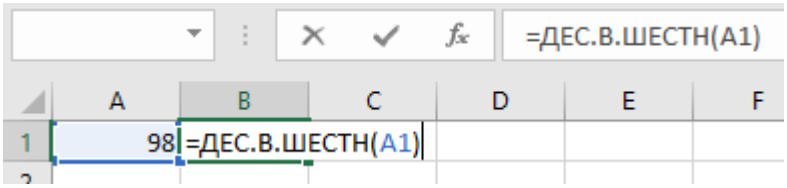
15-rasm.

Bu yerda A1 yacheykaga 98 - 10 likda son kiritilgan. B1 yacheykaga esa = ДЕС.В.ВОСЬМ(A1) funksiyasi yozilganidan so‘ng Enterni bossak quyidagi natija xosil bo‘ladi.



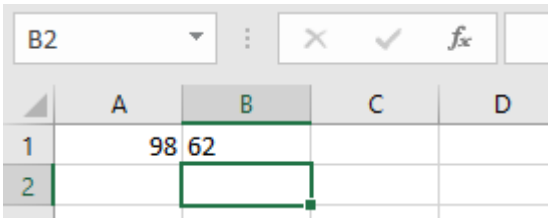
16-rasm.

- 10 lik sanoq sistemasida 16 lik sanoq sistemasiga o‘tkazish uchun ДЕС.В.ШЕСТН() funksiyasidan foydalaniladi. Buni excelda quyidagicha foydalanamiz:



17-rasm.

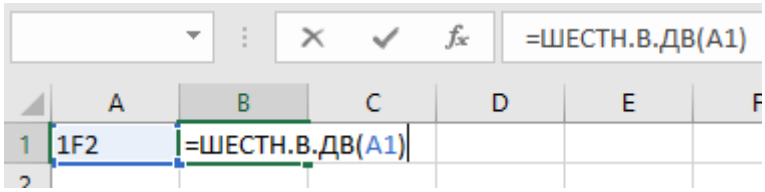
Bu yerda A1 yacheykaga 98 - 10 likda son kiritilgan. B1 yacheykaga esa = ДЕС.В.ШЕСТН(A1) funksiyasi yozilganidan so‘ng Enterni bossak quyidagi natija xosil bo‘ladi.



18-rasm.

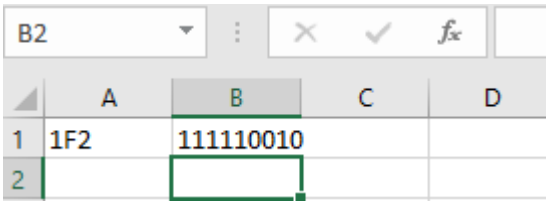
4. O‘n oltilik sanoq sistemasidan 2, 8, 10 lik sanoq sistemasiga o‘tkazish.

- 16 lik sanoq sistemasida 2 lik sanoq sistemasiga o‘tkazish uchun ШЕСТН.В.ДВ() funksiyasidan foydalaniladi. Buni excelda quyidagicha foydalanamiz:



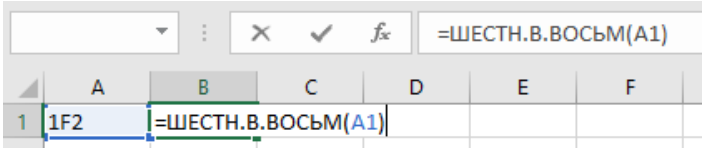
19-rasm.

Bu yerda A1 yacheykaga 1F2 - 16 likda son kiritilgan. B1 yacheykaga esa = ШЕСТН.В.ДВ(A1) funksiyasi yozilganidan so‘ng Enterni bossak quyidagi natija xosil bo‘ladi.



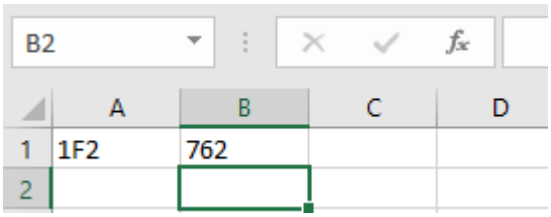
20-rasm.

16 lik sanoq sistemasida 8 lik sanoq sistemasiga o‘tkazish uchun ШЕСТН.В.ВОСЬМ() funksiyasidan foydalaniladi. Buni excelda quyidagicha foydalanamiz:



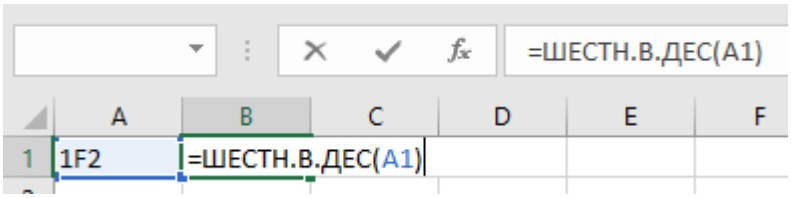
21-rasm.

Bu yerda A1 yacheykaga 1F2 - 16 likda son kiritilgan. B1 yacheykaga esa = ШЕСТН.В.ВОСЬМ(A1) funksiyasi yozilganidan so‘ng Enterni bossak quyidagi natija xosil bo‘ladi.



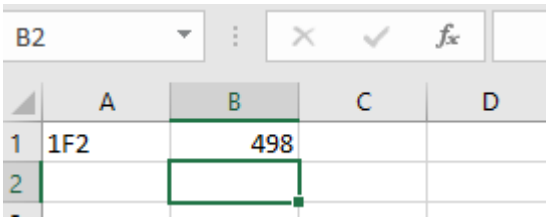
22-rasm.

- 16 lik sanoq sistemasida 10 lik sanoq sistemasiga o‘tkazish uchun ШЕСТН.В.ДЕС() funksiyasidan foydalaniladi. Buni excelda quyidagicha foydalanamiz:



23-rasm.

Bu yerda A1 yacheykaga 1F2 - 16 likda son kiritilgan. B1 yacheykaga esa = ШЕСТН.В.ДЕС(A1) funksiyasi yozilganidan so‘ng Enterni bossak quyidagi natija xosil bo‘ladi.



24-rasm.

Adabiyotlar.

1. *Ильичева О.А. Богачева М.Н. Основные понятия, Системы счисления, Логика, Лабораторный практикум в Excel. Мутодическая указания. – Ростов-на-Дону, 2019, -72 с.*
2. *https://exceltable.com/funkcii-excel/primery-funkcii-osnovanie*
3. *https://math.semestr.ru/inf/excel.php*

ДОИРАЛАРНИ КООРДИНАТАСИ ВА ДИАМЕТРИГА МОС РАВИШДА ЭКРАНГА ЧИҚАРУВЧИ СТАНДАРТ ПРОГРАММАНИ ОСМОНДАГИ ЮЛДУЗЛАР ТАҚСИМОТИГА ҚЎЛЛАШ

М. К. Махкамов, М.М. Муминов, Б.Ў.Назарова, Андижон Давлат Университети, Ўзбекистон.

Абстракт. *Мақолада SI+ программалаш тилида тузулган стандарт программа асосида астрономик илмий тадқиқотларида долзарб бўлган юлдузлар осмони харитасини тузиш масаласи ечими ёритилган. Бирламчи материал сифатида юлдузлар тартиб рақами, Х,У координатаси ва юлдуз кўринма диаметридан иборат жадвал бўйича тақсимот қилиш программаси тузилиб, компютерда намойиш этиш имконияти яратилди.*

Абстракт. *В статье опысивается стандартная программа на языеке SI+, расчитанная на составению карту звездное небо. Первычном материалом для прораммы служать каталог порядковые номера звезд, Х,У координаты и видимые звездные величина, на основание которых демонструются карта распределения звезд на небо.*

Abstract. *The article explores the standard SI + language program designed for mapping the starry sky. The primary material for the program is a catalog of serial numbers of stars, X, Y coordinates and visible magnitudes, on the basis of which a map of the distribution of stars to the sky is displayed.*

Калитли сўзлар: *программма, юлдуз, стандарт дастур, диаметр, тақсимот, координаталар, юлдуз катталиги.*

Ключевые слова: *программма, звезда, стандартная программа, диаметр, распределение, координаты, звездная величина*

Keywords*: program, star, standard program, diameter, distribution, coordinates, star magnitude*

У

шбу мақолада ST.exe кўринишда номланган программа ва EGAVGA.BGI, LITT.GHR ёрдамчи ёйилмалар ва NGC.DAT, SCALE.DAT маълумотлар базаси ёрдамида жадвалда келтирилган маълумотлардан юлдузлар харитаси намойиши жараёни кўрсатилади.

Программада

initgraph(&dr,&md,"");
settextstyle(SMALL_FONT,HORIZ_DIR,2);
setbkcolor(WHITE);
setcolor(color); каби стандарт программалардан фойдаланилади.

Ушбу программада NGC.DAT маълумотлар қуйидагича ташкил этилди:

Биринчи устун - доирачалар тартиб рақами,

Иккинчи устун – доирачалар Х-координатаси,

Учунчи устун – доирачалар У- координатаси,

Тўртинчи устун – доирачалар диаметри.

Биз бу маълумотларни юлдузлар χ ва h Per юлдуз тўдаси жадвалининг 19317 та объекти учун шакллантирдик ва жадвалнинг ихтиёрий 12 юлдузи учунт қуйидаги маълумотлар мисол сифатида келтирамиз:

40	-84.9	-3148.4	5.91
89	-79.9	-3127.6	5.75
92	-92.9	-3126.7	6.69
6458	-1781.9	-909.2	6.98
8901	-218.9	-165.2	6.97
8909	405.1	-163.2	3.29
9560	1326.1	15.8	6.55
9581	162.1	20.8	5.90
9626	-2684.6	31.8	4.59
11516	-569.9	450.8	6.98
11554	333.1	459.8	6.68
1 17919	-2708.8	2606.8	5.56

SCALE.DAT яратиладиган доиралар тасвирининг ўлчамлари ва кўринишини белгилайди. Унда учта бирлик киритилиб, биринчи ва иккинчиси юлдуз координатасининг абсолют қиймати катталигига, учинчиси юлдуз диаметрига мос бирликка мосланади.

Масалан, юқоридаги жадвалга мос маълумотлар қуйидагича танланди:

3350.0

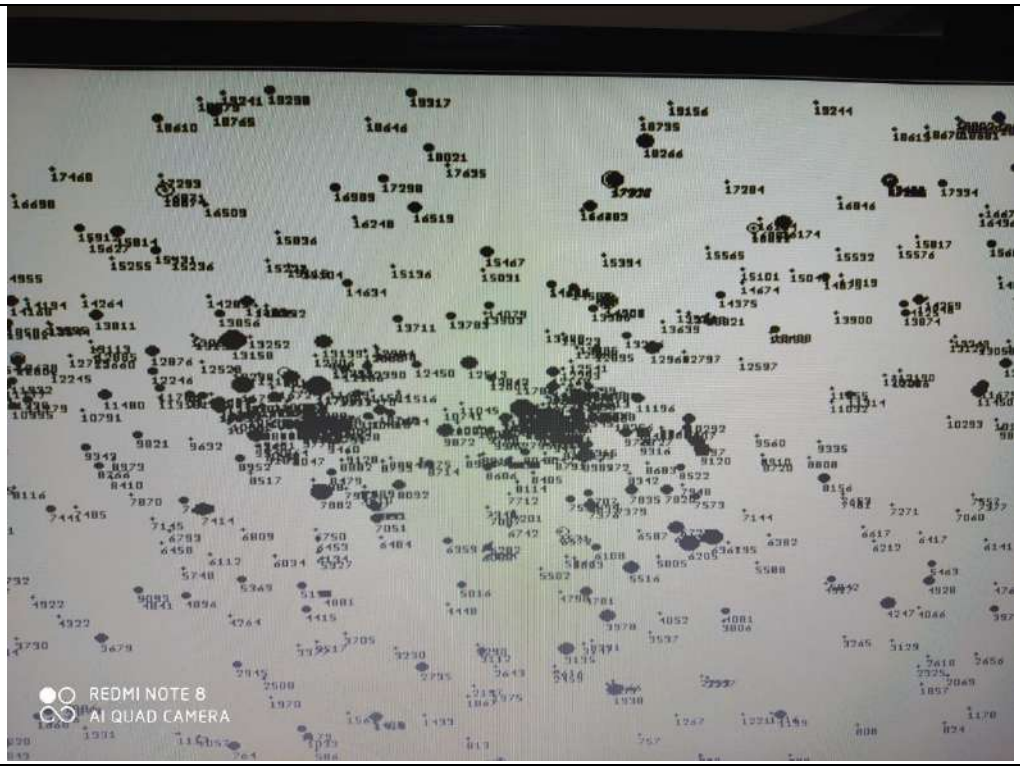
3331.0

7.0

Программа ST.exe ни босиш билан ишга туширилади.

Бунда экранда киритилган NGC.DAT маълумотлар бўйича SCALE.DAT га мосравишда юлдузлар харитаси пайдо бўлади.

Расмда ҳар бир тасвир киритилган юлдуз координаталари ва юлдуз катталигига мос равишда чизилиб, ҳар бир доира остига юлдуз тартиб рақами қўйилади.



1-расм. χ and η Per тарқоқ юлдузлар тўдасининг осмон сферасида тақсимотини программаси ёрдамида тасвирлаш натижаси.

Программадан фойдаланиш жараёнида расмнинг масштаби ва юлдузлар катталиги SCALE.DAT маълумотлар файли қийматлари орқали бошқарилади.

X ва У қийматлари ортгани сари экрандаги расм сиқилиб боради ва киритилган юлдуз диаметрига мос равишда тасвирлар катталтиги ўзгаради. Белгиланган миқдорлар нотўғри киритилган холатда экран қора тусга киритилади ва бу холда Esc тугмачани босиб, SCALE.DAT F4 тугмача ёрдамида кириб, қийматларнинг ўзгартирилиши зарур.

Хулоса: Мазкур программа ёрдамида жадвалда берилган сонли маълумотларни компьютер экранида юлдузлар харитасини шаклида намойиш қилиш мумкун. Натижада ўқув жараёнида кўргазмалиликка эришилиб, талабалар юлдузлар осмони тузулиши бўйича дунёқараши шакллантирилади.

Программанинг тўлиқ матни АДУ информатика кафедраси компютер базасида сақланган.

Адабиётлар:

1.Программирование на языке СИ. Kpolyakov.spb.ru/school/c.htm

2.M.Muminov Star Catalogue in region Open Star Cluster χ and η Per. Publ. Odessa obs. 2018, N 102, pp. 23-34

МОДЕЛИРОВАНИЕ ГИДРОГЕОЛОГИЧЕСКИХ ПРОЦЕССОВ НА ОСНОВЕ ГЕОИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Егамбердиев Х.С., Ишанходжаев О.А., Ёсинов С.Д., Ирмухамедова Н, Бегимкулов Д.К., Жуманов Ж.Ж., Ташкентский государственный технический университет, Узбекистан

Қисқа мазмуни. *Мақолада геоахборот технологиялар асосида гидрогеологик тизимларнинг геофильтрация жараёнларини моделлаштиришни масалалари ҳақида фикр юритилади. Ер ости сувининг шаҳар ҳудуди ва унинг атрофини захланиши ҳолатини моделлаш жараёнида дастлабки маълумотларни йигишда геоахборот технологиялар асосида таҳлил, бошқариш ва такомиллаштириш усуллари таклиф қилинади.*

Калит сўзлар: *математик моделлаштириш, фильтрация жараёнлари, геоахборот технологиялари, геофильтрацион схемалаштириш, геомаълумотлар базаси.*

Аннотация. *В стате рассмотриваются вопросы моделирования процессов геофильтрации гидрогеологических систем на основе геоинформационных технологий. Предлогается методы анализа, управлении и усовершенствование сбора исходаных данных на основе геоинформационных технологий при моделирования процессов подтопления городских территорий и их агломерации.*

Ключевые слова: *математическое моделирование, фильтрационные процессы, геоинформационные технологии, геофильтрационные схематизации, базы данных.*

Annotation . *In the article is considering modeling issues and geofiltration of processes hydrogeological systems on the base of geoinformation technologies. Offers the development of methods for the analyses, management and collecting initial data on the base of geoinformation technologies to the modeling of flooding of processes cities territory and this agglomeration.*

Key words: *mathematical modeling, filtration of processes, geoinformation technologies, geofiltration schematics, data base.*

ВВЕДЕНИЕ. Результатом интеграции математического моделирования гидрогеологических процессов и технологий геоинформационных систем (ГИС) стало формирование нового методического направления работы с геопространственными данными, как в сетевом режиме мониторинга, так и графической визуализации геологических объектов, которое получило название ГИС-системы. Основным достоинством этих систем является то, что она вяжет между собой геоданных, рассредоточенные по всему объекту изучения с географической привязкой, т.е., распределенная географическая информация. Из их множества стоит выделить следующие основные особенности [1]:

- Справочно-информационное картографическое обеспечение;

-Визуально-картографическое представление цифровых баз геоданных в интересах их распространения;

- Математическо-тематическо-картографическое и информационно-аналитическое обеспечение;

- Проблемы проведения исследований по повышению скорости обработки геоданных, формирования и составление схематически картографических изображений, повышения функциональности предлагаемых моделей;

-Совершенствования способов хранения больших объемов географической информации и повышения качества картографической визуализации данных.

Как видно, все перечисленные особенности характеристик опираются на картографическое представление геолого-гидрогеологических данных или его результата, что позволяет считать практически все интерактивно точно привязанные данные. В свою очередь, это делает возможными простые рабочие процессы для выполнения всех видов интересных и сложных аналитических операций и математических вычислений.

В результате анализа многочисленных экспериментов математического моделирования гидрологических процессов на примере г.Карши (Рис. 1) рассматриваются изменение режима потоков поверхностных и подземных вод долины р.Кашкадарьи, и повышение требований к рациональному использованию имеющихся на фоне общего подтопления города и его агломерации привели к выделению особого класса задач моделирования геофильтрационных процессов гидрогеологических объектов в нарушенных природных и техногенных условиях. В отличие от традиционного представления, массив исходных данных и электронный вариант результатов изображен интеграции средствами ГИС, реализована возможность разработки содержания электронной карты в полуавтоматическом режиме.

Одним из малоизученных аспектом данной проблемы является учет пространственного характера взаимосвязей составляющих гидрогео-логических условий потока подземных вод, инфильтрационных каналов, орошаемых территорий и т.п. До недавнего времени в качестве основного источника информации считались данные режимных наблюдений, представленные в основном в числовой форме. Особое внимание при схематизации гидрогеологических условий уделялось геофильтрационным характеристикам основного водоносного горизонта четвертичных отложений – по территории, площади и мощности водоносных горизонтов.

Слои карты активизируют мощные аналитические возможности, например, гидрографический и наблюдательный сеть, а также гидроизогипсы из нескольких накладывающихся друг на друга схем позволяет выполнять интеграцию и полезные операции наложения. И соседние ячейки внутри населенного пункта использованы для расчета зональной статистики, близости к выбранным объектам, моделирования поверхности и функций потока. Включая 3D и время в аналитику различных возможностей [2].

Различные слои и инструменты в ГИС объединяются в прогрессивную модель. Каждый инструмент пространственного анализа выполняет с географическими данными небольшую, но существенную операцию, например, объединение слоев с взвешенным наложением, вычисление расстояний от каждой ячейки до определённых пространственных объектов или отслеживание пути прохождения потока через поверхность. В свою очередь, эти производные слои могут передаваться в дополнительные инструменты, которые генерируют дополнительные результаты. Это позволяет собирать последовательности операций и создавать свои собственные алгоритмы пространственного анализа. С их помощью данные использованы для моделирования на базе программного комплекса ModFlow.

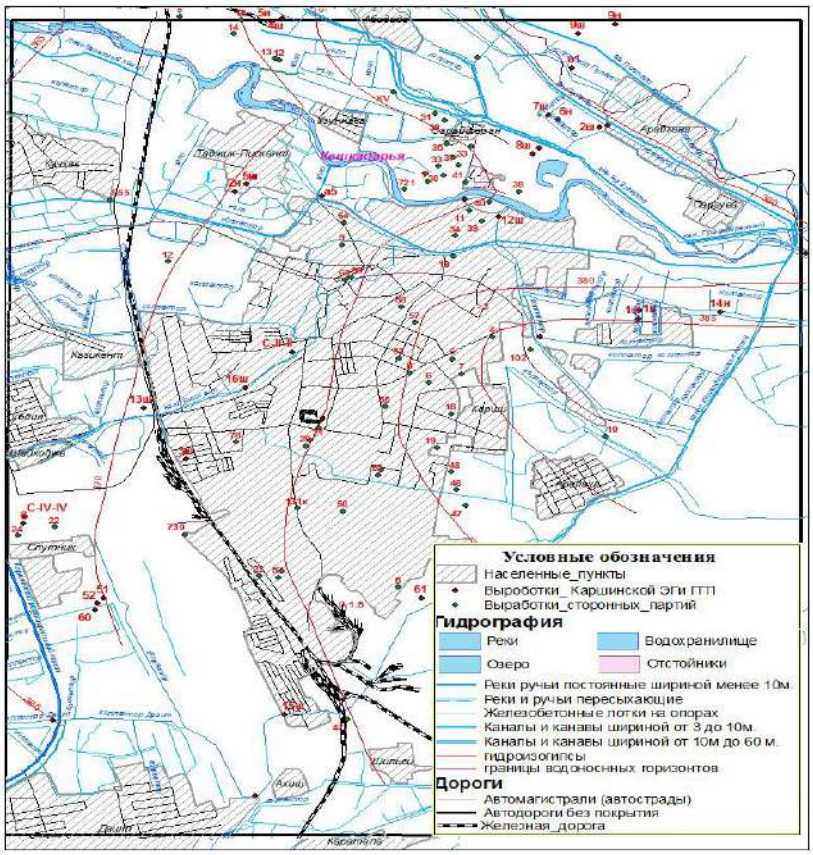


Рис. 1. Геофильтрационная схематизация гидрогеологических условий на примере г.Карши на основе ГИС.

По мнению авторов необходимо изменить ход гидрогеологических исследований и придать им комплексный характер, соответствующий требованиям гидрогеологических моделей. Кроме создания "геобазы данных", ведение работ по созданию "банка моделей", так как результаты гидрогеологических исследований будут представлены в виде моделей.

Необходимость интеграции ГИС в решении задач моделирования и принятие решений в рамках единой вычислительной системе объясняется недетерминированностью, непредсказуемостью поведения характеристик гидрогеологических объектов и их параметров. В этой связи рассматриваются аспекты применения геоинформационного подхода в процессы моделирования измененных, т.е. антропогенных условий. Однако, в

условиях интенсификации влияния техногенных факторов на гидрогеологические объекты и потока подземных вод, качества этих данных не всегда отвечает требованием адекватности, точности, оперативности и надежности. К тому же в процессах поддержки принятия решений характер пространственного распределения характеристик и их параметров учитываются в основном косвенно. В этой связи перспективным является создания сети автоматизированного мониторинга на базе приборов замера уровни и разработка модели процессов геофильтрации подземных вод в многослойных средах основанной на обработке данных пространственного характера [3]. Такой подход представляется более перспективным в плане визуализации и принятия решений.

Заключение. Разработана методика исследований по моделированию геофильтрационных процессов с интеграции ГИС, поддерживающие решения значительного объема задач геофильтрации. По этим подходам определены баланс подземных вод и рассчитаны по ней гидрогеологические параметры водоносного горизонта территорий г.Карши. Рассмотрены особенности и обоснованы граничные и начальные условия.

Геообработка представлялся собой методическое выполнение последовательности операций с географическими данными, в результате которого созданы новая базы данных и введены информации. Сходство данных между объектом и его модельным результатом достигается на основе географических распределенных информаций, и модель воспроизводит режим уровня подземных вод с точностью, определяемой количеством и качеством натурной информации, а также предпосылками идеализации и схематизации. Функции и дополнительные модули ГИС обладает самым богатым набором инструментов для схематизации, моделирования и комбинирования слоев данных. Геофильтрационная модель объекта исследований является примером возможностей интеграции моделирования и анализа данных в ГИС.

Литература:

1. *Джуманов Ж.Х. Геоинформационные технологии в гидрогеологии // Монография, Изд. ГП «Институт ГИДРОИНГЕО», 2016. 258 с.*
2. *Мавлонов А.А., Джуманов Ж.Х. Гидрогеоинформационная модель подземных вод в геоинформационных системах//Геология и минеральные ресурсы. –Т. 2006. № 2. С.55-59.*
3. *Хабибуллаев И., Джуманов Ж.Х. Об информационно-коммуникационной технологии в гидрогеологии. Геология и минеральные ресурсы. –Т.: 2014. № 1. С.48-54.*

LEARNING FOREIGN LANGUAGES WITH THE USE OF INFORMATION AND COMMUNICATION MEDIA

Normatova M.N. Gulistan State University, Uzbekistan.

Abstract*: The article deals with the teaching of foreign languages in a communicative situation with the use of new information technologies, the organization of teaching processes using multimedia software tools of modern information and communication technologies, functional linguistics, psycholinguistics, methodology. the organization of students' learning activities, the use of emotional tools in language teaching in accordance with the needs of life and communicative situations, the need to introduce new information technologies in this process. A new approach to foreign language teaching, the organization of the teaching process with the help of new information technologies, the use of interactive methods that shape the linguistic and communicative competence of students in this regard can overcome some shortcomings in the traditional teaching process.*

Key words*: information-communication technologies, multimedia tools: audio, video and animations, creative activity, modern education.*

Integrated education in the education system of the Republic of Uzbekistan ensures not only the intellectual development of the country, but also its spiritual development. In this regard, President Mirziyoyev pays attention to this issue and emphasizes the following: We must not forget that the foundation of our future will be laid in educational institutions, in other words, what will be the future of our people, how will our children be educated today? depending on. Our country is carrying out comprehensive reforms to require continuous learning of foreign languages at all levels of the education system. Presidential decrees introduce advanced methods of teaching foreign languages using modern pedagogical and information and communication technologies, radically improve the system of training specialists fluent in foreign languages and, on this basis, their achievements in world civilization and use of world information resources. a wide range of tasks has been identified to create conditions and opportunities for the development of cooperation and communication. As the first President I.Karimov said in his book "High spirituality is an invincible force": we must first nurture future great writers, great poets, great creators. Today, it is important to ensure the priority of the education system in the training of specialists in many educational institutions in our country, including the introduction of new information technologies in foreign language teaching. It is known that the expression of thinking, communication between peoples, spiritual and other communication is expressed through language.

Of course, new information technologies play an important role in teaching foreign languages. Mankind has entered a period of development of a new

information society. In developed countries, 60% -70% of the working age population is currently employed in the field of information processing. Informing the public has led to changes based on the introduction of new information technologies. In the process of teaching foreign languages, new levels of information technology can be used at different levels, depending on the nature of the activity and the national economic opportunities and characteristics of the language learner. Although new information technologies are used to facilitate difficult operations, they do not negatively affect the types of creative activities. At the same time, the introduction of new information technologies in the language learning process, as a rule, does not lead to a complete revision of the requirements for the language learning process, but to the complexity of the nature of the language learning process with the participation of new information technologies. this leads to an increase in demand. Currently, all higher education institutions in the country are connected to the national high-speed optical e-learning network. This allows them to hold regular video chats. Multimedia is a computer technology that allows the exchange of information between the user and the computer using sound (speech, music, noise), graphics (pictures, photos, drawings), animation (video, cartoons). Multimedia technology allows you to use several ways to present information at the same time: text, graphics, animation, video and audio. The most important feature of multimedia technology is interactivity - the ability to influence the user in the operation of the information environment.

The main purpose of multimedia technology is to create software products with sound, video, animation and other visual effects. It includes interactive interfaces and controls for multimedia software products. In addition, multimedia technology allows the user to design, as well as create static (motionless) and dynamic (moving) images, and distribute the results of their creative work to the external environment through communication channels.

Areas of application of multimedia:

- in training
- in medicine
- in the military field
- in business and management
- in art and creativity
- In archival work
- in artificial intelligence systems
- vertual being and others.

The first multimedia encyclopedia was created in education in 1986. Reading efficiency increased by 30%. Started using an interactive test system. Today, the only and most modern form of information presentation can be software in the form of textual information, images, slide shows, voice-enriched, video and animation, three-dimensional graphics. The main difference between presentations and other forms of information presentation is that they are content-rich and interactive, that is, they tend to change in a defined way and reflect the user's

attitude. Presentations created using multimedia technologies are understandable and effective for listeners and users.

Today, the process of teaching foreign languages is promoted by the disciplines of functional linguistics, psycholinguistics, methodology, teaching foreign languages in communicative situations using new information technologies, co-organization of students' learning activities, language teaching in accordance with life needs and communicative situations. the use of emotional impact tools highlights the need to incorporate new information technologies into this process. Therefore, the creation of a methodological system for the use of functional, communicative, emotional capabilities of a foreign language with the help of new information technologies, thereby increasing the effectiveness of education is promoted as one of the most pressing issues of today. A new approach to the process of teaching foreign languages, the organization of the teaching process with the help of new information technologies, the use of interactive methods that shape the linguistic and communicative competence of students in this regard can overcome some shortcomings in the traditional teaching process. There is no doubt that it will be new.

Today's student must not only learn, but also learn to work on themselves and evaluate their knowledge correctly, because today's student is a hopeful, free-thinking, independent thinker of independent Uzbekistan. is a generation of intellectual potential.

References:

1. *Karimov I.A. - "High spirituality is an invincible force" Tashkent —Manaviyat-2008*
2. *Begimkulov U.Sh. Organization of pedagogical education in the environment of modern information technologies. G'G —Pedagogical education jur, № 1, 2004 - 25-25 pages.*
3. *Sabirova D.A. Multimedia systems and technologies. Textbook -T: TGEU, 2013.*

CREDIT CARD FRAUD DETECTION USING MACHINE LEARNING

Muhmmad Umar[a], Qaiser ali[b] Imran Ahmed. [a]Riphah international university lahore. [b]Dept. Of Computer science, Lahore, Pakistan.

Abstract. *Credit card fraud is frequent and can lead to significant financial loss. Credit card fraud detection, data mining cases, are questioned for two main reasons: fraudulent behavior in common practice, and second, credit card fraud data collection. Variable selection and data set identification techniques. This article investigates the performance of logistic regression on disproportionate data, close neighbors, and credit card fraud. Each trick is created using a set of engine training models, which is chosen by the best diagnostic method. Three techniques are used for raw and pre-generated data. This project is implement in python. The work of this technology is estimate on the basis of accuracy, duration and level of permanent diagnosis. Results of this shows of best accuracy for logistic regression and random forest classifier, classifiers are 99.92%, and 99.96% respectively. The comparative results show that logistic regression performs better than naïve bayes and knearest neighbour techniques.*

Keywords: *machine learning, logistic, regression random forest classifier.*

INTRODUCTION

Credit card fraud is the major problems in today's business. However, in order to effectively prevent fraud, it is important to understand the mechanisms that implement fraud. Credit card fraud is a great way to get rid of fraud. Many procedures have been adopted. To sum up, credit card fraud is defined as: If a person uses another person's card owner card holder not know about issuer no card. - Except for use, the card holder has no connection with the card holder or issuer, and does not contact the cardholder or pay for the purchase. This perform in the following ways:

- illegal fraud (intentional misuse) uses unauthorized accounts and / or personal information
- Use your account illegally or for personal gain. no
- Misrepresentation of account information to obtain goods and/or services. Credit fraud is divided into two types:
- Offline.
- Online.

Offline fraud is done using an original card stolen from your store or call center. In most cases, the card issuer can block the card before using it for fraud. Online scams are not available for the web, phone purchases or card holders. You only need the card information, no manual signing or card printing at the time of purchase.

Location: Purchase made from different location
Items you buy: If you deviate from your regular buying pattern or time
Frequency: Make a large number of transactions in short period of time
Amount: Suddenly if the costly items are purchased.

lITERATURE REVIEW

4. You Dai, et. al: This article discusses the forest algorithm used to identify fraud. Random forests consist of random trees and wagon-based random forests. They describe 91.96% and 95.77% and accuracy respectively. The second type of paper is better than the other.

5. Suman Arora: In this article, many machine learning monitoring algorithms apply to 70% training and 30 test test datasets. Random Forest,Decision Tree, Beginner Algorithm and KNN (compare to each other), ie 94.59%, 95.27%, 94.59%, 93.54%, 93.24%,90.54% and 94.25%. To summarize this article, the highest rating for SVM is 0.5360 FPR and the lowest rating for the overall range is 0.0335.

6. Kosemani Temitayo Hafiz: In this paper, they describe flow chart of fraud detection process i.e., data Acquisition, data pre-processing, Exploratory data analysis and methods or algorithms are in detail. Algorithms are K- nearest neighbour (KNN), random tree, AdaBoost and Logistic regression accuracy are 99.92%, 94.32%, 57.73% and 98.24% respectively.

RESEARCH QUESTIONS

4) How is credit card fraud detected?

5) How to control large amount of people truncations through Credit Card?

6) Is it possible to find out Credit Card Fraud Detection Using Machine learning?

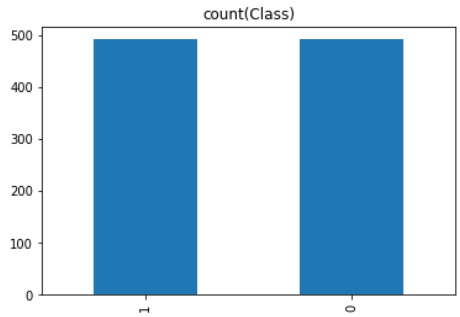
METHODOLOGY

MACHINE LEARNING AND ITS ALGORITHMS

A. Machine learning

Machine learning is a set of methods that can automatically identify

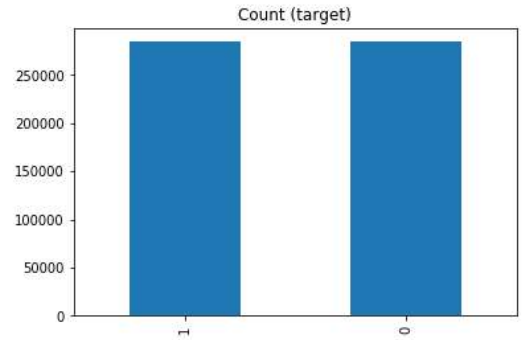
```
Random under-sampling:
1    492
0    492
Name: Class, dtype: int64

: <matplotlib.axes._subplots.AxesSubplot at 0x1e05963c7f
```



patterns in your data and then use those patterns to predict future results or make other types of decisions under certain conditions. There are a number of algorithms in machine learning, which allow the machine to be based on a machine that can understand its current state and make the right choices. Machine learning works independently and decides for itself. There are two basic types of machine learning: supervised learning and non-supervised learning. This is called learning to learn. When models are applied to new datasets, learn from the training dataset, create models, and provide predictive results. Examples of learning to monitor, such as important trees, are the basis of the proposal. Unskilled education: Non-social education is a place where there is only input data and no matching variables. The main task of informal education is to write your own class labels. An unbiased survey algorithm can be used to detect correlations between statistics to determine whether the data produces grouping features. This group is known as clusters. Unsupervised learning can be also described as cluster analyses‖. K Means Clustering, KNN etc. are the examples of unsupervised learning.

Using target the value

```
Random over-sampling:
1    284315
0    284315
Name: Class, dtype: int64
```



A. Selected online dataset

Analyze this data by grouping it into multiple transactions. In addition, analysis of key components of PCA is performed to transform the data into two-

dimensional space. The dataset had a transaction with European card holders through credit cards in September 2013, which took two days, 492 out of 284,807 fake. Dataset is too inconsistent. In all trades, the positive class (fraud) is 0.172%. It only includes numeric input variables that are the result of PCA changes. Unfortunately, due to privacy concerns, we are unable to provide real data and a lot of background information about this data. The V1, V2, and V28 attributes are the main components acquired by the PCA, and the only features not changed by the PCA are 'time' and 'quantity'. The 'Time' function includes the last seconds between each transaction and the first transaction in the data set. The 'money' feature is a transaction count, and this feature can be used as an example of cost-sensitive learning.

C. Amount per transaction by classis

There are to type of transaction are made by the difference of two types

- Normal
- Fraud

Normal that is real transactions made by the card holder by its own

Fraud transactions that is made by theft or unknown person that is not made by card holder and thy are not now bot that transactions

C. Transaction VS Amount classis

Transaction time that are doing in that time not in pattern time means that is not made by organal card holder on the basis of time check real and fraudent transactions

This data set train on 201109 samples, validate on 56962 *samples*

D. Select algorithm for implementing

The literature reviews many algorithms on the detection of fraud. The questionnaire base includes 22, logistic regression, and more. The nearest neighbor is better than any other algorithm that detects fraud.. randomforest classfier ,logistic regression are used in this articals and implementing the fraud detection using these mdels

1) RANDOM FOREST CLASSIFIER

The Random Forest Classifier Classification. The algorithm organizes the theory of bias. The Indiana Ada algorithm Yang Sedrana Dan Sangat Kot.

The accuracy of Random Forest Classifier

```
[[113714     10]
 [    41    158]]
Accuracy: 99.96%
```

Basic Theorem: The Prayer Theorem Cummington Tragedy Cateca Peristua Lane MoundP (A/B) = (P (B/A) P (A)) / P (B)

Where,

P (A) – Priority of A P (B) – Priority of B

P (A/B) – Posteriori priority of B

Naïve Bayes algorithm is easy and fast. This algorithm need less training data and highly scalable.

2) LOGISTIC REGRESSION

Logistic regression using the functional approach to estimate the probability of a binomial response based on one or more changes (attributes). Find the best parameter for a nonlinear function called sigmoid.

The vector Z is the input data and the optimal coefficient is W. This coefficient is multiplied by each factor and then added to get a number that determines the classification of the target category. A sigmoid value greater than 0.5 is considered 1. 0 otherwise. The optimization method is used to train the classification and find the most suitable criteria. Gradient climbing (9) and modified stochastic gradient climbing compatibility methods experimented to evaluate for their performance in classifier.

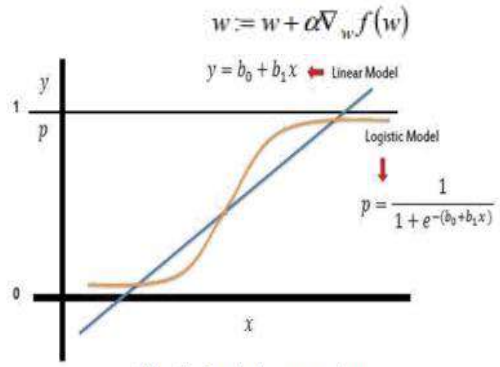


Fig. 1. Logistic regression

```
[[170557     29]
 [   105    194]]
Accuracy: 99.92%
```

This algorithm is similar to the linear regression algorithm. However, linear regression is used for predictive values and logistic regression is used to classify.

- Linear regression classified as
- Binomial – 2 Possible types (i.e. 0 or 1) only

results

After implementing the algorithm, the logistic regression gives the highest accuracy. The timeframe for logistic regression is quite high, but in this case accuracy is primarily considered as a result. Random forest classifier result 99.96%, logistic regression 99.92%, Beas, Logistic Relativity. Therefore, the results of the comparisons show that random forest classifier is more efficient than sophisticated bias and near-peripheral techniques. Then you can use logistic regression technology to detect credit cards.

REFERENCES

[25] J. Hand, G. Blunt, M.G. Kelly, and N.M. Adams, "Data Mining for Fun and Profit," Statistical Science, vol. 15, no. 2, pp. 111-131, 2000.

[26] "Statistics for General and On-Line Card Fraud," http://www.epaynews.com/statistics/fraud.html, Mar. 2007.

[27] S. Ghosh and D.L. Reilly, "Credit Card Fraud Detection with a Neural-Network," Proc. 27th Hawaii Int?l Conf. System Sciences:Information Systems: Decision Support and Knowledge-Based Systems, vol. 3, pp. 621-630, 1994.

[28] M. Syeda, Y.Q. Zhang, and Y. Pan, "Parallel Granular Networks for Fast Credit Card Fraud Detection," Proc. IEEE Int?l Conf. FuzzySystems, pp. 572-577, 2002.

[29] S.J. Stolfo, D.W. Fan, W. Lee, A.L. Prodromidis, and P.K. Chan, "Credit Card Fraud Detection Using Meta-Learning: Issues and InitialResults," Proc. AAAI Workshop AI Methods in Fraud and Risk Management, pp. 83-90, 1997.

[30] S.J. Stolfo, D.W. Fan, W. Lee, A. Prodromidis, and P.K. Chan, "Cost-Based Modeling for Fraud and Intrusion Detection: Results fromthe JAM Project," Proc. DARPA Information Survivability Conf. and Exposition, vol. 2, pp. 130-144, 2000

[31] Abdelhalim, A, and I Traore. "Identity Application Fraud Detection using Web". International Journal of Computer and Network Security1, no. 1 (October 2009): 31-44.

[32] E. D. Y. Sahin, "Detecting credit card fraud by decision trees," in

[33] Proceedings of the international multiconference of engineers and

[34] computer science, Hong Kong, 2011.

[35] 18. P. M.,. S. H. Geoffrey F.Miller, "Designing Neural networks using

[36] genetic algorithms," [Online]. Available:

[37] M. H. O. Ekrem duman, "Detecting credit card fraud by genetic

[38] M. K. A. N. Alireza Pouramirarsalani1, "Fraud detection in E-

[39] Banking by using the hybrid feature selection and evolutionary

[40] A. T. P. K. M. G. P. N. Priya Chougle, "Genetic K- Means Algorithm

[41] A. O. A. S. A. O. John o. Awoyemi, "credit cars fraud detection using

[42] science and information technologies, vol. 6, pp. 1724-1727, 2015.

[43] O. O. O. A.,. W. Stephen Fashoto, "Hybrid Methods for credit card

[44] fraud detection," Kampala International University, Kampala,

[45] Uganda, University of Abuja, Nigeria, Redeemer's University, Ede,

[46] Osun State, Nigeria, [Online]. Available:

[47] E. Aji M. Mubarek, "Multilayer perceptron neural network technique

[48] (Caelen, Multiple perspectives HMM-based feature engineering for credit, 2019)

МАСОФАВИЙ ТАЪЛИМ ТИЗИМЛАРИ ТАҲЛИЛИ

Жўраев М.Д., Андижон Давлат Университети, Ўзбекистон.

Annotation*: This article provides a comparative analysis of some of the existing education systems currently operating around the world. The analysis of software and technical requirements of the system, the activities of the internal facilities of the educational institution, the planning and control of the educational process, the database and the opportunities created for users were analyzed. As a result, the requirements for the introduction of a system that provides electronic services and facilitates the effective use of services in an interactive learning environment have been identified.*

Keywords*: ICT, education, education system, HEI, software, information management system, interactive learning environment, learning process, module, database, distance learning, academic sciences, electronic services, academic group.*

Аннотация*: в данной статье приводится сравнительный анализ некоторых существующих систем образования, действующих в настоящее время по всему миру. Анализируются программные и технические требования системы, деятельность внутренних средств образовательного учреждения, планирование и управление учебным процессом, база данных и возможности, созданные для пользователей. В результате были определены требования к внедрению системы, которая предоставляет электронные услуги и облегчает эффективное использование услуг в интерактивной учебной среде.*

Ключевые слова*: ИКТ, образование, система образования, вуз, программное обеспечение, система управления информацией, интерактивная учебная среда, учебный процесс, модуль, база данных, дистанционное обучение, академические науки, электронные услуги, академическая группа.*

Аннотация*: ушбу мақолада ҳозирги кунда дунё бўйлаб фаолият кўрсатиб турган мавжуд таълим тизимларининг баъзиларини қиёсий таҳлили амалага оширилган. Бунда тизимнинг дастурий техник талаблари, таълим муассаса ички объектлар фаолияти, ўқув жараёнини ражалаштириш ва назорат қилиш, маълумотлар базаси ҳамда фойдаланувчилар учун яратилган имкониятлар бўйича таҳлил қилинди. Натижада, интерактив таълим муҳитида элестрон хизматларни кўрсатувчи ва хизматлардан унумли фойдаланишга кўмаклашувчи тизимни жорий этиш учун талаблар аниқланди.*

Kalit so'zlar*: AKT, ta'lim, ta'lim tizimi, OTM, dasturiy ta'minot, axborotlarni boshqarish tizimi, interaktiv ta'lim muhiti, o'quv jarayoni, modul, ma'lumotlar bazasi, masofaviy ta'lim, o'quv fanlari, electron xizmatlar, akademik guruh.*

Ҳозирги АКТ ривожланган даврда таълим тизимини умумий концепциялар асосида сифатли ташкил этиш ва марказлашган шаклда бошқариш долзарб муаммолардан бири ҳисобланади. Таълим тизими фаолиятини бошқариш ва унинг автоматлаштирилган тизимларини яратиш ва жорий этиш ишлари жадал суръатлар билан амалга оширилмоқда. Қуйида ривожланган давлатлар таълим тизимида, жумладан ОТМларида жорий қилинган замонавий АБТлари ва мамлакатимиз таълим тизимида амалда бўлган ДТларнинг айримлари билан танишиб чиқайлик.

Ўқув жараёнини коплекс автоматлаштириш тизими - «GS-Ведомости». Бу тизим Россия Федерациясида ОАЖ «Гуру-Софт» компанияси томонидан ишлаб чиқилган замонавий ДТ бўлиб, унда ўрта махсус ва олий таълим муассасалари фаолиятини бошқариш билан бирга улар ўртасидаги алоқалар автоматлаштирилган. «GS-Ведомости» тизимининг мақсади таълим муассасалари асосий фаолиятларини шакллантириш ҳамда уларнинг кучли ва кучсиз жиҳатларини ҳар тарафлама ўрганиш ва таҳлил қилишдан иборат.

Тизим модулли архитектурага асосланган бўлиб, улар “Ўқувчилар контингенти”, “Абитуриент”, “Ходимлар хисоби”, “Ўқув режа ва ўқув юкламаларни шакллантириш”, “Кафедра”, “Дарс жадвали конструктори”, “Тест синови”, “Ҳужжат айланиши”, “Директ” (ходимларнинг иш вақтини назорат қилиш), “АИЖ раҳбари”, “Ҳужжатларни сақлаш”, “Кутубхона”, “Талабалар ётоқхонаси”, “Малака ошириш курси”, “Администратор”, “Маълумотларни импорт/экспорти”, “Ҳисоботлар конструктори”, “Шаблонларни таҳрирлаш” ва “Online” (тизим web-интерфейс шаклида бўлиб, унинг таркибига дарс жадвалларини ўзгартириш, давомат, талаба ўқув фаолиятлари боғлиқ маълумотлар тўлами, талабалар билан боғланишга доир маълумотлар, ахборотлар блоги, гуруҳларнинг алоҳида форумлари киради) модулларидан иборат.

Ҳозирги кунда «GS-Ведомости» тизими Россия ва МДҲ давлатларидаги йирик таълим муассасаларда жорий қилинган. Тизим 2009 йил май ойида ДТ сифати давлат стандартига жавоб бериши ҳақидаги сертификатга лойиқ деб топилган. Тизимнинг ишлаши учун МББТ FireBird, ОТ Windows ёки Unix/Linux, Pentium 3 ва ундан юқори параметрли компьютерлар зарур бўлади. Online модули глобал тармоқда, бошқа модуллар локал тармоқда ишлаш учун мўлжалланган.

«Галактика Управление Вузом» тизими. “Галактика” корпорацияси томонидан ишлаб чиқарилган таълим тизимини бошқариш тизими бўлиб, ўқув жараёнини бошқариш ва иқтисодий-хўжалик фаолияти масалаларини самарали ечишга мўлжалланган. Бу тизим қуйидагиларни ўз ичига олади:

- таълим муассасасида ўқув жараёни ва бўлимларни бошқариш ҳамда фаолиятларини режалаштириш;
- талабалар контингентини бошқариш ва ўзлаштиришини таҳлил қилиш;
- амалиётга тадбиқ этишдаги натижаларни таҳлил этиш
- молиялаштириш, шартнома ва логистикани бошқариш;
- аудитория ва талабалар турар жойлари жамғармасини бошқариш;

- ходимларни бошқариш (бунда штат бирликлари, табель ҳисоби);
- иш ҳақи ва стипендияларни ҳисоблаш;
- бухгалтерия, солиқлар ҳисоблари ва бошқалар.

Тизим Россия Федерациясидаги таълим тизими муассасаларидаги меъёрий ва ҳуқуқий қонун-қоидалар ва молиявий бошқарувига мослаб яратилган.

ОТМларни “Соатларини ҳисоблаш” тизими. Тизим локал тармоқда веб интерфейс кўринишда ишлаб, ОТМлари ўқитувчиларнинг ўқув юклама ҳажмларини ҳисоблашда қўлланилади.

ДТ ОТМда мутахассисликларга ўтиладиган ўқув фанларини кафедраларараро академик гуруҳларнинг бирлашмалари ва бўлинишларини ҳисобга олган ҳолда тақсимлайди. Кафедранинг умумий ўқув юкламасини ҳисоблаб чиқаради ва штат бирликларини аниқлаб беради. Кафедраларга штат бирликлари асосида тегишли ўқитувчилар киритилиб, уларга фан турлари тақсимланади. Мутахассисликлар ишчи ўқув режаси, кафедра ва ўқитувчиларга ўқувтюкламаларидан семестрлараро турли шаклдаги ҳисоботларни чиқариб беради.

“АИС ИНТЕГРАЛ” – ўқув жараёни ва таълим муассасасини комплекс бошқариш масаласини ечишда жаҳон талабларига жавоб берадиган ДТлардан бири. Дастур модуллар тузилишида яратилиб, модуллар ўзининг тегишли ахборот турларидан келиб чиқади (“Талабалар”, “Дарс жадвали”, “Баҳолаш”, “Факултатив”, “Кутубхона”, “Архив” ва бошқа модуллар). Ҳар бир модуль ўзига тегишли маълумотларни турли шаклларда тақдим этади. Дастурнинг асосий характерли томони ундаги модуллар бир-бирига боғлиқ эмас. Яъни фойдаланувчи ОТМ исталган модулларини қўшиб, кичик бошқарув тизимини яратиши мумкин. Дастур содда ва ундан исталган даражадаги фойдаланувчи ишлаши мумкин.

«АИС Интеграл» ДТ пуллик хизмат кўрсатиб, исталган операцион тизимида (Windows, Linux, MacOS, FreeBSD ва ҳ.к.), локал тармоқ орқали ишлайди ва маълумотлар базаси таълим муассасасининг алоҳида серверида сақланади.

Moodle тизими. Moodle – бу ўқитувчилар томонидан онлайн-курсларни яратиш учун махсус ишлаб чиқилган сайт таркибини бошқариш тизими (Content Management System - CMS) дир. У кўпинча e-learning тизимларни ўқитишни бошқариш тизимлари (Learning Management Systems - LMS) ёки виртуал ўқитиш воситаси (Vertual Learning Environments - VLE) деб аталиб, алоҳида онлайн-курслар каби, таълим веб-сайтларини ишлаб чиқувчи ускуна ҳисобланади.

Бу бепул тарқаладиган дастурий мажмуа ўзининг функционал имкониятлари, ўрганишдаги соддалиги ва ишлатишдаги қулайлиги билан электрон ўқитиш тизимларидан фойдаланувчиларнинг кўпгина талабларини қаноатлантиради.

Moodle масофавий ўқитиш жараёнини тўла қўллаб-қувватлаш учун кенг доирадаги имкониятларни беради. Унинг ёрдамида ўқув материалларини

турли усулларда бериш, билимларни текшириш ва ўзлаштиришни назорат қилиш мумкин. Ҳозирда Moodle тизимини дунёнинг йирик университетларида фойдаланилмоқда. Moodle ДТи GPL лицензияси остида очиқ жорий коди билан тарқалмоқда. Очиқ дастурий кодли Moodle га ўзгартиришлар киритиш, уни такомиллаштириш, модификациялаш учун http://www.moodle.org манзил бўйича мурожаат қилиши мумкин.

Moodle Unix, Linux, FreeBSD, Windows, Mac OS X, Netware операцион тизимларида ва PHP ишлатилиши мумкин бўлган ихтиёрий бошқа тизимларда ҳам модификацияларсиз қўланилиши мумкин. Маълумотлар MySQL ва PostgreSQL маълумотлар базасида сақланади. Шунингдек, уни маълумотлар базасини бошқаришнинг тижорат тизимлар учун ҳам ишлатилиши мумкин.

Бугунги кунда Moodle тизимидан Андижон давлат университети ҳамда Республикамиз бошқа ОТМларида талабаларнинг мустақил билим олишлари учун “Масофавий таълим тизими” шаклида фойдаланилмоқда.

«Talaba-S». Ўзбекистон Миллий университети томонидан «Talaba-S» локал тармоқда ишловчи дастури яратилган бўлиб, унда ўқув жараёни рейтинг тизимини, давомат, ўқув режалар каби қисмлари автоматлаштирилган.

«Unicos». Олий ва ўрта махсус таълим вазирилиги (ОЎМТВ) томонидан ишлаб чиқилган «Unicos» АБТ ҳам талабаларнинг давомати ва рейтинг тизимини ўз ичига олади.

Юқорида келтирилган ва бошқа таълим тизими фаолиятини автоматлаштиришга оид ДТларни таълим фаолиятини тўлиқ камраб олиши ва фойдаланувчиларга яратиладиган имкониятлари каби омиллар асосида таҳлил қилиш зарур. Таҳлил омиллари синфларга қуйидаги жадвалда кўрсатилгандек келтирайлик.

омиллар	A									B									C						D				E			
ДТ номи	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	c_1	c_2	c_3	c_4	c_5	c_6	d_1	d_2	d_3	d_4	e_1	e_2	e_3	e_4
Blackboard	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
GS-Ведомости	0	1	0	1	0	1	0	0	0	1	1	1	1	1	0	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0
ДекАр	0	1	0	0	0	1	0	0	1	0	1	1	1	1	0	0	0	1	1	1	1	1	1	0	1	0	0	1	0	1	0	0
Галактика Управление Вузом	0	1	0	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0
АСУ УЗ "МКР"	0	1	0	1	1	1	0	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	0		0	0	1	1	1	0	0
АИС ИНТЕГРАЛ	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	0	1	0	1	1	1	1	0	0
АКАДА ВУЗ	0	1	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	0	1	0	1	1	1	1	0	0
«Universys Web Server 5»	1	1	0	1	1	1	0	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
Naumen University	0	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1
Комплексное управления вузом «Аксиома»	0	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
ORBIT4Edu	0	1	0	1	1	1	0	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
1С:Университет	0	1	0	0	1	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
Система «Университет»	0	1	0	1	1	1	0	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
Tandem University	0	1	0	1	1	1	0	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	0	1	1	0	1	1	1	0	0
ИСУ ВУЗ ("ТИСБИ")	0	1	0	1	1	1	0	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	0	1	1	1	0	0
Moodle	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	1	1	1	0	0
SAMMER	0	1	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	1	0	0	1	0	0	0	1
Talaba-S	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	1	0	0	0	0
Unicos	0	1	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0
KARMAT	1	1	1	1	1	1	0	0	1	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	1	1	1	1	1	1	0	1

1- жадвал. Таълим ахборот тизимларининг қиёсий таҳлиллари.

Бу ерда жадвал қаторларда таълим муассаларда АБТ дастурий таъминотлар ҳамда устунларда таҳлил қилишда синфларга ажратилган омиллар берилган, $A=\{a_1,...,a_n\}$, $B=\{b_1,...,b_m\}$, $C=\{c_1,...,c_l\}$, $D=\{d_1,...,d_j\}$.

ДТларни таҳлил қилувчи омиллар қуйидагилардан иборат бўлсин:

А - дастурнинг техник, дастурий, тармоқ талаблари, яъни *А* = {a_1-очиқлилик, a_2-фойдаланувчига қулайли интерфейс; a_3-фойдаланиш текин; a_4-машинага боғлиқмаслиги; a_5-аниқ ва қисқа; a_6-мослашувчан ва тежамкорлик; a_7-кенгаювчан; a_8-янгиланиши, тармоқланганлиги; a_9-глобал тармоқда ишлаши};

В – таълим муассасаси ички объектлар фаолиятини қамраб олиши, яъни *В*={b_1-ректорат; b_2-факультет; b_3-кафедра; b_4-ўқув бўлими; b_5-ходимлар бўлими; b_6-ахборот-ресурс маркази; b_7-ахборот технологиялар маркази, b_8-бухгалтерия; b_9-илмий бўлим ва бошқа бўлимлар}

С – ўқув жараёнини режалаштириш ва назорат қилиш яъни *С* = {c_1-давомат; c_2-ўзлаштириш ва сессия; c_3-дарс жадвали; c_4-ўқув режа; c_5-юкламалрни бўлиштириш; c_6-виртуал ва мустақил таълим};

Д – маълумотлар базаси, яъни *Д* = {d_1-тизим доирасида объектлараро электрон ресурслар киритилиши, сақланиши ва узатилиши; d_2-маълумотлар базаси интеграцияланувчи; d_3-импорт/экспорт; d_4-ҳимояланиши};.

Е – фойдаланувчилар учун яратилган имкониятлар яъни *Е* = {e_1-рухсат этилган ходимларнинг ҳуқуқ даражалари аниқланиши; e_2-шахсий иш столлари; e_3-интерактив ишлаш имконияти; e_4-хабарлар алмашинуви}[].

Энди юқоридаги 1-иловадаги адвал таҳлилларига асосланиб қуйидаги умумлашган хулосаларни келтирамиз.

Биринчидан, хорижий ДТларини республикамиз таълим тизими талабларига мос равишда ўзгартириш учун маблағ талаб қилади.

Иккинчидан, аксарият ДТлар пуллик бўлганлигини келтириш мумкин.

Учинчидан, таълим муассасаларини бошқариш ва ташкил этиш модулларини мазкур ДТлар тўлиқ қопланмаган.

Тўртинчидан, ДТларни имкониятлари чегараланган ва уни кенгайтириш ёки ўзгартириш мумкин эмас.

Бешинчидан, аксарият дастурлар фақат локал тармоқда ишлайди.

Олтинчидан, мазкур ДТларда таълим тизимини бошқаришда профессор-ўқитувчилар, ишчи ходимлар ва талабаларга масофавий, кенг қамровли автоматлаштирилган иш жойилари яратилмаган.

Еттинчидан, бу ДТларда тизимдан фойдаланувчилар фаолиятини ҳамда ресурсларни бошқариш, баҳолаш, таҳлил қилиш, ягона тузилмали ахборот маконини ташкил қилиш моделлари йўқ.

Саккизинчидан, мазкур тизимларда ички ва ташқи фойдаланувчиларни табақалаштириш орқали талаб даражасидаги ЭХлар кўрсатилмайди.

Хулоса қилиб айтадиган бўлсак, юқорида келтирилган ДТлар ҳозирги вақтда ҳукуматимиз томонидан ЭХ кўрсатиш тизимларини ишлаб чиқариш учун қўйилган талабларга тўлиқ жавоб бера олмайди.

Кўрсатилган таҳлил натижаларини ҳисобга олган ҳолда, таълим тизимида ЭХларни кўрсатиш тизими қандай талаблар асосида барпо этилиши лозимлигини қараб чиқамиз. Мамлакатимиз таълим тизими бошқарувини автоматлаштиришни такомиллаштириш, бу ИТМда ЭХларни кўрсатувчи ва хизматларни унумли фойдаланишга кўмаклашувчи тизимни жорий этиш учун унда қуйидагиларни ўзида мужассамлаштириши зарур:

1. таълим муассасаси тузилмасини аниқ белгилаб олиш ва унинг ягоналикни таъминловчи ИТМ тузилмасини аниқ моделларга таянган бўлиши;
2. таълим муассасасига алоқадор барча объект ва шахсларнинг назардан четда қолмаслиги;
3. модулли архитектурага таяниб масофавий бошқаришда АИЖлари ва таълим беришнинг амалга оширилиши;
4. фойдаланувчиларнинг тизимга таъсирларини назорат қилиниши;
5. ресурсларни сақлаш, хизматларни кенгайтириш, бошқариш, баҳолаш ва рейтингини аниқлаш моделларини яратиш ва таҳлил қилиниши;
6. таълимда ЭХлар кўрсатиш талабларининг асоси ишлаб чиқарилиши;
7. реал вақт оралиғида эҳтиёжлар асосида ҳисоботларни шакллантириш, қарорлар қабул қилишга таклифлар бериш имконияти мавжудлиги;
8. Турли ахборот тизимлар билан интеграцияланиши;
9. Қўшимча хизматларни ўзида қўллаб-қувватлаши;
10. Таълимнинг барча босқич ва қатламлари билан боғлиқлиги;
11. Маълумотларни химоялаши;
12. Тизимнинг адаптивлиги.

Демак, таълим тизимида кенг қамровли ИТМни яратиш ва унда аниқ мақсадли эҳтиёж-мурожаатлар бўйича электрон хизматларни кўрсатувчи тизимини ишлаш, хизматлардан самарали фойдаланиш механизмларини жорий этиш ҳозирги куннинг муҳим масаласи ҳисобланади.

Адабиётлар

1. *Андреев А.А., Солдаткин В.И. Дистанционное обучение: сущность, технология, организация. – М.: Издательство МЭСИ, 1999. – 196 с.*
2. *Барсегян А.А., Куприянов М.С., Степаненко В.В., Методы и модели анализа данных: OLAP и Data Mining. БХВ- Петербург, 2004.–336 с.*
3. *Гура В.В., Дикарев С.Б. Система проектирования электронных образовательных ресурсов. Ростов-на-Дону: ООО "ЦВВР", 2003. 125с.*
4. *Кулагин В.П., Найханов В.В., Овезов Б.Б., Информационные технологии в сфере образования. - М.: Янус-К, 2004. - 248 с.*
5. *Ҳамдамов Р., Бегимкулов У., Тайлоқов Н., Таълимда ахборот технологиялари. Олий таълим муассасалари учун. -Т.: 2010, 120 б.*

LUNG CANCER PREDICTION FOR CLASSIFICATION ALGORITHM BY USING WEKA

Nudrat Batool, Department of Computer Science, Ripha International University, Lahore, Pakistan.

Abstract. Lung cancer is a leading cause of cancer-related death in men and women, and is increasing worldwide. Lung cancer is the uncontrolled growth of abnormal cells that begin in one or both lungs. Early detection of cancer is not an easy process, but if it is detected, it is curable. We analyzed lung cancer prognosis using a classification algorithm such as the Nave Bayes, Bayesian network, logistic regression and the J48 algorithm. Initially, data from 100 cancer and non-cancer patients were collected, pre-processed, and analyzed using the classification method for predicting lung cancer. The dataset contains 100 instances and 25 attributes. The main purpose of this study is to provide users with early warning and performance analysis of classification methods.

Keywords: Data Mining, Lung Cancer Prediction, Classification, Naive Bayes, Bayesian Network, Logistic regression, and J48.

1. INTRODUCTION

Data mining is an important step in finding knowledge from large data sets. Data mining has found its significant grip on every sector, including health care. Data processing plays an important role in obtaining hidden information on the medical data site. Mining process rather than data analysis involves classification, clustering, association rule mining and prediction. Lung cancer is the most common cause of cancer death worldwide. If the original Lung cancer had spread, a person may experience symptoms elsewhere in the body. Lung cancer diagnosis is used to predict risk. The main objective of this study was to quantify the risk of lung cancer using the WEKA tool.

Lung cancer is the second most common cancer in both men and women and is the leading cause of cancer-related deaths worldwide. Lung cancer is basically classified as non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC).

Non-small cell lung cancer (NSCLC) is the main form. Squamous cell carcinoma (SCC) and adenocarcinoma (ADC) are the two main forms of NSCLC, accounting for 80% of NSCLC cases.

The number of genes in a human cell varies from 20,000 to 25,000 and are hereditary elements of life [8]. The central dogma of life describes gene expression as a process in which a deoxyribonucleic acid (DNA) sequence produces ribonucleic acid (RNA), which then converts a protein sequence, as shown in Figure 1. Figure 1. Central dogma of life Due to various biological and environmental factors, mutations in genes disrupt expression levels, causing irregular levels of ribonucleic acid (RNA) and proteins.

MicroRNAs (a class of short endogenous RNA molecules) that play an important role in regulating gene expression are also found at modified levels. These results lead to uncontrolled cell growth, which leads to the development of cancer [9]. In recent years, efforts have been made to recognize biomarkers (biomolecules that can distinguish diseases). From the normal state) at different genetic and proteomic levels. The evolution of high throughput methods in genomics such as microarrays and mass spectroscopy in proteomics has progressed in cancer research.

They provide a wealth of data that can be used to understand the development and progression of the disease. However, identifying the genotype-phenotype correlation remains a complex problem. To solve this problem, we used an improved technique to create the decision tree (a modified version of the J48 classifier of the WEKA tool).

The classifier used in our approach can accurately classify genes to differentiate cancer phenotypes. Chapter II provides an overview of similar studies in the literature. Part III explains the uniqueness of our training dataset and provides a detailed view of the data mining methodology applied for estimation. In Chapter IV we discuss the results of the decision tree and the last, Section V gives the result of our study and summarizes the future perspectives of this study.

2. RELATED WORKS

YongqianQiang, YouminGuo, Xue Li, Qiuping Wang, Hao Chen, &DuwuCuic [6] conducted clinical and imaging diagnostic rules of peripheral lung cancer by data mining technique, and to explore new ideas in the diagnosis of peripheral lung cancer, and to obtain early-stage technology and knowledge support of computer-aided detecting (CAD). The data were imported into the database after the standardization of the clinical and CT findings attributes were identified. The diagnosis rules for peripheral lung cancer with three data mining technology is same as clinical diagnostic rules, and these rules also can be used to build the knowledge base of expert system.

The demonstrated the potential values of data mining technology in clinical imaging diagnosis and differential diagnosis. Krishnaiah V,Narsimha G, Subhash Chandra N [7]proposed to a model for nearly detection and correct diagnosis of the disease which will help the doctor in saving the life of the patient. Using generic lung cancer symptoms such as age, sex, wheezing, shortness of breath, Pain in shoulder, chest, arm, it can predict the likelihood of patients getting a lung cancer disease. PrashantNaresh [8] applied a pattern prediction tools for a lung cancer prediction system, lung cancer risk prediction system should prove helpful in detection of a person's predisposition for lung cancer. The early prediction of lung cancer should play a pivotal role in the diagnosis process and for an effective preventive strategy. Thangaraju P, Karthikeyan T, Barkavi G [9] conducted smoking is the biggest risk factor of lung cancer.

The more years and larger number of cigarettes smoked the greater the risk of developing lung cancer. The average age of someone diagnosed with lung

cancer is 65 to 70 years old, but people who are younger can develop lung cancer. Young adults who have never smoked also can develop lung cancer. Ravi Kumar G.,Ramachandra.A, Nagamani.K,[10] conducted breast cancer is one of the major causes of death in women when compared to all other cancers.

Breast cancer has become the most hazardous types of cancer among women in the world. Early detection of breast cancer is essential in reducing life losses. The comparison among the different data mining classifiers on the database of breast cancer Wisconsin BreastCancer (WBC), by using classification accuracy. The aims to be establish an accurate classification model for breast cancer prediction, in order to make full use of the invaluable information in clinical data, especially which is usually ignored by most of the existing methods when they aimfor high prediction accuracies. It is compare six classification techniques in WEKA software and comparison results that Support Vector Machine (SVM) has higher prediction accuracy than those methods. Different methods for breast cancer detection are explored and their accuracies are compared.

With these results, The SVM are more suitable in handling the classification problem of breast cancer prediction, and use of approaches in similar classification problems. Tapas RanjanBaitharu, Subhendu Kumar Pani [11] Conducted the most important cause of death for both men and women is the cancer lung cancer is a disease of uncontrolled cell growth in tissues of the lung. Data classification is an important task in KDD (knowledge discovery in databases) process. It has several potential applications. The performance of classifiers is strongly dependent on the data set used for learning. It leads to better performance of the classification models in terms of their predictive or descriptive accuracy, diminishing of computing time needed to build models as they learn faster, and better understanding of the models.

A comparative analysis of data classification accuracy using lung cancer data in different scenarios is presented. The predictive performances of popular classifiers are compared quantitatively.

2.1 Methodology used of proposed work

Based on previous studies on the classification of cancer biomarkers, we analyzed these data sets and can make improvements by studying combinations of cancer subtypes. And cancer classification.

The structure includes four main stages: data collection, preparation of data collection, data analysis, evaluation and implementation. The detailed process is described below.

Research Group	Cancer types	Data set Type	Methodology applied
Junkui Chen et al., [14]	Colon Cancer	Gene expression data	Signal-to-Noise ratio, Bhattacharyya distance, Support vector machine
Chun-Hui Wu et al., [15]	Prostate cancer	PSA blood test	Decision tree
Lihua Lia et al., [16]	Ovarian cancer	Proteomic data	Receiver operator characteristic
Shelly Gupta et al., [11]	Breast Cancer	Review article	Review of techniques
Shital Shah et al., [12]	Ovary, Prostate and Lung cancer	Gene expression data	Decision tree, Support vector machine
Jinn-Yi Yeh et al., [13]	Lymphoma, Lung and colon cancer	Gene expression data	Genetic algorithm

3. DATA MINING TECHNIQUE

Data mining is the process of automatically collecting large volumes of data, aiming to discover hidden patterns and analyze the relationships between different types of data to create predictive models. Classification techniques and prediction are two forms of data analysis that can be used to extract models describing important data classes or to predict future data trends. Such analysis will help us to better understand the data.

4. DATASET DESCRIPTION

The dataset used in this study is highly accurate and accurate in order to improve the predictive accuracy of data mining methods. The diagnostic attributes of the disease must be handled efficiently to obtain optimal results from the data mining process. Age, sex, air pollution, alcohol use, dust allergies, occupational hazards, genetic risk, chronic lung disease, balanced diet, obesity, smoking, inactive smoker, chest pain, blood cough,

Fatigue, weight loss, shortness of breath, shortness of breath, difficulty swallowing, finger nailing, frequent cold, dry cough, and snoring are considered to predict lung cancer. WEKA implements algorithms for classifications such as data pre-processing, feature reduction, Now Bayes, Bayesian network, logistic regression and J48. The efficacy of lung cancer mechanisms is analyzed using visualization tools.

Table 4.1 Lung cancer factors

Factors

Factors
Age
Gender
Air pollution
Alcohol use
Dust allergy
Occupational hazards

Genetic risk
Chorionic lung disease
Balanced diet
Obesity
Smoking
Passive smoker
Chest pain
Coughing of blood
Fatigue
Weight loss
Shortness of breath
Wheezing
Swallowing Difficulty
Clubbing of Fingers nails
Frequent Cold
Dry Cough
Snoring

The table 4.1 shows that dataset description and these factors have the equivalent numeric value based on the symptoms.

5. PERFORMANCE ANALYSIS

This study mainly uses classification algorithms such as the Naive Bayes, Bayesian network, logistic regression and the J48 algorithm to predict lung cancer from a given set of data, and the proposed methods in the WEKA tool in the Lung Cancer Database. And performance is measured.

All the attributes in data set are displayed in row format in the left half and on the right side is bar graphs represent the distributions of the different attributes for data mining. Class is predicting the risk attributes for 4 distinct in Label Low, Medium, and High.

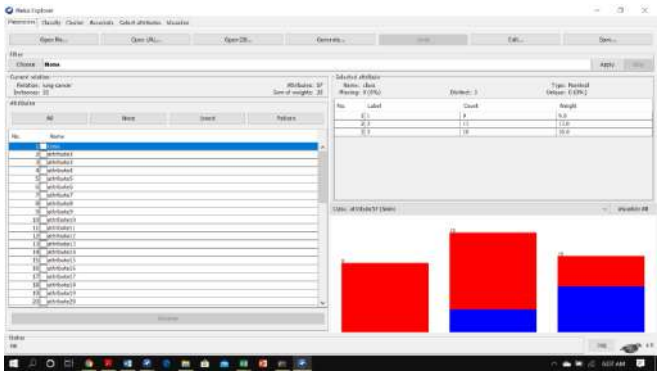


Figure 5.2 Lung cancer risk prediction

The Figure 5.2 shows that risk for Low, Medium and High Level of Lung cancer disease. It's observed that to predict 30 patients in low level risk, 33 patients in medium level risk, 37 patients in high level risk.

The Figure 5.3 shows that the Naive Bayes algorithm builds the prediction in 0.01 seconds and the 100 instances are correctly classified.

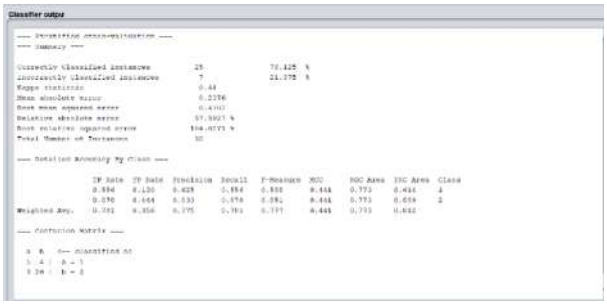


Figure 5.3 Naive Bayes

The Figure 5.4 shows that the Bayesian network algorithm builds the prediction in 0.03 seconds, and all the instances are correctly classified.

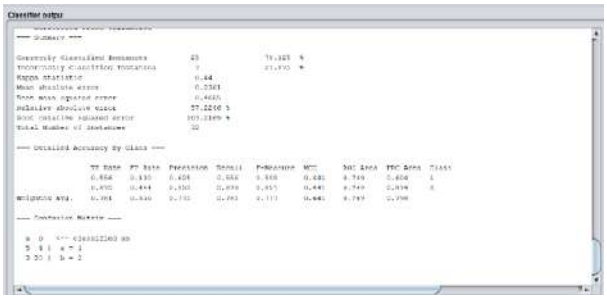


Figure 5.4 Bayesian Network

Confusion matrix is shows that predict 30 is low risk level of patient, 33 is medium risk level of patient, 37 is high risk level of patient is predicting risk level

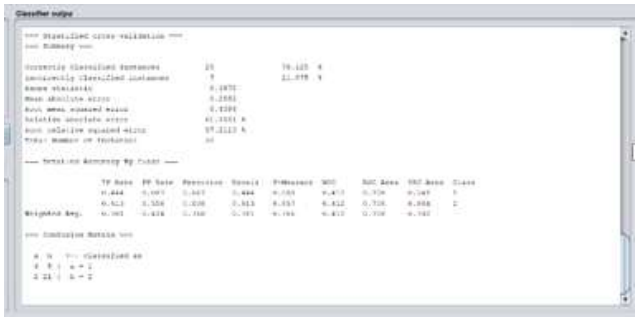


Figure 5.5 J48 algorithm

The Figure 5.5 is shows that a J48 algorithm builds the prediction in 0.06 seconds, and all the instances are correctly classified.

The Confusion Matrix for the classification algorithms such as Naive Bayes, Bayesian Network, Logistic regression and J48 can be given as follows based on the execution of the algorithm using WEKA tool. The table 5.1 explains about the confusion 4x4 matrixes of the Naive Bayes, Bayesian Network, Logistic regression and J48.

Table 5.1 Confusion Matrix for J48

A	B	C	Classified

30	0	0	a = low
0	33	0	b = medium
0	0	37	c = high
0	30	0	d = low

The proposed method is used for predicting the lung cancer disease risk level using different algorithms such as Bayesian Network, Naïve Bayes, Logistic regression and J48 are applied on lung cancer disease data set in the WEKA tool. The algorithm performance can be obtained based on the time taken to build model.

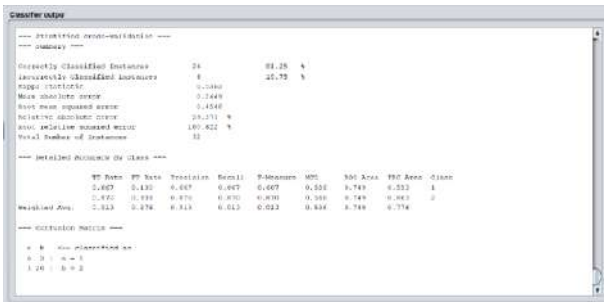
The Table 5.2 is shows that classification algorithm is perform to predicting the lung cancer disease from dataset instances and attributes the proposed model contains three different types of algorithms such as Naïve Bayes, Bayesian Network, Logistic regression and J48 are applied on lung cancer disease dataset in the WEKA tool.

Table 5.2 Time taken by the algorithms

Algorithm	**Time Taken to build the model (seconds)**
Bayesian Network	0.03 seconds
Naïve Bayes	0.01 seconds
J48	0.06 seconds
Logistic regression	0.03 seconds

Figure 5.6 Comparative analysis

The Figure 5.6 shows that details about the time taken by the algorithms (Naive Bayes Bayesian Network, Logistic regression and J48) to build model in WEKA tool. The graph shows that Naïve Bayes algorithm is the best performance algorithm based on the time.



The Figure 5.7 is shows that a logistic regression algorithm builds the prediction in 0.03 seconds, and all the instances are correctly classified.

The Figure 5.7 shows that details about the time taken by the algorithms (Naive Bayes Bayesian Network, logistic regression, and J48) to build model in

WEKA tool. The graph shows that Naïve Bayes algorithm is the best performance algorithm based on the time.

6. RESULTS AND DISCUSSION

In this study, we identified the expression of different biomarkers in two histological subtypes of non-small cell lung cancer. This method of identifying co-expressed and differentially expressed markers will help us understand the main differences in the molecular levels involved in the development of various cancer phenotypes. These results will help create a matrix differential diagnostic platform for SCC and ADC.

Most studies that identify cancer biomarkers report that hundreds of genes or miRNAs are considered differentially expressed. However, they forgot to regulate most of them for various forms of cancer. Therefore, these results cannot establish a model for the differential diagnosis of cancer. We designed a classification model to be able to identify the characteristic genetic mutations or the expression changes of the different types of cancer. Therefore, we tried to combine the knowledge of validated biomarkers and use the enhanced fleet cabin and J48 to predict the category of cancer phenotype. This classification methodology reveals an approach for differential cancer diagnosis based on knowledge of rejected / unregulated biomolecules. Treatment options for patients with NSCLC do not differ between these two histological subtypes. So we searched further, studying the identified genes differentially expressed for drug development studies to determine the appropriate therapeutic target for both types of cancer.

7. CONCLUSION AND FUTURE WORK

Data mining in healthcare management is not similar to other fields because the data here is heterogeneous in nature and many ethical, legal and social restrictions apply to private medical information. The experiment was performed using WEKA tools and different data mining classification techniques, and found that the Naive Bayes algorithm performs better than other classification algorithms (such as Bayes, Logistic Regression, and J48). The long-term cancer prediction system can be improved and expanded. It can also include other data extraction methods, such as time series, grouping rules, and associations. Continuous data can also be used.

REFERENCE

1. Ada[1], R. K. (2013). Early detection and prediction of lung cancer survival using neural network classifier.

2. Ayyadurai, P., Kiruthiga, P., Valarmati, S., & Amritha, S. (2013). Respiratory cancerous cells detection using TRISS model and association rule mining. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, *2*(3), 1030-1035.

3. Baitharu, T. R., & Pani, S. K. (2015). A comparative study of data mining classification techniques using lung cancer data. *Int J Comput Trends Technol*, *22*(2), 91-95.

4. Kaur, H., & Wasan, S. K. (2006). Empirical study on applications of data mining techniques in healthcare. *Journal of Computer science*, *2*(2), 194-200.

5. Kharya, S. (2012). Using data mining techniques for diagnosis and prognosis of cancer disease. *arXiv preprint arXiv:1205.1923*.

6. Krishnaiah, V., Narsimha, G., & Chandra, D. N. S. (2013). Diagnosis of lung cancer prediction system using data mining classification techniques. *International Journal of Computer Science and Information Technologies*, *4*(1), 39-45.

7. Kumar, G. R., Ramachandra, G. A., & Nagamani, K. (2013). An efficient prediction of breast cancer data using data mining techniques. *International Journal of Innovations in Engineering and Technology (IJIET)*, *2*(4), 139.

8. Naresh, P., & Shettar, D. R. (2014). Early detection of lung cancer using neural network techniques. *Prashant Naresh Int. Journal of Engineering Researchand Applications*, *4*(8), 78-83.

9. Park, S. M., Lim, M. K., Shin, S. A., & Yun, Y. H. (2006). Impact of prediagnosis smoking, alcohol, obesity, and insulin resistance on survival in male cancer patients: National Health Insurance Corporation Study. *Journal of clinical Oncology*, *24*(31), 5017-5024.

10. Priyanga, A., & Prakasam, S. (2013). Effectiveness of data mining-based cancer prediction system (DMBCPS). *International Journal of Computer Applications*, *83*(10).

11. Qiang, Y., Guo, Y., Li, X., Wang, Q., Chen, H., & Cuic, D. (2007). The diagnostic rules of peripheral lung cancer preliminary study based on data mining technique. *Journal of Nanjing medical university*, *21*(3), 190-195.

12. Rani, V. M. K., & Mary, M. S. (2014). Predicting Progression Of Primary Stage Cancer To Secondary Stage Using Decision Tree Algorithm. *gene*, *3*(6).

13. Sowmiya, T., Gopi, M., Begin, M., & Robinson, L. T. (2014). Optimization of lung cancer using modern data mining techniques. *International Journal of Engineering Research*, *3*(5), 309-314.

14. Thangaraju, P., Barkavi, G., & Karthikeyan, T. (2014). Mining lung cancer data for smokers and non-smokers by using data mining techniques. *International Journal of Advanced Research in Computer and Communication Engineering*, *3*(7), 7622-7626.

15. Vijaya, G., Suhasini, A., & Priya, R. (2014). Automatic Detection of Lung Cancer in CT Images. *IJRET: International Journal of Research in Engineering and Technology*, *3*(7), 182-186.

IMAGE SEGMENTATION AND CLASSIFICATION

Mehwish Shabbir. MSCS Student , Dept. of Computing &IT Riphah International University of Lahore, Pakistan.

Abstract. *Segmentation and marking remains the most fragile advance in numerous clinical vision applications. Research outlines a methodology dependent on the watershed, and which is controllable through adjustment of their parameters. Two of these modules are introduced: the lung malignant growth detection, for the segmentation of disease areas from CT images, a watershed calculation for image segmentation and brain tumor detection from MRI images. Different GLCM includes alongside some facts that are utilized for characterization utilizing the Neural system and Support Vector Machine (SVM).*

Keywords*: Brain tumor detection, Segmentation, Lung cancer , MRI images, CT images, GLCM, Neural network*

1. INTRODUCTION

Recently, image preparation procedures have been widely used in certain medical areas to improve the image in the previous detection and treatment phases, in particular in various malignant tumors, for example lung diseases, etc. The quality and precision of the images are the main components of this exploration. In this review, the main distinguishing features for accurate image statistics are the measurable characteristics of the pixels. The segmentation of medical images is the programmed detection or automatic loading of the limits in a 2D or 3D image. A major problem with imaging segmentation is the high inconsistency in imaging. A wide range of modalities (radiography, computed tomography, magnetic resonance imaging, microscopy, PET, SPECT, endoscopy, OCT and a few others) are used to create medical images.

The research covers two uses of medical segmentation, in particular the detection of malignant lung growth and the segmentation of mental cancer. In the research, an improved technique for the artificial neural network (ANN) and the SVM classifier model are proposed to dissect the areas of the lung removed from CT images of the human chest. A combination of different factual features is used to detect projections of the pulmonary district.

This review presents a brain tumor segmentation technique for magnetic resonance imaging, the logic of which is to efficiently divide tumor masses into MRI images with reproducible results. As a result, our tube can dissect tumor masses, which means that manual determination of a starting area is not necessary, unlike segmentation plans with automatic loading. We use surface examination strategies to eliminate useful data that allows segregation between different types of tumor tissue.

2. RELATED WORK

i. Lung Cancer Segmentation:

Over the years, several PCs have been useful for diagnosis (CAD) and lung nutrition patterns have been developed and tested in the area of this scan.Garkan et al. [1] proposed a multivariate CAD model that uses linear regression and linear discriminant analysis (LDA) to divide the distal lung segment using k-k for the detection of lung and lung cancer.). Each reports 84% effectiveness and a normal 1.74 false dreams. 4 McCulloch [2] also proposed a multidimensional CAD structure in which each phase carries specific anatomical data and lung structure. Its structure produces 8.3 false facts per element. Ge et al. [3] The 3D image of the lens needed to detect lactation was removed. Within these high-impact products are 80% effective, with about 0.77 counterfeit products per area. When you find the shoulder of a figure or outline, it shows a specific x, y, z axis. In this regard, Lee et al. [4] He proposed a slide method for specifying the moving window. They reported 100% effect and 1.4 false positives per cut and attempted 20 cuts. Many PC works have been completed on CNN due to advances in image-format broad-spectrum (CNN) neural networks created by Alex Krizhevsky using AlexNet [5]. Since then, the strategy has always been called "deep learning." They have also been extraordinarily successful in the application of medical imaging technology, as demonstrated by Kayalbay in his work. Thus, the CAD method for detecting posterior lung tumors uses the CNN model. Yang et al., Because CNN always needs more guidance to successfully broadcast it. [7] We investigated whether the dissemination of information improved the implementation of 2 CNNs C. Its basic decision is that the enrichment of information improves the performance of CNN TV. An ongoing study by Hamidian et al. [8] CNN 3D was investigated and has an 80% impact and 22.4 false positives in each test. CT filters are 3D in size and are often a problem. Deep vertual networks also appear to be suitable for 3D delivery, as described by Milletari et al. [9]. It is also effective for 2D to 3D, which is an enhanced effective looga by adding others. [10]

ii. Brain Tumor Segmentation:

The author shows that the segmentation of brain tumors can be achieved by simultaneously separating the discharge basins by collecting image pixels based on the intensity of their respect, dealing with problems with high resolution, confusion or varying degrees of intensity. [eleven]. The seed should be selected due to the local segmentation-based client program. Therefore, more customer selection is required in seed selection. Sowing between cancer cells is an advantage. The intensity of the local development strategy may be the cause of an imperfect problem. In addition, it should give accurate results in all images [12]. The identification of the lane provides careful, clear boundaries or extremities. However, its weight is that if the difference in intensity between normal and cancer cells is small, it cannot be differentiated. As it stands, it gives the volume a certain size when detected. The Edge Discovery strategy works very well in high contrast images and requires room in low contrast images due to the low angle size. Fluid C refers to the same method used to dissect and dissect irregular tissues by cutting the magnetic resonance imaging [14]. This strategy has many functions and numerical equations, and therefore has a higher computational prediction, such as

the average number of moves, while FCM is appropriate when the size of the groups is hidden from the previous one [15]. In the index control basin, the internal and external indicators required to divide objects are processed based on the relevant pixel elements. Internal indicators are related to pixel elements associated with conspiracy objects. External indicators are relatively parts of pixels that are related to the base of objects. These markers are used for anatomical purposes to separate the tumor bundle. Indicator-based partitioning requires a natural selection of indicators, so that a planned selection of indicators is no longer required.

3. WATERSHED TRANSFORM FOR IMAGE SEGMENTATION

Segmentation is used to divide the image into several areas or small elements. It has numerous clinical applications for segmentation. There are different strategies for image segmentation. The studies use watershed segmentation methods with threshold control and marker. Threshold is the best tool due to image segmentation. It is used to change the grayscale image. The watershed is well known in areas such as biomedical and clinical imaging as well as CV vision [17]. In the topography of the basin. In the basin lines, limits are defined that distinguish the image areas. A change in the catchment basin records catchment basins and ridge lines (also known as water separation lines), where catchment basins are compared to image areas and outlet lines that define the boundaries of the area [18] . Watershed segmentation illustrates a host of ideas from three procedures, for example, boundary-based segmentation, boundaries, and locations. Basin calculations, depending on the change in the basin, have mainly two classes. Numerous calculations were proposed in two classes, but river basin calculations based on related parts generally show excellent performance compared to all others. Follow the rainwater harvesting approach. This provides excellent segmentation results and is consistent with models with less computational unpredictability for team execution.

4. PROPOSED METHODOLOGY

The principal research consists of the following four steps of the form of a proposition, in the manner of what is written below

I. The sound of the MRI or CT image of the Image preprocessing filtering by means of the removal of the intermediate there is;

II. The multiple CT lung cancer, lung or brain MRI images or a brain tumor images using the fast change watershed;

III. extraction of features of divided brain tumors using a variety of statistical and persistent pulmonary cancer GLCM features.

IV. SVM and NN classify lung cancer, brain tumor or a picture.

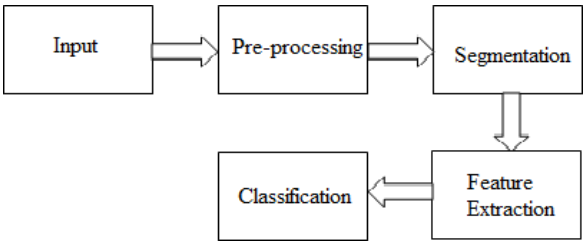


Figure 1.Proposed System Diagram for Image Processing

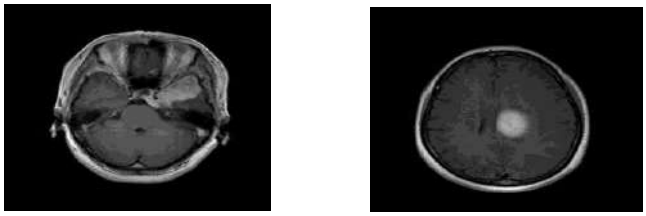
A. Input images

The CT scan images used for the intervention were taken from the emergency centers. The data collection for this image included CT scan images of the tumor without tumor. Figure 2 shows part of the CT scan images with one tumor and no tumors.

(a) (b)

Figure 2. CT scan image (a) Lung without tumor (b) Lung with tumor

This brain tumor data set, including 3064 T1 weighted complexity, was used in 233 patients with three types of mental tumors: meningioma (708 cuts), glioma (1426 cuts), and pituitary tumor (930 cuts). Enhanced image. Figure 3 shows an image of a brain tumor from the dataset.



(a) (b)

Figure 3. MRI Scan Image with Brain Tumor

B. Preprocessing

The central channel is a nonlinear isolation strategy, often used for vocalizing or expressing. Reducing this type of violence is a set of activities to improve the consequences of preparing (for example, at the corner of the picture). Moderate release is generally used in computer video management in light of the fact that, in certain situations, gelatin removes violence (and also follows the story below), it also has applications that show preparation. The central idea of the central passage is to cross the sign at the entrance and stop at any point of entry in the nearby street. The neighboring case is called "window", which overlaps, crosses, the whole sign. In 1D symbols, the most visible window is only the beginning of the first and subsequent sections, while 2D (or higher order) features are visible, for example images, more complex window designs (e.g., "box" ") or" cross "design). Remember that in the event that a window has unexpected numbers, at that point the interval can be easily identified: it is only the stimulus center after all the components of the window have been carefully prepared. In a large number of sections, there is more than one function that can be traced by Matlab B = medfilt2 (A) performs an average classification of the image on one of the two dimensions. Each pixel rendering consists of intermediate stimuli in the neighborhood 3 by 3 around the comparison of the pixels in the information display. medfilt2 evokes the image with 0 edges, so the central quality focused on half the width of the area ([m n] / 2) at the edges may seem interchangeable.

C. Image Segmentation

The segmentation with the watershed method works excellently if the elements of the front and cellar area are different or stamped. This is a basic, insightful technique and is quick. The water portal method produces seeds that indicate the proximity of objects or the creation of certain zones. Marking areas are then defined as the minimum dimensions of the neighborhood within the topological surface and an assessment of the threshold is made. The advantage of the portal is that it offers an excellent response to certain input information. When the portal image is received, diseases appear in the documentary image. Fig. 4 shows a saw image of the malignant growth.

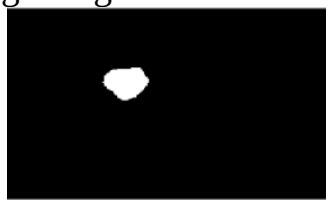


Figure 4.Cancer Spotted Image

For the undeniably obvious site of cancer, the red imprint shading is used to distinguish the absolute No. cancer present in a given double or informational picture. Figure 5 shows cancer cells highlighted by a red stripe.

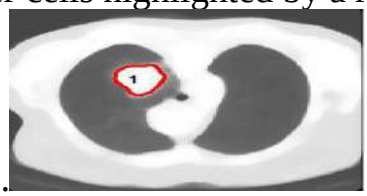


Figure 5. Cancer Outlined By Red Color In Input Image

Likewise watershed calculation is utilized for mind tumor division. Figure 6 shows the mind tumor division utilizing the watershed change.

Figure 6. Brain Tumor Segmentation

D. Feature Extraction

The characteristic is the important information extracted from an image which provides a more detailed understanding of the image. of the image.

a. GLCM Features

The origin of the gray line is a second-order statistical model. GLCM is a dust-conditioned co-growth matrix (GLCM), also known as a dust-dependent spatial matrix. GLCM operations define the image of an image by determining how many times the pixel and its values are calculated and in a network corruption situation that occurs in an image, creating a GLCM and then compiling the statistics. of this matrix. The statistics calculated from the GLCM values are as follows: The matlab function GLCM = greycomatrix (I) creates a dust-conditioned co-growth matrix (GLCM) from the images.

The matlab graycomatrix function calculates GLCM from the image measurement. In general, if they are a binary image, the gray matrix resizes the image in white high. If I were a strong figure, the Graycomatrix would be calculated in eight positions of dust. stats = greycoprops (GLCM, property) calculates the statistic defined in the properties of the recurrence matrix of empty states GLCM. GLCM is an m-e-n-e-p line of allele status indicators. If the GLCM is a derivative of the GLCM, the statistics are a set of statistics for each GLCM. Grycoprops characterizes the hair position co-growth matrix (GLCM) so that the sum of its elements is equal to 1.Each of the elements (r, c) in the normalized GLCM is the joint probability of pixel pairs anda descriptive damage ratio has the gray values r and c in the image. greycoprops uses augmented GLCM to determine properties.

Property	Description	Formula
'Contrast'	It returns a measure of the intensity contrast between a pixel and its neighbor throughout the image. Range = [0 (size (GLCM, 1) - 1) ^ 2] The contrast is 0 for a stable image.	$\sum_{i,j} \lvert i-j \rvert^2 p(i,j)$
'Correlation '	Returns a measure of the correlation of a pixel with its neighbor over the entire image. Range = [-1 1] The correlation is 1 or -1 for a perfectly positive or negatively correlated image. The correlation is NaN for a constant image.	$\sum_{i,j} \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma_i \sigma_j}$
'Energy'	Returns a summation of the correlation of a pixel with its neighbor over the entire image.Range = [-1 1] The correlation is 1 or -1 for a perfectly positive or negatively correlated image. The correlation is NaN for a constant image.	$\sum_{i,j} p(i,j)^2$
'Homogene ity'	GLCCM returns a value that approximates the object. Directly on the GLCM diagram. Range = [0 1] GLCM jacket for rehab 1	$\sum_{i,j} \frac{p(i,j)}{1+\lvert i-j \rvert}$

Table 1: GLCM properties Along with GLCM following features are extracted

b. Mean: M = mean (A) returns the average value of elements over different dimensions of the array.

c. Standard Deviation: s = std (X), where X is a vector, using the above (1), restores the standard deviation. As long as X is free and indistinguishable. Std (X) is the standard deviation of the component during the first non-singular measurement of X.

d. Entropy: E = Entropy (I) returns E, a very valuable attribute of the grayscale element I. Entropy is characterized by the sum (p. * Log2 (p)), which

includes the histological thickness returned by the imhist.

e. Variance:V = var (X) Returns the difference of X for the vector. For the matrix, the character (x) is the row vector that contains the variation of each column of x. For the n-dimensional matrix, the character X works with the first non-singleton dimension. The result is approximate. However, depending on the population variation from which X is drawn, X contains evenly distributed independent samples. Normalizes var V to N - 1 if N> 1, where N is the sample size. For N = 1, V is normalized by 1

f. RMS: The RMS block calculates the real square root square (RMS) value of the input signal. This is calculated using a window of the cyclical mean of the specified fundamental frequency

g. Kurtosis:k = Kurtose (10) reports the measured sample kurtosis of the 10 vectors, and kurtosis (x) is cos x kurtosis of the components. For networks, kurtosis (10) reports, which are ideal for the kurtosis of the N-dimensional 10 singleton kurtosis measurement of group 10, do not work during the first.

h. Skewness:y =the skewness (X) returns the example bias of X. For vectors, the bias (x) is the bias of the elements of x. For matrices, asymmetry (X) is a row vector that contains the sample asymmetry for each column. For N-dimensional matrices, the asymmetry operates along the first non-singleton dimension of X

E. Classification using Neural Networks and SVM

Neural networks are made up of flat elements that work equally well. These elements are motivated by systems of organic sense. As in nature, the relationships between the elements largely decide on the net. You can prepare a Neural and SVM network to steal a certain capacity by changing the estimates of the relationships (loads) between the elements. In general, neural networks are balanced or prepared in such a way that certain information directs a specific objective efficiency. The following figure summarizes such a situation. Here, in light of the correlation between efficiency and target, the network is balanced until the efficiency of the network coordinates the target. Normally, a large number of such sets of information / targets should prepare a network. Figure 7 shows the media technique of the neural network.

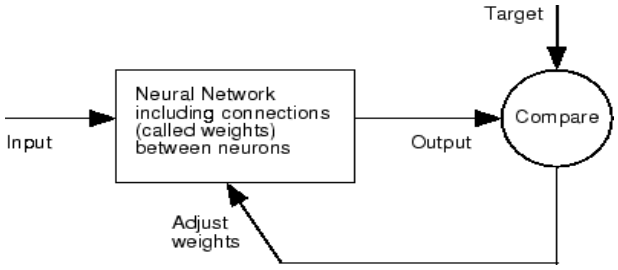


Figure 7.NN Structure

a. Artificial Neural Network (ANN)

ANN is one of the hierarchical strategies generally used in imaging techniques. ANN are different digital models that copy the true neural structure of the brain into three layers. These are the information layer, the folding layer, and the performance layer. A general ANN design is shown in Figure 8.

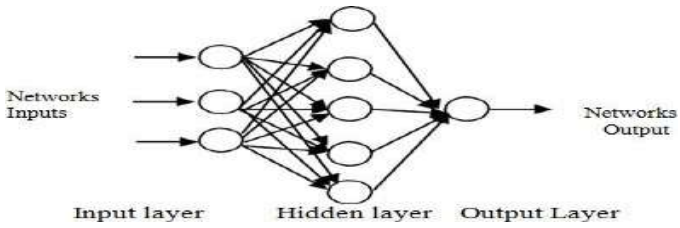


Figure 8. ANN Structure

b. Support Vector Machine (SVM): To classify administrative training models with corresponding training calculations, they are used in studies that are also referred to as reference vector machines. It collects a lot of information and predicts a predefined class between the two available solutions for any given information, which makes it less problematic. paired direct classifier. To display this information in an alternative room, SVM uses the capacity by parts, e.g. B. polynomial, BF, square perceptron, multilevel (MLP). Partitions can also be made with complex restrictions.

3. EXPERIMENT AND RESULTS

This experiment was carried out in two parts: detection of lung cancer from computed tomographic images and segmentation of brain tumors from magnetic resonance images.

a. Lung Cancer Segmentation and Classification

The resources are tested in a diagnostic framework for lung cancer, the sources of which are provided. CT The CT image of the lung receives several methods of image enhancement and the resulting function is obtained. Several ways to update the image used are partitioning and the threshold in the middle. The output of the image update process is used to support the image splitting module. In this article, the central channel function is used as information. The watershed division controlled with the marker is used for the final purpose of image segmentation. The efficiency of the partitioning method is generated and estimated. The results are shown in Figure 9

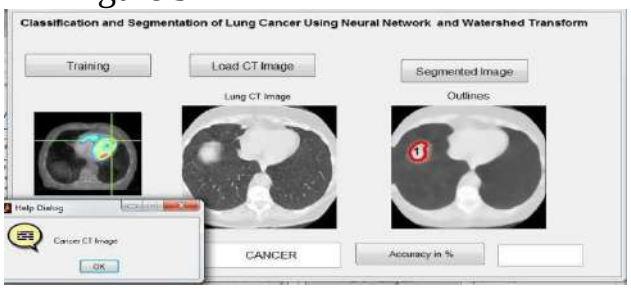


Figure 9.Lung Cancer Detection

Some shades are taken from the image classification. Neural systems use these forces. It uses a neural system to detect and diagnose lung cancer with 100% accuracy.

b. Brain Tumor Segmentation and Classification

This test is based on a brain tumor detection framework, the source of which is the MRI of the brain. Ideally, magnetic resonance imaging is effectively prepared for each progression of the brain tumor localization framework. Magnetic resonance imaging of the brain is used in various image update strategies to calculate yield. The various image enhancement methods used are intermediate

screening and thresholding. The output from the image enhancement strategy is used as a contribution to the image segmentation module. In this task, the performance of the intermediate channel is used as information. The ultimate goal of image segmentation uses marker-controlled watershed segmentation. The results obtained are shown in Figure 10.

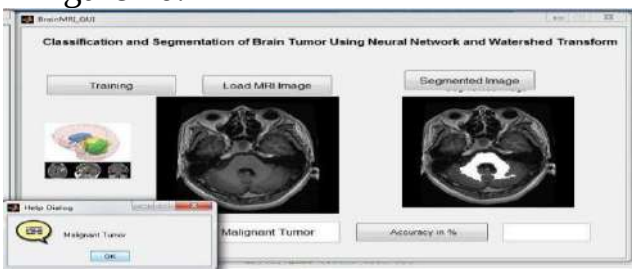


Figure 10.Brain Tumor classification and Segmentation

Neurological systems are developed using the most popular. The tumor in the brain is called formal meningioma, glioma and tumor. Using neurological systems, we obtain 100% certainty for diagnosis and grouping, Figure 11, show Neurological system .

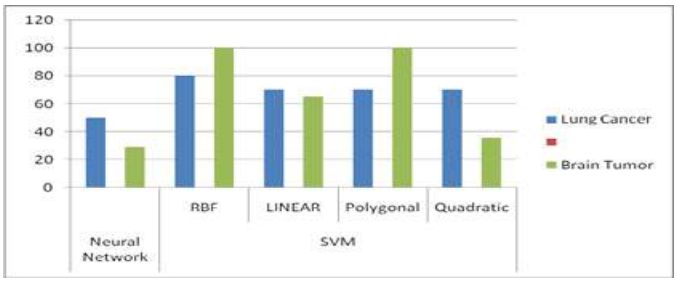


Figure11: System of SVM and Neural Network

4. CONCLUSION

The purpose of this study is to develop a framework that can detect and limit lung nodules in lungs on people's CT scans and identify brain tumors from brain magnetic resonance images. The framework proposed in the study can be implemented with expressed precision. A variety of nodule detection and brain tumor identification calculations were reviewed and addressed in the study. After exploring different strategies, we proposed a flexible strategy to detect three lung nodules and brain tumors from brain MRI images, including pretreatment, segmentation, and knob discovery based on three-stage characterization. It is proposed to use middle channels for SVM groups that use Clammer movements, brain tumor basin changes, or malignant pulmonary nodules and nervous system divisions, and actual headings between them. With 100% accuracy, SVM can be claimed to improve the implementation.

ACKNOWLEDGMENT

I would like to express great fullness to P.G.Dept. of Electronics and Telecommunication, J T Mahajan College of Engineering,Faizpur.

REFERENCES

[1] "Glucan MN, etc." "Detection of pulmonary nodules in chest CT images: preliminary assessment of the computer-aided diagnostic system" "In: ("

2001 ").

[2] "Mc Culloch CC et al." "Model-based detection of pulmonary nodules in computed tomography. Computer-assisted chest diagnosis ", 2004).

[3] Shanu Ge et al. "Computer detection of respiratory nodules: reduction of false positives using the 3D gradient field method and 3D ellipsoidal construction". 2005).

[4] "SLA Lee, AZ Kozani, EJ Hu". "" Automatic detection of respiratory nodules "". In 2008 ").

[5] Alex Krzyzewski, Ilya Sutzkever, Jeffrey E Hinton. "ImageNet classification by conventional neural network". From: (2012), pp. 1097-1105.

[6] Baliscali Bay, Gray Digensen, Patrick van der Smout. "Segmentation of medical image data on CNN". In: CoRR abs / 1701.03056 (2017).

[7] He is Yang, Heng Yong Yu, Ge Wang. "A detailed study for the classification of pulmonary nodules". By: (2016).

[8] Hamidian S et al. "3D convolutional neural network to automatically detect pulmonary nodules on thoracic CT". By: (2017).

[9] Faust Milletari, Nasir Nawab, Dit Ahmad Ahmadi. "V-Net: a completely conventional neural network for volume segmentation of medical images". In: CoRR abs / 1606.04797 (2016).

[10] Özgün Çiçek et al. 3D U-Net: learning of dense volume segmentation from sparse interpretations. In: CoRR abs / 1606.06650 (2016).

[11] Puja Thakur, Dr. Dr. Kurdeep Power November 16, 2012

[12] .S. Charm, m. International conference held in July 2014, p. 1193 to 1199.

[13] In Computer and Communication Technologies (WCCT), World Congress 2014, February 2014, pages 30-33.

[15]. Alan Jose 1, para. Ravi, M Volume 2, Number 3, March 2014

[16]. C $ 31.00 2015 IEEE

[17]. MD. Shakowat Saman Sarkar, Tan Wui Ho, Rajeswaran Lokeswaran, Morphologically Based Image Segmentation Method, International Journal of Information Technology, Vol. N ° 14 and n ° 1.

[18] Manisha Bugwat, R. K Krishna and Vivek Pais, Simple Transformation of Watersheds, International Journal of Computer Science and Communication, Vol. 1, n ° 1, pp. 175-177,2010

LOGISTICS SUPPLIER INTEGRATION IN AUTOMOTIVE INDUSTRY IN PAKISTAN

Mehwish Shabbir. MSCS Student , Dept. of Computing &IT Riphah International University of Lahore, Pakistan.

Abstract. *The automotive industry has created various types of logistical integration between vehicle assembly and its suppliers on the first level. The purpose of this document is to determine which assembly plants (OME) of the most important device manufacturers for logistics integration communicate with the suppliers and classify different types of logistics suppliers and compare them with suppliers. This article explains how buyers and suppliers can be effectively managed in Pakistan's growing automotive industry. And how managers can shape the buyer-supplier relationship strategy so that goals can be easily achieved and there is a need to develop buyer-supplier relationships in order to achieve maximum profit, to achieve production goals and to meet market needs goals.*

1. INTRODUCTION

The Pakistani car industry is one of the quickest developing ventures in the country. It represents 4% of Pakistan's GDP and utilizes more than 1.8 million individuals. There are at present 3,200 car industrial facilities in the country with a speculation of RS 92 billion (USD 650 million), which produce 200,000 vehicles every year. His commitment to the treasury is almost $ 350 million (professional Pakistani official statement, 2015). The segment utilizes 3.5 million individuals and assumes a key job in advancing the development of the flexibly industry. The Pakistani vehicle showcase is viewed as one of the littlest yet quickest developing in Asia. In excess of 180,000 vehicles were sold in the 2014/15 money related year, and 206,777 units were accomplished in the 2015/16 budgetary year. Today the vehicle advertise is commanded by the vehicle brands Honda, Toyota and Suzuki. The car industry is one of Pakistan's developing ventures, overwhelmed by Japanese producers, the vast majority of which have get together plants in the country. The market is overwhelmed by three Japanese vehicle organizations, Honda, Toyota and Suzuki, every one of which has a get together office in Karachi and Honda has its Lahore plant, which are all mutually possessed by neighborhood accomplices.

In 2007, the automobile industry represented 2.8% of Pakistan's GDP, which is relied upon to ascend to 5.6% throughout the following five years. It as of now contributes 16% to assembling, which is relied upon to develop by 25% throughout the following 7 years. Numerous vehicles in the country have two fuel alternatives and Touch CNG, which is less expensive and less expensive than gas in the country.

Pak Suzuki Motors: Pak Suzuki Motors was before the market head in Pakistan with over 60% piece of the overall industry and complete strength of the

little vehicle part. Be that as it may, this number has diminished altogether and JDM vehicles have beated Pak Suzuki because of the way that second rate class vehicles with zero security and negligible solace highlights have been sold at significant expenses in the previous 30 years.

The organization was established in September 1982 as a joint endeavor between the Pakistani government and the Japanese car organization Suzuki and planned the understandings under which awabi auto Ltd had produced the Suzuki SS80, and since 1982. Suzuki initially possessed 25% of the offers and bit by bit expanded their property, which they currently own, by 73.09%. Pak Suzuki was a joint endeavor between Pakistani state-claimed car organizations (PACO), which recently directed nearby gathering utilizing units. Every single strong vehicle produced by Pak Suzuki are reviewed models from Suzuki Japan worldwide that don't fulfill customer security guidelines or present day emanations gauges.

Honda: Honda Atlas Car is a joint endeavor between Honda Motor Company Ltd. Japan and the Atlas Honda Pakistan gathering. The company was included to the rundown November 4, 1992, and a joint endeavor understanding was marked on August 5, 1993. The company is recorded on the Karachi, Lahore and Islamabad stock trades. On July 14, 1994, the Honda vehicle stock began at six vendors in Karachi, Lahore and Islamabad. From that point forward, the company has created 21 3S vendors (deals, administration and extra parts) and 22 sellers (administration and extra parts) just as six seller systems (save parts) in the principle urban communities of Pakistan.

The car business in Pakistan began in 1950 and has experienced different stages, from 1950 (private division) to 1960 (private part). 1970 (nationalization). From 1980 (privatization and passage of the private division).

Toyota: Indus Motor Company Limited works since Toyota Indus has been a Pakistani automaker and a joint endeavor between Habib's home, Toyota Tsusho and Toyota Motors, situated in Karachi, Pakistan, since 1990.

Indus Motor has been the authorized repairman and maker of vehicles, car parts and adornments from Toyota and Daihatsu in Pakistan on a region of 105 hectares since July 1, 1990. Creation office in the industrial port container Qasim on the edges of Karachi.

Indus Motor was established in December 1989 as a constrained obligation organization and started business creation in May 1993. The organization's offers are recorded on the Pakistani stock trade. Toyota Motors and Toyota Tsusho Corporation have a 25% stake in the organization, while Habib's greater part investor is at home.

2. LITERATURE REVIEW:

Pakistan's car industry history is one of the most seasoned in Asia. The industry began delivering Cargo Bath (Bedford) in 1949 by General Motors. This denotes the start of mechanical history after the freedom of British India. Since this year, the industry has not been developing consistently and has retreated to take over other Asian nations, for example, China, Thailand and India, and its situation in the worldwide market is likewise addressed. Business speeds don't bolster

Pakistan's car industry, so purchasers face more significant expenses, settle for the easiest option and have the right to be progressively serious in this area. Being the 6th biggest maker in the car industry, utilizing 3.5 million laborers, contributed 2.8% of GDP in 2012, indicating more than Rp300 billion (KPMG Investment in Pakistan 2013) . The market had a creation limit of 751,887 units in the 2012-13 season, yet is a concentrated market with just eight models representing all household car deals. The models are made by three Japanese organizations: Honda, Suzuki and Toyota.

There are in excess of 100 organizations in Pakistan that collect cars, for example, cars, trucks, transports, bikes, toy cars and tractors. Around 1,700 vehicle parts manufacturers are occupied with Pakistan's car industry. Part manufacturers are regularly the biggest neighborhood organizations, yet OEM manufacturers unite significant Japanese organizations. Creation and specialized frameworks face numerous shortcomings because of absence of seriousness in the industry because of natural needs. Remote venture is additionally low in certain areas of the industry.

After the principal autos were delivered in Pakistan-In the mid 1990s, deforestation after confronted with exacting guidelines from the earliest starting point, Japanese manufacturers made high rivalry by entering the market. At the 1999 Ford Car Summit, little cars and enormous blended race cars (LCV) including Daihatsu were drawing nearer, and Hino, Mazda, and Suzuki 1984 showed up. , 1993 Toyota and 1994 Honda. Most Pakistani cars have two fuel choices, so they run on gaseous petrol (CNG), which is less expensive and more impressive than Pakistan's gas. The industry just arrangements with deals in developing urban communities, specialized collaboration understandings and moves from China, Japan, Korea and Europe. Pakistan falls into the classification of a few manufacturers around the globe who create or coordinate a wide range of vehicles including overwhelming vehicles and transports, 2/3 transports, cars, traveler cars, traveler trailers, trailers, trailers and trailers. , LCV. Major worldwide brands incorporate Honda, Suzuki (at present the market head with 60% piece of the overall industry), Toyota, Nissan and general motors. The primary coal industry comprises of CNG, workshops, vehicle parts stores and tire stores. Pakistan (PAMA), Pakistan Automobile Manufacturers Association (PAAPAM), Pakistan Automobile Manufacturers (APMA) biggest producer of automakers. All associations give individuals fundamental and valuable data. Moreover, the Engineering Development Board (EDB) holds data about the car industry.

3. RESEARCH GAP/PROBLEM IDENTIFICATION:

For certain years, numerous specialists have taken in evolved nations with respect to provider the executives in created nations. The specialists found that numerous elements, including governmental issues, society, innovation, lead to provider the board. These created nations capitalize on the overview results for their association. We presently trust it is important to examine provider the executives in the Pakistani setting.

4. RESEARCH OBJECTIVE:

The fundamental target of this examination is to investigate the helpful and enlightening components that really influence provider the executives with provider and purchaser in car industry in Pakistan, likewise another goal is to explore the accompanying exploration questions,

- Does coordinations affect the car business?
- Does strategic provider impact on car industry?
- Does strategic provider impacts the presentation of the car business??

5. METHODOLOGY:

The information in this article originates from the writing survey and the subjective research technique. Also, we visited three distinctive vehicle producers, including SUZUKI, HONDA and TOYOTA. We likewise led interviews with their specialists and utilized a semi-organized survey.

6. ANALYSIS/DISCUSSIONS:

One respondent said that, specifically as coordinations or providers, it isn't our decision. It has just been chosen by our organization. They pick this, and afterward we prohibit them. It is sent to providers due to our shipments come only for them. When the shipment contacted us. We discover vehicles and afterward we check and convey to clients. It would appear that we have a devoted space for us, and it is moving every now and then and that the following level is our actual clients.

Keep in mind, we are in Pakistan, so doing however much as could be expected is troublesome. To locate the correct dealer or devotee, make more calls, ask, ask and ask, so you may need to reply.

They will never arrive at the most significant level; they will consistently arrive at the primary level one. So when they arrange, they identify with the youthful. This is the standard system.

Another respondent stated: Pak Suzuki has not many providers for providing modest Suzuki items If item quality issues or providers end up being inadequate or not usable, the organization gets client criticism and tells the provider and the organization dispatches a re-call battle and clients are gathered. from the deal understanding. In the review crusade, calls are sent to providers for development and afterward purchasers are at long last advised of changes made to the item.

For instance, if 10,000 units are offered to clients, clients will get a severe greeting, with merchandise traded for nothing and damaged parts came back to providers.

Moreover, these providers offer various items to various clients. Pak Suzuki concentrates more on consumer loyalty and this is their essential objective.

The motors control motors (ECMs) from similar providers if the incase section is imperfect and the objection has experienced a few unique sellers, making the nature of the item be accused and supplanted by providers all things considered.

To get more consumer loyalty, Pak Suzuki, Toyota and Honda furnish clients with a 3-year guarantee, empowering them to hang tight for the acquisition of items with security and guarantee to be fixed and supplanted.

7. CONCLUSIONS/FUTURE RECOMMENDATIONS:

In view of the model, a considerable lot of the respondents expressed that we don't pick providers on the grounds that the primary site chooses the provider and afterward the provider disseminates the requests to different sellers. What's more, at times the provider conveys the request on schedule, yet different occasions it isn't because of low correspondence. Effect of the coordinations supplier on the exhibition of the car business. Save parts conveyance time, similar to each request time, can influence the entire procedure because of clients.

Furthermore, there are three significant vehicle sales centers available; Honda, Toyota and Suzuki. Suzuki has an imposing business model on the little vehicle portion, while confronting practically no opposition aside from the Daihatsu Cuore-molded Toyota Indus motor. In the Honda City, Honda Civic, Toyota Corolla subcompact car portion at present the main vehicles in the assembling procedure. A privately fabricated games utility vehicle (SUV) isn't yet accessible, nor is a full or medium size car.

Pakistan as a creating nation, with moderate innovative turn of events, requires a steady government approach and a sound traffic plan for the automobile business to lessen vulnerability between nations. Nearby producers and merchants in Pakistan. Innovation updates, very much prepared providers and solid government money related help are required for the accessibility of vehicle vendors and constructing agents in Pakistan. Pakistan's automobile industry depends intensely on imports from Korea, Japan, China, and the United States, and independence stays a fantasy for Pakistan's vehicle industry. In any case, there is promise for an update in the automobile business. Pakistan's auto showcase has gotten increasingly serious as of late, and there is likewise an unmistakable move towards a littler vehicle fragment. The aftereffects of the audit and definite proposals gave to the Pakistani vehicle industry are relied upon to without a doubt advantage the car business in the nation.

REFERENCES:

1. M.N. Abbasi, M. Afzal, "Effect OF E-SOURCE APPLICATION ON KEY BUSINESS STRATEGIES – EVIDENCES FROM AUTOMOTIVE INDUSTRY OF PAKISTAN", Journal of Quality and Technology Management Volume VII, Issue II, December, (2011).

2. .Abeer Imam, Noor Hussain, Aiman Raza, "Powerful Supplier Management in Automobile Industry in Pakistan", Journal of Applied Environmental and Biological Sciences, (2015).

3. David Bennett Florian Klug, (2012),"Logistics provider reconciliation in the car business", International Journal of Operations and Production Management, (2012).

4. Hafiz Pasha, Zafar Ismail, "An Overview of Trends in the Automotive Sector and the Policy Framework", Automotive Sector in Pakistan Phase I Report, (2012).

5. Juthathip Sararaksa and Kwang Sup Shin, "Near Analysis of Factors for

Supplier

6. Determination and Monitoring: The Case of the Automotive Industry in Thailand" Sustainability (2019)

7. .Usman Qadir, Pakistan's Automotive Industry: A Case of Stalled Development, PIDE Working Papers No. 137, PAKISTAN INSTITUTE OF DEVELOPMENT ECONOMICS, July (2016).

ИЗУЧЕНИЯ ИК-ПОТОКА В УСТАНОВКАХ ДЛЯ СУШКИ ВИНОГРАДА

Юсупов М.Т., Саримсаков Х.У., Андижанский машиностроительный институт, Тожибоев А.М., Ученик средней школы Алтинкульского района Андижанской области, Узбекистан.

Annotation: *Currently, the process of drying fruits and vegetables is carried out in several ways. One such method is to use infrared (infrared) rays to transfer heat to both the product and yourself. This article highlights the effect of infrared light flux on drying grapes in devices.*

Keywords: ***heat-mass exchange process, infrared, range, IR exposure, experimental device.***

Аннотация: *В настоящее время процесс сушки фруктов и овощей осуществляется несколькими способами. Один подобный метод заключается в использовании инфракрасных (инфракрасных) лучей для передачи тепла как продукту, так и себе. В этой статье освещена влияние инфракрасного светового потока на сушку винограда в устройствах.*

Ключевые слова: ***тепло масса обменный процесс, инфракрасный, диапазон, ИК воздействие, экспериментальное устройства.***

Аннотация: ***Ҳозирги кунда мева сабзаотларни қуритиш жараёни бир қанча усуллар билан амалга оширилади. Худди шундай усуллардан бири бўлиб, махсулотга, шунингдк узимга иссиқлик ўтказишда ИҚ(инфра қизил) нурлардан фойдаланишдир. Ушбу мақола қурилмаларда узимни қуритишда ИҚ нурлар оқимини таъсирини ёритган.***

Калит сўзлар: *иссиқлик модда алмашинуви, инфрақизил, оралиқ, ИҚ – таъсири, тажриба қурилмаси*

В существующих технологических линиях в процессе переработки винограда для получения сушеных продуктов ТМО процессы организуется в основном конвективным способом подвода тепла. Имеющиеся аппараты и установок для реализации процесса сушки винограда не достаточно подкреплений теоретическими и экспериментальными предпосылками, а также имеет существенные недостатки при организации процесса. Данный способ сушки основывается на передаче тепла высушиваемому продукту за счет энергии нагретого сушильного агента.

Однако, по этому способу характерное недостатки, существенно снижающие качество конечного продукта. Высокая температура и большая продолжительность сушки способствуют развитию окислительных процессов.

Целесообразность применения новых физических методов подвода энергии в поле ИК- диапазона обусловлено высокой эффективности при предварительной обработки и сушки сельскохозяйственных продуктов. Следовательно в качестве источника ИК- излучения необходимо

использовать генераторы, длина волны излучения которых находится в диапазоне соответствующем наибольшей пропускательной способности продукта в зоне предварительной обработки а в зоне сушки наиболее поглощательной способности влаги. С этой точки зрения, на основе проведенных экспериментальных и теоретических исследований определены оптимальные значение тепловых потоков в процессе переработки сельхозпродуктов.

Таким образом, выбранные режимы для предварительной обработки и сушки, является целенаправленным для интенсификации процесса сушки. Применение новых физических методов сушки дает возможность получить следующие результаты:

- уменьшить затраты энергии на единицу готовой продукции за счет
- эффективности процесса сушки на этапе ИК- воздействия;
- повысить качество сушеных продуктов в результате сохранения ценных компонентов.

Для равномерного удаления влаги по всей объемам высушиваемых продуктов необходимо требуется равномерность подачи тепла, в противном случи, удлиняется процесса сушки и приводит к увеличение удельного затрата энергии. Например, для получения 1 кг сушеного продукта, с помощью существующих сушильных установок следует, удалит от 4,0 до 9,5 кг влаги. При этом удельный расход энергии составляет от 20 тыс. до 27тыс кДж/кг или 0,65 до 0,91 кг условного топлива.

С этой точки зрения для разработки энергосберегающей технологии по переработке винограда нами разработано нетрадиционный способ сушки новым физическим методом подвода энергии. Однако для широкой реализации предлагаемой способ сушки необходимо решить проблемы воздействия источника ИК- излучателей и ее распределение на поверхности высушиваемого продукта.

На основе экспериментальных и теоретических предпосылок определены пропускательные и поглощательные способности высушиваемого продукта в области ИК- диапазона.

Путем математического моделирование изучены взаимовлияния всех систем в процессе сушки и определены оптимальные значение влияющих факторов такие, как длина волны излучения λ, плотность теплового потока q, толщина слоя продукта δ, удельная нагрузка σ на единицу площади F сушильной камеры а также режим сушки в области ИК- диапазона.

Следует отметить, что для проведения процесса сушки в электромагнитном поле ИК- диапазона требуется равномерное распределение лучистого потока на поверхности изделий. Исходя этого нами теоретически обосновано расположение ИК- излучателей в предлагаемой сушильной установки.

В нашем примере на поверхности высушиваемого продукта высоты расположения излучателей относительно продукта Zu, шаги между излучателям S, расстояния от отражателей hp.

При горизонтальном расположении над плоским конвейером и одностороннем облучении, плотность потока от одного излучателя определяется по формуле [1, 2].

$$E_i = \frac{E_э \cdot d_э \cdot Z_u}{2 \cdot \beta} = \left| \frac{x - v \cdot \tau}{(x - v \cdot \tau)^2 + \beta^2} + \frac{1}{\beta} \cdot arctg \frac{x - v \cdot \tau}{\beta} \right|_{x_1}^{x_2}$$

где υ- скорость движения конвейера, м/с; τ-время, с.; x_1, x_2 – координата начального и конечного излучателя; i- порядковый номер излучателя в блоке; $d_э$- эквивалентный диаметр ИК- излучателя, м; $E_э$- энергетическая светимость излучателя, Вт/м^2.

Графическое изображение подающего потока излучения на точку $A(X_a, У_a)$ от i-го излучателя приведен на рис.1.

Результирующий поток излучения на поверхность продукта определяется по формуле [3].

$$E_p = E_{pu} + E_{pu}(R_p)$$

где E_{pu} -плотность результирующего потока, который определяется:

$$E_{pu} = B_э \cdot d_э \cdot K_{осв} \cdot \left(\frac{A}{1 - R \cdot R_p} \right)$$

где $B_э$- энергетическая яркость излучателя; A- интегральная поглощательная способность продукта; R- интегральная отражательная способность продукта; $R_р$-интегральная отражательная способность рефлектора; $К_{осв}$- коэффициент освещенности поверхности материала, который определяется по следующей формуле:

$$K_{осв} = \left(\frac{E}{B_э \cdot d_э} \right)$$

E- энергетическая освещенность излучателя;

E_{pu} и $R_р$- отраженная от рефлектора составляющая плотности результирующего потока:

$$E_{pu}(R_p) = B_э \cdot R_p \cdot d_э \cdot K_{осв} \cdot \frac{A}{1 - R \cdot R_p}$$

Суммарный коэффициент облученности обрабатываемого материала от блока, содержащего n-го излучателей определяется по выражению

$$K_{осв} = \sum_{i=0}^{n-1} \cdot K_{осв}$$

Расчеты выполнены на программы Turbo-Pascal и MATLAB. В результате расчетов установлено, что для выбранного способа переработки в новым физическим подводом энергии с учетом терморадиационных характеристик высушиваемых продуктов необходимая плотность потока ИК-

излучений данного материала, (в зоне сушки) равная q=1,5-2,0 кВт/м^2 достигается при $К_{осв}$=96 м$^{-1}$. Как видно из кривых, такую величину $К_{осв}$ можно получить при n=9, Zu=0, 17-0, 20 м.

Анализ расчетов показали, что при расположении ИК- излучателей с одинаковым шагом, не обеспечивается равномерное распределение лучистого потока на поверхности обрабатываемого материала. На краях конвейера плотность падающего потока составляет 2-2,8 кВт/м^2 а в центре около 7-7,5 кВт/м^2. Это свидетельствует о том, что при таких распределениях лучистого потока, в процессе предварительной обработки и сушки происходит неравномерность удалении влаги по всей поверхности изделий, в результате чего удлиняется самого процесса и влияет на качество сушеного продукта.

Для более равномерного распределения целесообразно расположить ИК -излучатели с переменным шагом. Такое расположение излучателей обеспечивает равномерное распределение плотности потока на краях конвейера как в центре так и на периферии, и плотность потока поверхности материала составляет 1,5-2,00 кВт/м^2 .

Полученные данные использованы для инженерного расчета ИК - конвективных сушильных установка непрерывного действия и при создании полупромышленной сушильной установки.

Литература

1. Левитин И.Б. Применение инфракрасной техники в народном хозяйстве. -Л.: Энергоиздат.1981.-264с.

2. Ильясов С.Г., Красников В.В. Методы определения оптических и терморадиационных характеристик пищевых продуктов. -М.: Пищевая промышленность. 1972. -175 с.

3. Плаксин Ю.М. Исследование процессов выпечки мучных кондитерских изделий в печах с инфракрасным излучением.: Дисс.канд.техн.наук .-М.: 1972.-254с.

ИЗУЧЕНИЕ ПРОЦЕССА СУШКИ ВИНОГРАДА ИСПОЛЬЗУЯ СИСТЕМНОГО АНАЛИЗА

***Юсупов М.Т , Парпиев Г.Г.*, Андижанский машинастроительний институт, Тургунова М., учителница средней школы город Чирчик, Узбекистан.**

Annotation: *Currently, around the world, a system analysis is used to solve problems and issues. And also, to provide the population with high-quality agricultural products, a number of works and devices are used.*

This article highlights the study of the drying process of grapes based on system analysis.

Key words: *technological line, subsystem, air distribution, process of IR - convective and vacuum drying*

Аннотация: *В настоящее время по всему миру, для решения проблемы и вопросов используется системного анализа. А также для обеспечения населения высококачественной сельскохозяйственной продукцией используется рад работ и устройств.*

В данной статье освещена изучения процесса сушки винограда на основе системного анализа.

Ключевые слова: *технологическая линия, подсистема, распределение воздуха, процесс ИК - конвективной и вакуумной сушки*

Аннотация*: Ҳозирги кунда дунё миқёсидаги муаммо ва масалаларни хал қилишда тизимли тахлилдан фойдаланлади. Шунингдек, қишлоқ хўжалиги махсулотларини ахолига сифатли тарзда етказиб бериш мақсадида бир қатор ишлар ва қурилмалар яратилган.*

Ушбу мақоламиз "Узумни қуритиш технологиясини тизимли тахлил асосида ўрганишга" бағишланган.

Калит сўзлар: *технологик тизим, тизим ости, хаво тақсимланиши, ИҚ жараён – конвектив ва вакуумда қуритиш*

В Узбекистане особый приоритет был отдан реформированию сельского хозяйства. Это исходило из той роли, какую играет сельское хозяйство и отрасли, связанные с переработкой сельскохозяйственного сырья, в современной экономике Республики Узбекистан.

Критический анализ существующих методов сушки, свидетельствует, что рекомендуемые способы энергоемки, материалоемки и малоэффективны, т.е. в каком – то мере отрицательно отражается на продолжительности сушки, в связи с чем, они не очень распространены.

Известно, что сушка сельхозпродуктов в обычном понимании этого слова представляет собой удаление из них влаги или других растворителей, молекулы которых не утратили своей индивидуальности. В общем случае в материале, кроме этой влаги, могут быть и такие, молекулы

которых входят в химическое соединение с молекулами вещества и образуют гидраты его окисов; в материале влага может также быть связана в виде гидроксильных ионов за счет сил главных валентностей, при этом молекула воды как таковая исчезает. Удаление этих видов влаги представляет собой нечто среднее между физико-химическим явлением и химической реакцией. После удаления химически связанной влаги существенно изменяются физические свойства вещества. Совершенно ясно, что в тех случаях, когда технологические свойства материала при сушке не должны быть изменены, эта влага в материале должна оставаться и при сушке, удалению подлежит только влага, расположенная внутри пор высушиваемого материала и пропитывающая стенки твердого скелета тела.

В связи со сложностью поставленной задачи важной становится проблема научно обоснованного подхода к разработке и проектированию современного производство сушки винограда. Предпосылки для ее решения, для перехода на новые методы проектирования и анализа промышленно - технологических процессов есть — это разработка теоретических основ процессов современной технологии, широкое применение вычислительной техники, а также внедрение в арсенал познавательных средств современной науки принципов и методов системного исследования [1,2,3].

В связи с этим нами были исследованы процесс сушки винограда с применением новых методов анализа и синтеза, т.е. современных научных принципов и методов системного исследования промышленно - технологических процессов [4].

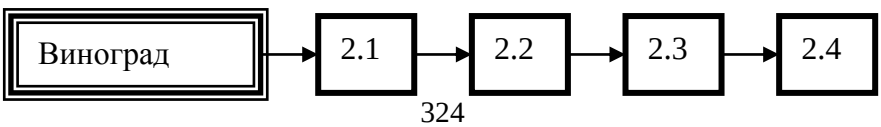
С целью совершенствования процесса сушки винограда, изучения тепло- и массообменных явлений, протекающих в продуктах при их обработке, и определения влияния различных факторов на качество сушеного продукта проведен комплексный анализ технологической линии и обоснована ее иерархическая структура.

На этот основе разработана рациональная технологическая схема, которая призвана позволить получать высококачественную сушеную продукцию.

При этом задача состоит в том, чтобы в процессе предварительной ИК - обработки и сушки, в первую очередь, максимально сохранить биохимический состав винограда.

Для этого требуется разработка общего подхода к исследованию процесса сушки с реализацией регулирования мощности теплового источника, поля скоростей и параметра теплоносителя.

Технологическую линию переработки винограда можно представит в виде большой системы. Здесь технологические процессы можно представить в виде подсистемы получения высушенного продукта из винограда. Данная система подвергается к декомпозиции и на втором этапе анализа имеет место ряд под систем, состоящих из следующих аппаратов:



Подсистема 2.1. – аппарат для мойки. В этом объекте протекают процесс очистки поверхности продуктов от наружных примесей с помощью промышленной воды, а также частичная сортировка и удаление наружных сопутствующих веществ.

Подсистема 2.2.- аппарат для предварительной обработки винограда, котором происходит обработка продуктов под ИК -воздействием с соответствующей длиной волны. Благодаря этому существенно возрастает коэффициент диффузии влаги в материале.

Подсистема 2.3.- аппарат для сушки. Здесь протекает процесс удаления влаги в воздушную среду.

Подсистема 2.4. представляет собой в процессе расфасовки и упаковки высушенного продукта.

В рассмотренных подсистемах второго уровня иерархии основные тепло массообменные процессы происходят в системах 2.2. и 2.3.

С целью комплексного анализа процесса сушки винограда нами составлена технологическая схема с детализацией по уровням иерархии (рис.1.а, б).

На втором уровне иерархии нами рассмотрена под система 2.3., которая представляет собой аппарат для предварительной обработки. Совершенствование организации процесса достигается за счет конструктивных параметров аппарата, а управление процессом осуществляется путем изменения технологических параметров.

Для детального изучения тепло масса обмена на третьем уровне иерархии под система 2.2. расчленена на подсистемы 2.2.1 и 2.2.2. Элементами подсистемы являются нагреватель и твердая фаза. Выбор нагревательного элемента осуществлен на основе изучения ТФХ и ТРХ винограда. Процесс нагрева твердой фазы в под системе 2.3 осуществляется под воздействием системы 2.2.1. Механизм действия излучения на облучаемые вещества (твердую фазу) касается превращения лучистой энергии в тепловую.

Математическая модель подсистем включает в себя математические описания процесса перемещение влаги и тепла в твердой фазе, которые описываются соответствующими дифференциальными уравнениями.

На четвертом уровне иерархии подсистема 2.2.2 (твердая фаза) расчленяется на подсистемы твердая и локальная сплошная фазы.

Подсистема локальная сплошная фазы представляет собой ленту транспортера. При анализа этого объекта рассматривается механизм движения транспортера, процесс нагрева и механизм пропускания T_λ, отражения R_λ материала ленточного транспортера. Однако при этом основным является равномерность распределения лучистого потока на поверхности облучаемого продукта. Оптимальная обработка винограда состоит в обеспечении времени пребывания ленточного транспортера в зоне облучения. В процессе обработки винограда за счет проникновения и

поглощения ИК - лучей происходит нагрев под системы, в связи, с чем частично удаляется влага из продукта.

Таким образом, элементами четвертого уровня иерархии являются твердая и локальная сплошная фазы.

Управление процессами осуществляется за счет регулирования равновесия в системе, выражаемого в виде разности температур и скорости воздуха.

Улучшение организации процесса достигается за счет обеспечения равномерного распределения твердой фазы в сплошной. Математическая модель подсистемы включает в себя математические описания процессов, протекающих в системах с твердой фазой, уравнения равновесной влажности, коэффициента и скорости сушки.

На пятом уровне иерархии происходит совокупность химических и теплофизических явлений имеющих место в элементах твердой фазы. Структурными элементами здесь выступают белки, сахара, витамины и микроэлементы, находящиеся в составе твердых частиц.

К полезному эффекту данного уровня относится повышение кинетической энергии молекул воды, обусловленное внешним тепловым воздействием. Отрицательным эффектом - это частичная денатурация белков и потери аминокислот.

Из внешних причин, влияющих на эффекты пятого уровня, следует выделить воздействие вышестоящих ступеней иерархии, проявляющиеся в виде различных кинетических, диффузионных и термодинамических эффектов.

Исходя из вышеизложенного, в качестве основного параметра, характеризующего поведение подсистемы данного уровня можно принять температуру. Образование и превращение продуктов описывается системой дифференциальных уравнений.

Процесс ИК - конвективной и вакуумной сушки плодоовощных культур (под система 2.3.) рассмотрен на втором уровне иерархии. Элементами системы здесь выступают рабочая камера, вентилятор, фильтр и вакуум насос. Подсистема 2.3.1. представляет собой рабочую камеру сушильной установки, которая снабжена нагревательными элементами, отражателями и сетчатым подом. Вспомогательным оборудованием являются под системы 2.3.2., 2.3.3. и 2.3.4.

Подсистема 2.3.2. – вентилятор, обеспечивающий равномерную подачу воздуха по всему объему конвективной сушильной установки. Подсистема 2.3.3 - фильтр - обеспечивает очистку поступающего воздуха.

В процессе ИК - вакуумной сушки для создания вакуума предположена подсистема 2.3.4, которая представляет собой вакуум - насос.

Улучшение организации процессов протекающих в отдельных элементах аппарата, достигается путем упорядочения параметров отдельных токов за счет конструктивного соответственного оформления элементов установки.

На третьем уровне иерархии под система 2.3.1. расчленена на подсистемы 2.3.1.1, 2.3.1.2 и 2.3.1.3. Подсистема 2.3.1.1 представляет собой элемент ИК - нагревателя, который аналогичен предыдущей подсистеме 2.2.1.

Подсистема 2.3.1.2 – газовая фаза. Здесь протекают процессы перемещения, поглощения влаги и теплообмена. Под система 2.3.1.3. – твердая фаза. Здесь протекают процессы, аналогичные явлениям в подсистеме 2.3.2. Однако необходимо отметить, что в ходе сушки основная часть ИК - энергии поглощается водой, в связи, с чем интенсифицируют, убила влаги из материала.

На четвертом уровне иерархии рассматриваются элементы подсистемы, связанные с твердой фазой. Для обеспечения равномерного прохода и распределения воздуха по всей поверхности высушиваемого продукта изготовлено в виде решетки из нержавеющей стали. Коэффициент прохода воздуха составляет 0,98.

В ходе сушки за счет тепло- и массообмена удаляется влага. В процессе нагрева происходят структурные изменения плодоовощных культур, имеющие место на пятом уровне иерархии. Основным процессом на данном уровне иерархии являются структурные превращение биологических веществ, явления денатурации белков, разрушения микроэлементов и потери витаминов. Математическая модель подсистем представляет собой математическое описание явлений протекающих в структуре плодовых культур и описываемых соответствующими дифференциальными уравнениями.

Декомпозиция системы ТМО процессов на подсистемы, обладающие целостностью и целенаправленностью, позволяет анализировать их с учетом взаимовлияния и взаимосвязи между собой и с внешней средой и вскрыть иерархическую структуру системы.

При этом на каждой иерархической ступени соответствующего ТМО процесса имеет место своя подсистема с соответствующими параметрами, для которой формируются способы управления и улучшения организации тепло- и массообменных процессов.

Полная математическая модель процесса сушки винограда, протекающих в масштабе аппарата в целом включает в себя математические описания процессов и явлений, происходящих на рассмотренных выше уровнях иерархи.

Таким образом, иерархическая структура рассматриваемых в данной работе тепломассообменных процессов представляет собой сложную систему, состоящую из пяти ступеней иерархии:

На первом уровне рассматривается процессы, протекающие в объеме сушильного аппарата.

На втором уровне рассматривается биохимические и ТМО процессы в объеме винограда.

На третьем уровне рассматривается биохимические и ТМО процессы, протекающие в слоях винограда.

На четвертом уровне рассматривается явления, относящиеся к сахарам и витаминам, которые вызывают фазовые превращения и биохимические реакции, для чего поглощается тепло.

На пятом, молекулярном уровне рассматривается изменения внутренней энергии материала.

Составление иерархической структуры дает возможность разделить все явления и процессы, протекающие на том или ином уровне иерархии, на необходимые эти явления и процессы, предусмотренные технологическим процессом и явления, подлежащие ликвидации, разложение витаминов и так далее. Ясно, что практически невозможно полностью ликвидировать один явления и процессы для обеспечения осуществления других.

Вместе тем уровень развитие научных исследований пока не может осуществить количественную оценку явлений и процессов. Однако представленное определение иерархической структуры процессов переработки сельхозпродуктов позволяет анализировать представляемых явлений и процессов и дать оценку им.

Литература

1. Общая химическая технология: Учеб. для техн. вузов /А. М. Кутепов, Т. И. Бондарева, М. Г. Беренгартен.—2-е изд.,испр. и доп. — М.: Высш. шк., 1990. — 520 с.; ил.15ВМ 5-06-000493-7

2. Кафаров В.В., Глебов М.Б. Математическое моделирования основных процессов химических производств. - М.: Высшая школа,1991.С. 400.

3. Закгейм А.Ю. Введение в моделирование химико-технологических процессов. М.: Химия, 1982. -288 с.

4. Анализ и синтез биотепломассоооменных прцессов// А.А. Артиков, А.Х. Маматкулов, Н.И. Хамидов–Ташкент. Фан. 1994.-134 с.

КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ПРОСТРАНСТВЕННОЙ ДИНАМИЧЕСКОЙ КРАЕВОЙ ЗАДАЧИ ТЕОРИИ УПРУГОСТИ ДЛЯ ОРТОТРОПНЫХ ТЕЛ

Каландаров А.А., Каландаров А., Аллаяров С.П., Адилов А.Н.
Гулистанский государственный университет., Узбекистан.

Аннотация. В статье рассматривается компьютерное моделирование на основе нового подхода к численному решению пространственной динамической краевой задачи теории упругости для ортотропных тел на основе неявных конечно разностных схем

Abstract. The article discusses computer modeling based on a new approach to the numerical solution of the spatial dynamic boundary value problem of the theory of elasticity for orthotropic bodies based on implicit finite difference schemes

Рассмотрим динамическую краевую задачу теории упругости для ортотропных тел, которая состоит из уравнения движения [1,2,3]

$$\sum_{j=1}^{3}\frac{\partial\sigma_{ij}}{\partial x_j}+X_i=\rho\frac{\partial^2 u_i}{\partial t^2} \tag{1}$$

закон Гука, определяющий соотношения между напряжениями и деформациями для анизотропных материалов

$$\sigma_{ij}=C_{ijkl}\varepsilon_{kl} \tag{2}$$

соотношения Коши

$$\varepsilon_{ij}=\frac{1}{2}\left(u_{i,j}+u_{j,i}\right) \tag{3}$$

с начальными

$$u_i\big|_{t=t_0}=\varphi_i\,,\quad \dot{u}_i\big|_{t=t_0}=\psi_i \tag{4}$$

и краевыми условиями

$$u_i\big|_{\Sigma_1}=u_i^0\,,\quad \sigma_{ij}n_j\big|_{\Sigma_2}=S_i^0 \tag{5}$$

где σ_{ij} – тензор напряжений, ε_{ij} – тензор деформаций, u_i – перемещения, X_i – объёмные силы, C_{ijkl} – тензор четвёртого ранга, определяющий механические свойства материала, ρ – плотность тела.

Во многих случаях задачи теории упругости решаются относительно перемещений. Рассмотрим задачу (1)-(6) в трёхмерном случае. Для этого, подставляя уравнение (3) в (2) и полученное в (1), будим иметь уравнение движения в перемещениях (делаем замену $u=u_1,\ v=u_2, w=u_3$), а именно

$$\begin{cases} C_{1111}\dfrac{\partial^2 u}{\partial x^2}+C_{1212}\dfrac{\partial^2 u}{\partial y^2}+C_{1313}\dfrac{\partial^2 u}{\partial z^2}+(C_{1122}+C_{1212})\dfrac{\partial^2 v}{\partial x\partial y}+(C_{1133}+C_{1313})\dfrac{\partial^2 w}{\partial x\partial z}=\rho\dfrac{\partial^2 u}{\partial t^2} \\ C_{1212}\dfrac{\partial^2 v}{\partial x^2}+C_{2222}\dfrac{\partial^2 v}{\partial y^2}+C_{2323}\dfrac{\partial^2 v}{\partial z^2}+(C_{2211}+C_{1212})\dfrac{\partial^2 u}{\partial x\partial y}+(C_{2233}+C_{2323})\dfrac{\partial^2 w}{\partial y\partial z}=\rho\dfrac{\partial^2 v}{\partial t^2} \\ C_{1313}\dfrac{\partial^2 w}{\partial x^2}+C_{2323}\dfrac{\partial^2 w}{\partial y^2}+C_{3333}\dfrac{\partial^2 w}{\partial z^2}+(C_{3311}+C_{1313})\dfrac{\partial^2 u}{\partial x\partial z}+(C_{3322}+C_{2323})\dfrac{\partial^2 v}{\partial y\partial z}=\rho\dfrac{\partial^2 w}{\partial t^2} \end{cases} \quad (6)$$

Начальные и краевые условия в трёхмерном случае будут иметь следующий вид

$$\begin{cases} u(x,y,z,t)\big|_{t=0}=0,\ \dfrac{\partial u}{\partial t}\Big|_{t=0}=0 \\ v(x,y,z,t)\big|_{t=0}=0,\ \dfrac{\partial v}{\partial t}\Big|_{t=0}=0 \\ w(x,y,z,t)\big|_{t=0}=0,\ \dfrac{\partial w}{\partial t}\Big|_{t=0}=0 \end{cases} \quad (7)$$

$$\begin{cases} u\big|_{\Gamma_1}=u^o,\ v\big|_{\Gamma_1}=v^o,\ w\big|_{\Gamma_1}=w^o \\ (\sigma_{11}n_1+\sigma_{12}n_2+\sigma_{13}n_3)\big|_{\Gamma_2}=S_1 \\ (\sigma_{21}n_1+\sigma_{22}n_2+\sigma_{23}n_3)\big|_{\Gamma_2}=S_2 \\ (\sigma_{31}n_1+\sigma_{32}n_2+\sigma_{33}n_3)\big|_{\Gamma_2}=S_3 \end{cases} \quad (8)$$

где $\Gamma=\Gamma_1+\Gamma_2$ граница рассматриваемой пространственной области.

Заменяя частные производные в дифференциальных уравнениях (6) соответствующими конечно-разностными отношениями, получим

$$C_{1111}\frac{u_{i+1,j,m,k+1}-2u_{i,j,m,k+1}+u_{i-1,j,m,k+1}}{h_1^2}+C_{1212}\frac{u_{i,j+1,m,k+1}-2u_{i,j,m,k+1}+u_{i,j-1,m,k+1}}{h_2^2}+$$

$$C_{1313}\frac{u_{i,j+1,m,k+1}-2u_{i,j,m,k+1}+u_{i,j-1,m,k+1}}{h_3^2}+(C_{1122}+C_{1212})\frac{v_{i+1,j+1,m,k}-v_{i-1,j+1,m,k}-v_{i+1,j-1,m,k}+v_{i-1,j-1,m,k}}{4h_1h_2}+ \quad (7)$$

$$(C_{1133}+C_{1313})\frac{w_{i+1,j,m+1,k}-w_{i-1,j,m+1,k}-w_{i+1,j,m-1,k}+w_{i-1,j,m-1,k}}{4h_1h_3}=\rho\frac{u_{i,j,m,k+1}-2u_{i,j,m,k}+u_{i,j,m,k-1}}{\tau^2}$$

$$C_{1212}\frac{v_{i+1,j,m,k+1}-2v_{i,j,m,k+1}+v_{i-1,j,m,k+1}}{h_1^2}+C_{2222}\frac{v_{i,j+1,m,k+1}-2v_{i,j,m,k+1}+v_{i,j-1,m,k+1}}{h_2^2}+$$

$$C_{2323}\frac{v_{i,j+1,m,k+1}-2v_{i,j,m,k+1}+v_{i,j-1,m,k+1}}{h_3^2}+(C_{2211}+C_{1212})\frac{u_{i+1,j+1,m,k}-u_{i-1,j+1,m,k}-u_{i+1,j-1,m,k}+u_{i-1,j-1,m,k}}{4h_1h_2}+ \quad (8)$$

$$(C_{2233}+C_{2323})\frac{w_{i,j+1,m+1,k}-w_{i,j-1,m+1,k}-w_{i,j+1,m-1,k}+w_{i,j-1,m-1,k}}{4h_2h_3}=\rho\frac{v_{i,j,m,k+1}-2v_{i,j,m,k}+v_{i,j,m,k-1}}{\tau^2}$$

$$C_{1313}\frac{w_{i+1,j,m,k+1}-2w_{i,j,m,k+1}+w_{i-1,j,m,k+1}}{h_1^2}+C_{2323}\frac{w_{i,j+1,m,k+1}-2w_{i,j,m,k+1}+w_{i,j-1,m,k+1}}{h_2^2}+$$

$$C_{3333}\frac{w_{i,j+1,m,k+1}-2w_{i,j,m,k+1}+w_{i,j-1,m,k+1}}{h_3^2}+\left(C_{3322}+C_{1313}\right)\frac{u_{i+1,j,m+1,k}-u_{i-1,j,m+1,k}-u_{i+1,j,m-1,k}+u_{i-1,j,m-1,k}}{4h_1h_2}+ \quad (9)$$

$$\left(C_{3311}+C_{2323}\right)\frac{v_{i,j+1,m+1,k}-v_{i,j-1,m+1,k}-v_{i,j+1,m-1,k}+v_{i,j-1,m-1,k}}{4h_1h_3}=\rho\frac{w_{i,j,m,k+1}-2w_{i,j,m,k}+w_{i,j,m,k-1}}{\tau^2}$$

Далее разрешая конечно-разностные уравнения (7)-(9) относительно $u_{i,j,m,k+1}$, $v_{i,j,m,k+1}$ и $w_{i,j,m,k+1}$ соответственно, получим

$$u_{i,j,m,k+1}=\left(C_{1111}\frac{u_{i+1,j,m,k+1}+u_{i-1,j,m,k+1}}{h_1^2}+C_{1212}\frac{u_{i,j+1,m,k+1}+u_{i,j-1,m,k+1}}{h_2^2}+\right.$$

$$C_{1313}\frac{u_{i,j+1,m,k+1}+u_{i,j-1,m,k+1}}{h_3^2}+\left(C_{1122}+C_{1212}\right)\frac{v_{i+1,j+1,m,k}-v_{i-1,j+1,m,k}-v_{i+1,j-1,m,k}+v_{i-1,j-1,m,k}}{4h_1h_2}+$$

$$\left.\left(C_{1133}+C_{1313}\right)\frac{w_{i+1,j,m+1,k}-w_{i-1,j,m+1,k}-w_{i+1,j,m-1,k}+w_{i-1,j,m-1,k}}{4h_1h_3}+\rho\frac{2u_{i,j,m,k}-u_{i,j,m,k-1}}{\tau^2}\right)/ \quad (10)$$

$$\left(\frac{\rho}{\tau^2}+\frac{2C_{1111}}{h_1^2}+\frac{2C_{1212}}{h_2^2}+\frac{2C_{1313}}{h_3^2}\right)$$

$$v_{i,j,m,k+1}=\left(C_{1212}\frac{v_{i+1,j,m,k+1}+v_{i-1,j,m,k+1}}{h_1^2}+C_{2222}\frac{v_{i,j+1,m,k+1}+v_{i,j-1,m,k+1}}{h_2^2}+\right.$$

$$C_{2323}\frac{v_{i,j+1,m,k+1}+v_{i,j-1,m,k+1}}{h_3^2}+\left(C_{2211}+C_{1212}\right)\frac{u_{i+1,j+1,m,k}-u_{i-1,j+1,m,k}-u_{i+1,j-1,m,k}+u_{i-1,j-1,m,k}}{4h_1h_2}+$$

$$\left.\left(C_{2233}+C_{2323}\right)\frac{w_{i,j+1,m+1,k}-w_{i,j-1,m+1,k}-w_{i,j+1,m-1,k}+w_{i,j-1,m-1,k}}{4h_2h_3}+\rho\frac{2v_{i,j,m,k}-v_{i,j,m,k-1}}{\tau^2}\right)/ \quad (11)$$

$$\left(\frac{\rho}{\tau^2}+\frac{2C_{1212}}{h_1^2}+\frac{2C_{2222}}{h_2^2}+\frac{2C_{2323}}{h_3^2}\right)$$

$$w_{i,j,m,k+1}=\left(C_{1313}\frac{w_{i+1,j,m,k+1}+w_{i-1,j,m,k+1}}{h_1^2}+C_{2323}\frac{w_{i,j+1,m,k+1}+w_{i,j-1,m,k+1}}{h_2^2}+\right.$$

$$C_{3333}\frac{w_{i,j+1,m,k+1}+w_{i,j-1,m,k+1}}{h_3^2}+\left(C_{3322}+C_{1313}\right)\frac{u_{i+1,j,m+1,k}-u_{i-1,j,m+1,k}-u_{i+1,j,m-1,k}+u_{i-1,j,m-1,k}}{4h_1h_2}+$$

$$\left.\left(C_{3311}+C_{2323}\right)\frac{v_{i,j+1,m+1,k}-v_{i,j-1,m+1,k}-v_{i,j+1,m-1,k}+v_{i,j-1,m-1,k}}{4h_1h_3}+\rho\frac{2w_{i,j,m,k}-w_{i,j,m,k-1}}{\tau^2}\right)/ \quad (12)$$

$$\left(\frac{\rho}{\tau^2}+\frac{2C_{1313}}{h_1^2}+\frac{2C_{2323}}{h_2^2}+\frac{2C_{3333}}{h_3^2}\right)$$

На основе полученных конечно-разностных уравнений организуем следующий итерационный процесс

$$u_{i,j,m,k+1}^{(n+1)} = \left(C_{1111} \frac{u_{i+1,j,m,k+1}^{(n)} + u_{i-1,j,m,k+1}^{(n)}}{h_1^2} + C_{1212} \frac{u_{i,j+1,m,k+1}^{(n)} + u_{i,j-1,m,k+1}^{(n)}}{h_2^2} + \right.$$
$$C_{1313} \frac{u_{i,j+1,m,k+1}^{(n)} + u_{i,j-1,m,k+1}^{(n)}}{h_3^2} + \left(C_{1122} + C_{1212}\right) \frac{v_{i+1,j+1,m,k} - v_{i-1,j+1,m,k} - v_{i+1,j-1,m,k} + v_{i-1,j-1,m,k}}{4h_1h_2} +$$
$$\left.\left(C_{1133} + C_{1313}\right) \frac{w_{i+1,j,m+1,k} - w_{i-1,j,m+1,k} - w_{i+1,j,m-1,k} + w_{i-1,j,m-1,k}}{4h_1h_3} + \rho \frac{2u_{i,j,m,k} - u_{i,j,m,k-1}}{\tau^2} \right) /$$
$$\left(\frac{\rho}{\tau^2} + \frac{2C_{1111}}{h_1^2} + \frac{2C_{1212}}{h_2^2} + \frac{2C_{1313}}{h_3^2} \right) \quad (13)$$

$$v_{i,j,m,k+1}^{(n+1)} = \left(C_{1212} \frac{v_{i+1,j,m,k+1}^{(n)} + v_{i-1,j,m,k+1}^{(n)}}{h_1^2} + C_{2222} \frac{v_{i,j+1,m,k+1}^{(n)} + v_{i,j-1,m,k+1}^{(n)}}{h_2^2} + \right.$$
$$C_{2323} \frac{v_{i,j+1,m,k+1}^{(n)} + v_{i,j-1,m,k+1}^{(n)}}{h_3^2} + \left(C_{2211} + C_{1212}\right) \frac{u_{i+1,j+1,m,k} - u_{i-1,j+1,m,k} - u_{i+1,j-1,m,k} + u_{i-1,j-1,m,k}}{4h_1h_2} +$$
$$\left.\left(C_{2233} + C_{2323}\right) \frac{w_{i,j+1,m+1,k} - w_{i,j-1,m+1,k} - w_{i,j+1,m-1,k} + w_{i,j-1,m-1,k}}{4h_2h_3} + \rho \frac{2v_{i,j,m,k} - v_{i,j,m,k-1}}{\tau^2} \right) /$$
$$\left(\frac{\rho}{\tau^2} + \frac{2C_{1212}}{h_1^2} + \frac{2C_{2222}}{h_2^2} + \frac{2C_{2323}}{h_3^2} \right) \quad (14)$$

$$w_{i,j,m,k+1}^{(n+1)} = \left(C_{1313} \frac{w_{i+1,j,m,k+1}^{(n)} + w_{i-1,j,m,k+1}^{(n)}}{h_1^2} + C_{2323} \frac{w_{i,j+1,m,k+1}^{(n)} + w_{i,j-1,m,k+1}^{(n)}}{h_2^2} + \right.$$
$$C_{3333} \frac{w_{i,j+1,m,k+1}^{(n)} + w_{i,j-1,m,k+1}^{(n)}}{h_3^2} + \left(C_{3322} + C_{1313}\right) \frac{u_{i+1,j,m+1,k} - u_{i-1,j,m+1,k} - u_{i+1,j,m-1,k} + u_{i-1,j,m-1,k}}{4h_1h_2} +$$
$$\left.\left(C_{3311} + C_{2323}\right) \frac{v_{i,j+1,m+1,k} - v_{i,j-1,m+1,k} - v_{i,j+1,m-1,k} + v_{i,j-1,m-1,k}}{4h_1h_3} + \rho \frac{2w_{i,j,m,k} - w_{i,j,m,k-1}}{\tau^2} \right) /$$
$$\left(\frac{\rho}{\tau^2} + \frac{2C_{1313}}{h_1^2} + \frac{2C_{2323}}{h_2^2} + \frac{2C_{3333}}{h_3^2} \right) \quad (15)$$

Выбрав начальные значения для $u_{i,j,m,k+1}^{(0)}$, $v_{i,j,m,k+1}^{(0)}$ и $w_{i,j,m,k+1}^{(0)}$ во внутренних точках рассматриваемой трёхмерной области, найдём значения искомых функций во внутренних точках пользуясь формулами (13)-(15). Итерационный процесс продолжается до достижения требуемой точности.

В разностных уравнениях (7)-(9) была использована полностью неявная схема, кроме этого могут быть использованы другие неявные конечно разностные уравнения.

На основе выше приведённых итерационных формул разработано программное обеспечение и получены численные результаты.

Литература

1. *Победря Б. Е. Численные методы в теории упругости и пластичности. -М.: МГУ, 1996. - 343с.*
2. *Khaldjigitov A.A., Nik M.A.Asri Long., Qalandarov A.A., Eshquvatov Z. Mathematical and numerical modelling of the thermoplastic coupled problem. International conference on mathematical sciences and statistics 2013. Singapore, Springer, pp. 69-75, (2014).*
3. *Qalandarov A.A., Khaldjigitov A.A. Mathematical and numerical modeling of the coupled dynamic thermoelastic problems for isotropic bodies. TWMS J. Pure Appl. Math. V.11, N.1, 2020, pp.119-126*

ULTRAISO DASTURIDA YUKLOVCHI FLASH DISK YARATISH

Matkarimov J.A, Mirzaaxmedov M.K., Abduvaliyev A.A., Andijon davlat universiteti, O'zbekiston.

Annotatsiya:*Ushbu ishda fayllarni arxivlovchi UltraISO dasturining flash disklarga xizmat qiluvchi imkoniyatlari va unda ishlash bo'yicha nazariy va amaliy ko'rsatmalar berilgan. UltraISO dasturi orqali flash diskni yuklovchi diskga o'tkazish ketma-ketligi yoritib berilgan.*

Kalit so'zlar: *UltraISO, .iso, .isz, .bin, .cue, .nrg, .mdf, .mds, .img, .ccd, .sub, HDD, ISO, flash disk, format, USB.*

Annotation:*In this work provides theoretical and practical instructions on how to use UltraISO software for flash drives.*

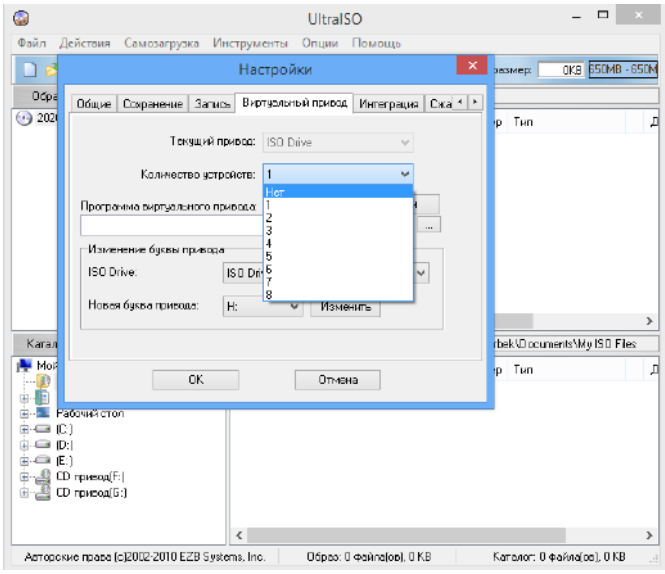
Keywords: *UltraISO, .iso, .isz, .bin, .cue, .nrg, .mdf, .mds, .img, .ccd, .sub, HDD, ISO, flash disk, format, USB.*

Аннотация: *В этой статье описываются возможности UltraIso, программы архивирования файлов, которая обслуживает флеш-накопители, и даются теоретические инструкции по работе с ней.*

Ключевые слова: *UltraISO, .iso, .isz, .bin, .cue, .nrg, .mdf, .mds, .img, .ccd, .sub, HDD, ISO, flash disk, format, USB*

Bugungi kunda kompyuterlarga compact disklardan operatsion tizimni o'rnatish sezilarli darajada pastlagan, sababi foydalanuvchilar zamonaviy kompyuterlarga operatsion tizimni o'rnatishni flash disklardan tez va qulay amalga oshirishmoqdalar.

Foydalanuvchilar orasida UltraISO dasturi keng ommalashgan. Bu dastur foydalanuvchilarga bir qancha imkoniyatlarni yaratib beradi, dastur yordamida nafaqat flash disklarni yuklanuvchi qilish, balki compact diskni yuklanuvchi xolatini ko'chirib olish, compact disklarga ma'lumotlarni yozish, fayllarni arxivlash va boshqa imkoniyatlari. Bir nechta arxivlangan fayllar bilan ishlash uchun xayoliy (Виртуальный привод) disk yurituvchi yaratib olish mumkin.

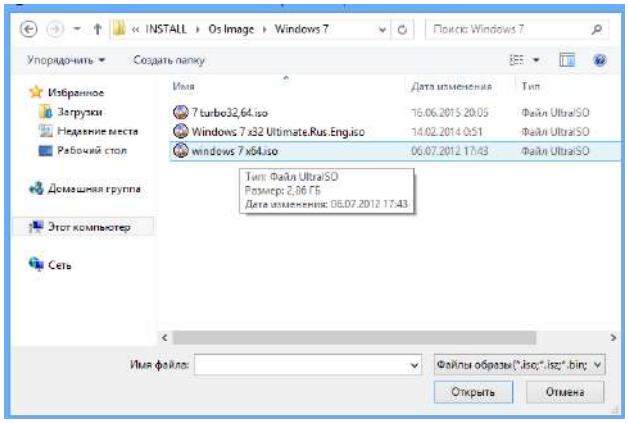


Dastur odatda .iso, .isz kengaytmalari bo'yicha arxivlaydi, lekin yana bir necha kengaytmalar ko'rinishida xam arxivlash va bu kengaytmalarni qo'llab-quvvatlashi xam mumkin (.bin, .cue, .nrg, .mdf, .mds, .img, .ccd, .sub).

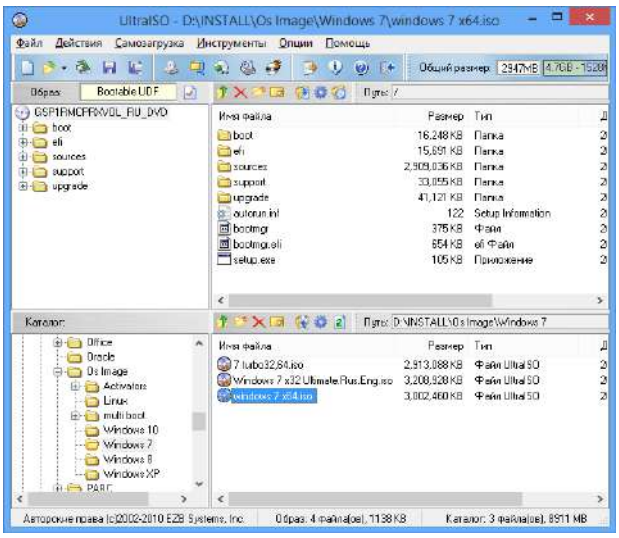
UltraISO dasturida yuklanuvchi flash disklarni yaratish oddiy va tezkor xisoblanadi. Yaratilgan yuklanuvchi flashka kompyuter va noutbuklarga o'rnatilishi mumkin, agarki flash disk virusdan zararlanib qolmasa.

Dastur yordamida operatsion tizimni nusxasidan (obrazidan) (Windows 7,8,10, Linux,…) yuklanuvchi flashkani yaratishingiz mumkin.

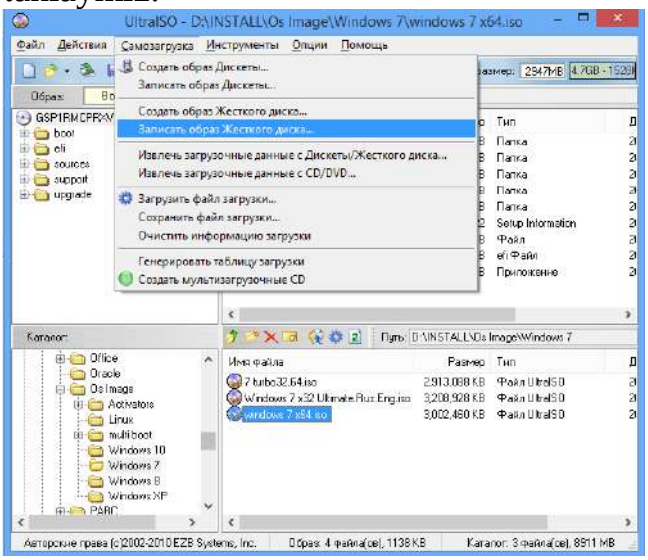
Flashkani yuklanuvchi qilishni Windows 7 operatsion tizimi misolida ko'rib chiqamiz. Buning uchun sizda operatsion tizimni yuklanuvchi xolatda olingan nusxasi (obrazi) bo'lishi kerak. Dasturni ishga tushiramiz, menyudan (Файл-открыть, Ctrl+O) tugmasini tanlaymiz va faylni ochamiz.



Fayl ochilganidan so'ng dasturni asosiy oynasida obrazga kiruvchi barcha fayllarni ko'rishingiz mumkin.

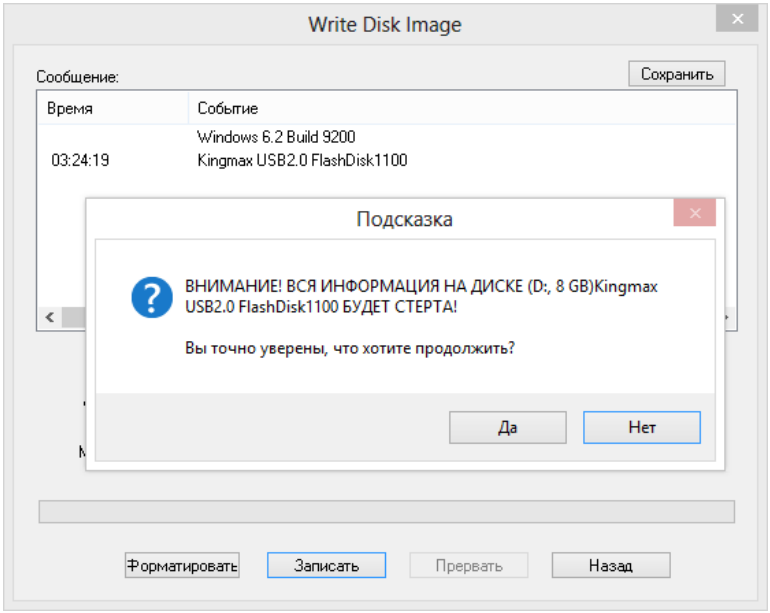


Dasturning bosh menyusida (Самозагрузка→Записать образ Жесткого диска… buyrug'ini tanlaymiz.



Write Disk Image oynasi ochiladi. Disk Drive maydonida yozish uchun mo'ljallangan flash disk yoki yana boshqa flash disklar ko'rsatiladi. Demak qaysi flash diskka operatsion tizimni yozishni tanlab olamiz. Aks xolda boshqa flash disk aktiv turgan bo'lsa, yozish oldidan flash diskni dastur formatlab tashlaydi.

Yozish usuli odatdagiday bo'lgani ma'qul. **USB-HDD+**. **Записать** tugmasini tanlaymiz.



Dastur ogoxlantiradi fleshkadagi barcha ma'lumotlar o'chirilishi haqida oyna paydo bo'ladi, va ISO obrazidan yuklanuvchi fleshkaning yozilishi boshlanadi, ushbu jarayon esa foydalanuvchidan bir necha daqiqa kutishni talab etadi.

Adabiyotlar

1. *Konstantina Matsoukas, Anca Meret, John T. Oliver, and Michael Purcell. Customized USB Flash Drives Used to Promote Library Resources and Services to First-year Medical and Dental Students.*

2. *Mills Elinor. USB Devices Spreading Viruses. [Accessed: December 2, 2010];CNET News. 2008 November 20; Available: < http://news.cnet.com/8301-1009_3-10104496-83.html>.*

MAKTABGACHA TA’LIM MUASSASALARI TARBIYALANUVCHILARININGMATEMATIK BILIMLARINI RIVOJLANTIRUVCHI ELEKTRON KONTENTLAR YARATISH

Absobirov S.Q. Navoiy davlat pedagogika instituti, O’zbekiston.

Annotatsiya: *Ushbu maqolada maktabgacha ta’lim muassasalari tarbiyalanuvchilarining matematik bilimlarini rivojlantirishga xizmat qiluvchi elektron ta’lim resurslarini yaratish va undan foydalanish jarayoni misollar yordamida ochib berilgan. Ushbu elektron ta’lim resurslarini yaratishda Flash dasturidan foydalanilgan. Maktabgacha ta’lim muassasalari tarbiyalanuvchilarining matematika bo‘yicha bilimlarini tekshiruvchi va rivojlantiruvchi dasturlar yaratilgan.*

Kalit so‘zlar: *Flash, Macromedia, Effekt, Web, Komponent, Tool, Dynamic Text, Input.*

Аннотация: *Аннотация: в данной статье описывается процесс создания и использования ресурсов электронного обучения для развития математических знаний дошкольников на примерах. Flash был использован для создания этих электронных учебных ресурсов. Программы дошкольного образования предназначены для проверки и развития знаний учащихся по математике.*

Ключевые слова: *Flash, Macromedia, Effect, Web, Component,* Tool, Dynamic Text, Input.

Annotation: *This article describes the process of creating and using e-learning resources to develop the mathematical knowledge of preschool learners through examples. Flash program was used. Programs for checking and developing mathematical knowledge of preschool learners’ were created.*

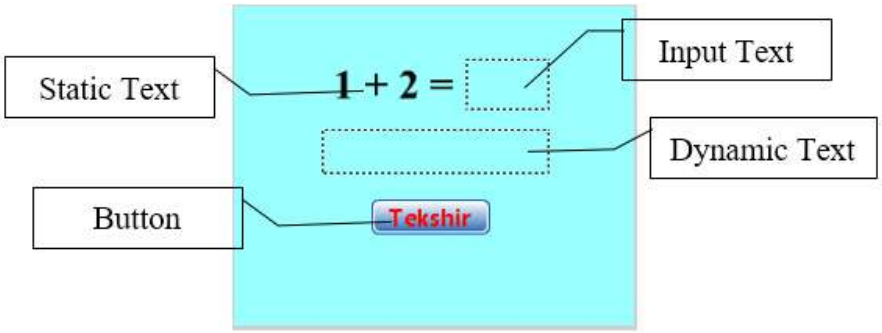
Key words: *Flash, Macromedia, Effect, Web, Component, Tool, Dynamic, Text, Input.*

Flash dasturi Macromedia kompaniyasi tomonidan ishlab chiqarilgan bo‘lib, u o‘quv materiallarini animatsiya effektlariga tayangan holda, o‘qitiladigan fanlar uchun dinamik ko‘rgazmali o‘quv qurollarini yaratishda juda qulay va keng imkoniyatlar yaratadi. Bugungi kunda jahonda ko‘plab foydalanuvchilar va dasturchilar Flash texnologiyasidan chiroyli bo‘lgan Web resurslar yaratishda foydalanadi. Dizaynerlarni Flashda grafika yaratishning yangi imkonyatlari o‘ziga tortadi, professional darajadagi yaratuvchilar bo‘lsa, o‘zlari istagancha scriptlar, shakl va server imkoniyatlaridan foydalanib ilovalar yaratish mumkin.

Biz quyida Flashning ActionScript dasturlash tili imkoniyatlaridan foydalangan holda maktabgacha ta’lim muassasalari tarbiyalanuvchilarining matematik bilimlarini tekshiruvchi amaliy dasturlar yaratish jarayonini ko‘rib chiqamiz.

Flash dasturini ishga tushiramiz. Ishchi oynaga 1 ta Button tugmasini hosil qilamiz. Text Tool uskunasi yordamida izohlar kiritamiz va 1 ta Input Text hamda 1

ta Dynamic Text yaratamiz. Input Text yaratish uchun Text Tool uskunasi yordamida matn sohasi hosil qilinadi. Properties oynasidan Text type bo‘limiga Input Text moslanadi. Dynamic Text yaratish uchun ham yuqoridagidek matn sohasi hosil qilamiz va Properties oynasidan Text type bo‘limiga Dynamic Text moslanadi. Input Textni “son” deb va Dynamic Textni esa “natija” deb nomlaymiz. Ya’ni Properties oynasining Variable bo‘limiga mos holda “son” va “natija” jumlalarini kiritamiz.

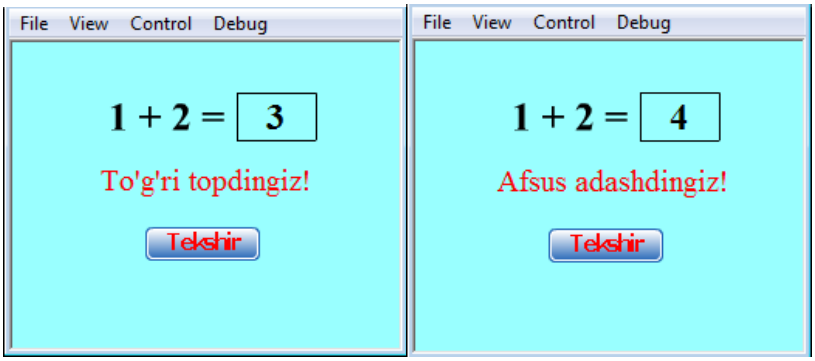


1-rasm.

Komponentalarni 1-rasmdagi kabi joylashtirib olamiz. Endi esa dasturlash qismiga o‘tamiz. Button tugmasini belgilab F9 tugmasini bosamiz. Hosil bo‘lgan Actions-Button oynasi dasturlash maydoniga quyidagilarni kiritamiz:

```
on (release)
{       a=son;
        if (a==3) natija = "To'g'ri topdingiz!"
        else    natija = "Afsus adashdingiz!";  }
```

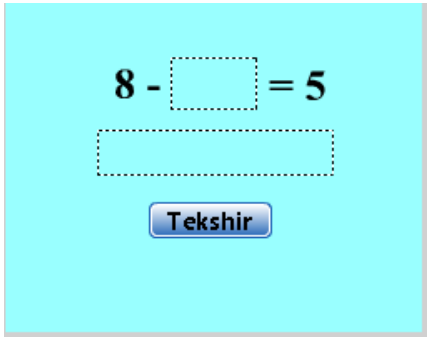
Faylni kompyuter xotirasiga saqlaymiz va Ctrl+Enter tugmalarini bosib ishga tushiramiz. Input Textga 3 sonini kiritsak “To‘g‘ri” so‘zi boshqa sonlarni kiritsak “Xato” so‘zi paydo bo‘ladi.



2-rasm.

Fayl menyusining Publish Settings... bo‘limi orqali yaratilgan loyihani *.html yoki *.exe fayl kengaytmada saqlaymiz.

Yuqoridagi formani nusxalaymiz va misol ko‘rinishini o‘zgartiramiz.

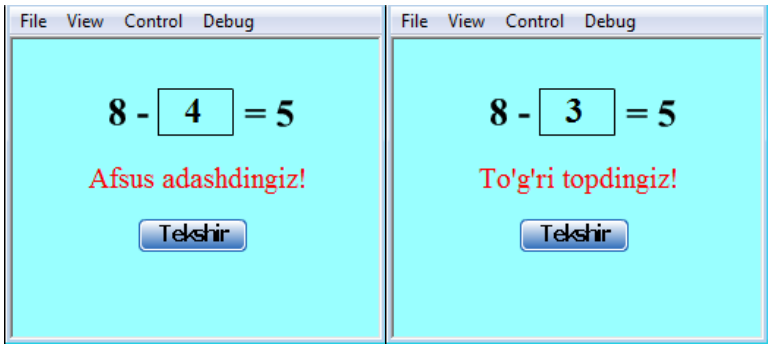


3-rasm

Button tugmasinining Actions-Button oynasiga quyidagilarni kiritamiz:

```
on (release)
{       a=son;if (a==3) natija = "To'g'ri topdingiz!"
        else    natija = "Afsus adashdingiz!";  }
```

Faylni kompyuter xotirasiga saqlaymiz va Ctrl+Enter tugmalarini bosib tekshirib koʻramiz.



4-rasm.

Shu tartibda bir nechta dasturlar yaratamiz.

Flash Shock Wave Flash formatli vektorli grafikdan foydalanishga asoslangan texnologiyadir. ShockWave Flash formati oʻzida grafika, animatsiya, Action Script tili scriptlarini, bundan tashqari rastirli grafika, video, ovoz va Flashga import qilinganlarni optimallashtirilgan va siqilgan holda saqlaydi. Ushbu formatni juda koʻplab grafik paketlar va video bilan ishlovchi dasturlar qoʻllab quvvatlaydi. Flash hujjatni Shock Wave Flash formatida saqlash uchun eksport yoki nashr etish (публикации) amalini bajarish zarur.

Adabiyotlar

1. *Алберт Д.И., Алберт Е.Э. Самоучитель Macromedia Flash Professional 8. – СПб.: БХВ-Петербург, 2006. – 736 с.: ил.*
2. *Дронов В.А. Macromedia Flash Professional 8. Графика и анимация. – СПб.: БХВ-Петербург, 2006. – 656 с.: ил.*

ДАСТУРЛАШ ТЕХНОЛОГИЯСИДАН ТАЪЛИМ ВОСИТАСИ СИФАТИДА ФОЙДАЛАНИШНИНГ БАЪЗИ БИР МАСАЛАЛАРИ

Маҳкамов М.К., Андижон давлат университети, Ўзбекистон.

Аннотация. *Мазкур иш сонларга оид масалани ечиш алгортми ва дастурини тузишга бағишланган бўлиб, бу жараёнда талабаларнинг билиш фаоллигини оширишга оид тавсиялар берилган*

Калит сўзлар: *моделлаштириш, математик моделлар, дастур, натижа, таҳлил*

Аннотация. *Данная работа посвящена разработке алгоритмов и программ для решения численных задач, в процессе которых даются рекомендации по повышению познавательной активности студентов*

Ключевые слова: *моделирование, математические модели, программа, результат, анализ*

Annotation. *This work is devoted to the development of algorithms and programs for solving numerical problems, in the course of which recommendations are made to increase the cognitive activity of students*

Keywords*: modeling, mathematical models, program, result, analysis*

Ҳисоблаш ва маълумотларга асослауган фан – тадқиқотнинг учинчи устуни сифатида назария ва тажриба ўртасини боғловчи муҳим воситадир. Шу билан бир қаторда, таълимнинг барча соҳаларини, илмий изланишларни, ишлаб чиқариш амалиётини, шунингдек, иқтисодиётни тубдан ўзгартириш имкониятига эга. Компьютер моделлари системанинг табиатини ўзига хос турли хил комбинацион математик моделларга асосланган яқинлашувлар ёрдамида тасвирлаб беради. Компьютерда моделлаштириш ёрдамида тажриба қилмасдан аввал маълум бир тизимлар учун қулай моделни тузиб чиқиш ва синаб олиш тажриба учун кетадиган вақт ва ҳаражатларни сезиларли даражада камайтиради. Моделлар ёрдамида тажриба учун сарфланадиган вақт ва ҳаражатларни тежаш ўз навбатида катта иқтисодий самарадорликка ҳам олиб келади. Компьютерда моделлаштириш талабаларни тадқиқотларга ундайди, уларда тадқиқот ва изланишга қизиқиш уйғотади. Талабалар турли хил параметрларнинг қийматларини ўзгартириб, уларнинг таъсирини кузатиш орқали илмий тушунчалар ва ғоялар тўғрисида ўз хулосаларини чиқарадилар. Талабаларни компьютерли моделлаштиришга, дастурлаш технологиясига жалб қилиш уларнинг билим олишини яхшилашда унинг афзалликларини ошириш учун хизмат қилади. Интерфаол компьютерли моделлаштириш талабаларга ўзларининг кашфиётларини бошқариш ва эгалик қилиш ҳиссини беради ва шу билан уларнинг маълумотларни тушуниш ва таҳлил қилиш қобилиятини оширади, мисолларни қайта тасаввур қилиш имкониятини беради [1].

Ҳозирги замон тажрибасини ҳисобга олган ҳолда айтиш мумкинки, Ўзбекистондаги барча ОТМлар компьютерда моделлаштириш ва дастурлаш технологияси йўналишларини ривожлантиришни жиддий ўйлаб кўришлари керак. Ўзбекистон дунёга бўй кўрсатаётган бир вақтда яқин 4-5 йил ичида компьтерда моделлаштириш мутахассисларига эҳтиёж сезиларли даражада ортади. Ушбу талабга мувофиқ ҳолда профессионал кадрлар тайёрлашни тизимли йўлга қўйиш мақсадга мувофиқ.

Қуйидаги содда мисолни кўриб чиқамиз:

Икки хонали сон билан худди шу сондаги рақамларнинг тескари тартибда ёзилиши билан пайдо бўлган соннинг йиғиндиси тўла квадратни берадиган барча сонларни топамиз [2].

```
Var a,b,k:integer;
Begin
 writeln('Qidirilaytgan ikki xonali sonlar');
  for a:=1 to 9 do
   for b:=1 to 9 do
    Begin
     k:=round(sqrt(10*a+b+10*b+a));
     if k*k=10*a+b+10*b+a then
     write(10*a+b:4);
    end
 end.
```

Талабаларга дастур натижасини олгандан сўнг round фнкциясини trunc фнкциясига алмаштириш ва ҳосил бўлган натижаларни текширишни тавсия қилинади.

Тузилган дастур натижасидан кўриниб турибдики, тузилган дастурдан фойдаланиш – таълим сифатини ошириш билан бирга вақтни тежаш, ҳисоблашлар муддатини қисқартириш, меҳнат сарфини тежаш, иш вақтини зое кетишини камайтириш, визуал тушунчалар ёрдамида тушунтириш оборотини тезлаштириш, тушунчалар кўламини кенгайтириш ва бошқа натижалар билан тавсифланади ҳамда юқори самарадорликка эришилади.

Ўқув жараёнида дастурлаш технологиядан таълим воситаларидан бири сифатида фойдаланиш талабалар учун ўзларини ўраб турган олам ҳақидаги билимларини ўзлаштиришларида муҳим рол ўйнайди.

Адабиётлар:

1. Ж.Раззоқов. Компьютерда моделлаштириш: бу қанчалик муҳим ва у ҳақда нималарни биламиз? https://kun.uz/58396204.

2. Ш.Х.Фозилов, С.С.Жуманазаров, Д.Т.Мухамедиева. Дастурлаш бўйича олимпиада масалалари. Тошкент.: Фан ва технологиялар, 2007.

3. Абдуқодиров А.А., Юлдошев И.А. Информатика ўқитиш методикаси. Тошкент.: Фан ва технологиялар, 2010.

PARKINSON'S DISEASE DETECTING WITH VOICE SIGNAL DATA USING MACHINE LEARNING

M.Shakeel Khan, Qaiser ali Imran Ahmed. Riphah international university lahore. Dept. Of Computer science, Lahore, Pakistan.

Abstract. *Parkinson's illness (PD) is de-generative neurologic turmoil characterize by lower levels off dopamine in the heart. Manifest through worsening movement, including shivering UK there is usually significant effects on speech, including breathing difficulties (hard to speak), throwing (volume). In addition, cognitive impairment and mood swings and the risk of dementia may occur. PD affects about 1-2% of people over 60 and most of the people are genetically affected with Parkinson's disease. In this article we are getting the voice signal dataset and analyzed the data and understand and visualized the dataset for better understanding for detecting disease and detect illness with Parkinson's and a technique that is used the machine learning algorithms like as a Random forest, Naive base, Constant, logistic regression These algorithms is clearly detect the Parkinson patients and some of non-Parkinson using for analysis of orange tool and applying the diff algorithms techniques. Hence they gives the best accuracy result of the disease is showing with the AUC, Precision, Recall .F1 of detecting the Parkinson's illness.*

Keywords: *Machine learning, logistic regression, Random forest, constant model, naïve bayes.*

INTRODUCTION

Parkinson's disease (PD) that affects neurons in the brain that causes dopamine to indicate a sign that includes muscle resistance, agitation, and changes in tone and speech. Parkinson's disease affects men's voices, [1]. PD is only the second for Alzheimer diseases in neurological disease.

The prediction performance **Substantia Nigra** is a range of pigmented cells in the middle brain that is responsible for generating the neurotransmitter dopamine.[2] This segment of the brain plays an important role in the control of movement,[5] and seems to interact additionally in the path of dependence, too[6]. When you see a brain in the movement section, the nigra substance is placed in the center of the brain, just at the top of the brains [8] we have also completed a problem analysis, which can be completed with the final and quantitative results.

Machine learning Algorithms

I will use the four different types of the algorithms they are gives the best performance in this research work, now showing below these algorithms in methods like as logistic regression, random forest, constant model, naïve bayes.

lITERATURE REVIEW

(Frid, Kantor, Svechin, & Manevitz, 2016) In this article, we teach patient diagnosis information in Parkinson's disease (BPP) and design version class/prediction to make the disease easier but Parkinson's disease research has

been fully exploited in the field, and we are able to create a platform that is not in compliance with your laws [7].

(Muthumanickam S1, 2018) Parkinson's disease its neurons degenerative factors that affect the development of neurons and dopamine in black substance.The main algorithms of automatic training and vocational training and general application in the treatment of Parkinson's diseases[10].

(C. Salvatorea & G. Arabia e, 2014) this article, which focuses on determining the possibility of Parkinson's disease (BP) using a control machine to help diagnose patients with a medical diagnosis of disease. Method: Morphologic T1 Magnetic Goose (MRA) of patients with BPC (28), patients with telephone and healthy control (28) used using learning algorithm based on a combination of joint analysis of the main component [5].

(Hashmi, 2013) detect the application of machine-learning algorithms for medical data.The overall learning algorithms to copy the results, methods of data stability, including the trend of processing in the form of unmarked data, the data-mid-don't travail training point [8].

RESEARCH QUESTIONS

During the Parkinson's disease treatment the complication to understand the report because this report is based on the voice signal frequencies and difficult to classified the report. These issues resolve with the ML.

7) Detecting the patients who are affected with Parkinson's disease?
8) How to understand the voice signal datasets and with graphically?

METHODOLOGY

Inside our paper, we be the aiming to concern the Machine learning models in arrange to predict the class of health of the patients in Parkinson dataset (getting from Kaggle site) (to is, if someone is suffer from Parkinson or none) with the analysis in orange tool this is the best tool as compared to python [1].So next to with categorization, we will be as well look into the nearly all important features with the intention of are dependable for making the choice and check condition the newer attribute are healthier than the conventional ones.

1.1Data understanding

Parkinson illness data is in the appearance of voice frequencies and data analysis and visual form on the orange tools. Now I'm describe with visualized the datasets f0r better understanding.

Description	Feature Label	Min.	Max.	Mean	Std. Dev.
Average vocal fundamental freq.	MDVP:F0(Hz)	88.33	260.11	154.23	41.39
Max. vocal fundamental freq.	MDVP:Fhi(Hz)	102.15	592.03	197.11	91.50
Min. vocal fundamental freq. frequency	MDVP:Flo(Hz)	65.48	239.17	116.33	43.52
Several measures of variation in fundamental frequency	MDVP:Jitter(%)	0.002	0.033	0.006	0.005
	MDVP:Jitter(Abs)	7E-06	26E-05	4.4E-05	3.48E-05
	MDVP:RAP	0.001	0.021	0.003	0.003
	MDVP:PPQ	0.001	0.020	0.003	0.003
	Jitter:DDP	0.002	0.064	0.010	0.009
Several measures of variation in amplitude	MDVP:Shimmer	0.01	0.119	0.03	0.019
	MDVP:Shimmer(dB)	0.085	1.302	0.282	0.195
	Shimmer:APQ3	0.005	0.056	0.016	0.010
	Shimmer:APQ5	0.006	0.079	0.018	0.012
	MDVP:APQ	0.007	0.138	0.024	0.017
	Shimmer:DDA	0.014	0.169	0.047	0.030
Two measures of ratio of noise to tonal components in the voice	NHR	0.001	0.315	0.025	0.040
	HNR	8.441	33.047	21.886	4.426
Two nonlinear dynamical complexity measures	RPDE	0.257	0.685	0.499	0.104
	D2	1.423	3.671	2.382	0.383
Signal fractal scaling exponent	DFA	0.574	0.825	0.718	0.055
Three nonlinear measures of fundamental frequency variation (Last one, PPE, is the proposed measurement of dysphonia)	Spread1	-7.965	-2.434	-5.684	1.090
	Spread2	0.006	0.450	0.227	0.083
	PPE	0.045	0.527	0.207	0.090

Figure 9:

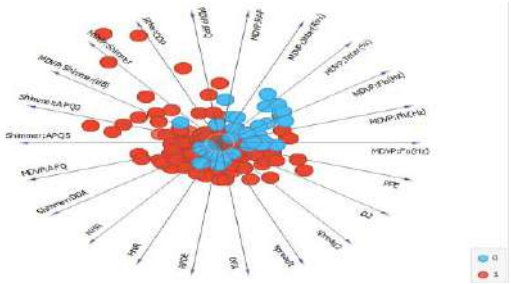


Figure 2:

In this tables in which showing the dataset description, features and min value, max values, mean values standard deviation values these are labels is a voice frequency data.

Now visualize the data set in free-vizz chart form the chart showing bellow. This fig-3 free-vizz chart is visualized the voice signal data frequencies and showing the red and blue circle of Parkinson disease.

1.2Analysis

In this analysis imp detecting the Parkinson disease patient and non-Parkinson disease patients to-count the average and percentage and predict with the machine learning algorithms. Before the detecting the results, we see the causes of the Parkinson disease effect on the brain showing bellow the figure.

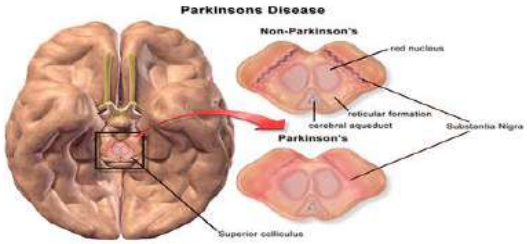


Figure 3:

This is medical images in which the mid-brain is effect with the line of **substantia Nigra** is the main cause of the Parkinson disease

a) Random forest model

It is This model is most powerful functionality of the others machine learning algorithms and the prediction performance to good and visualize the results now, I'm showing you the calculation result in the form of the charts and confusion matrix form.

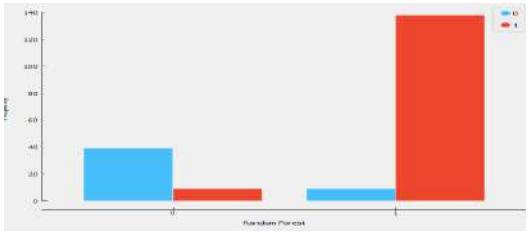


Figure 4:

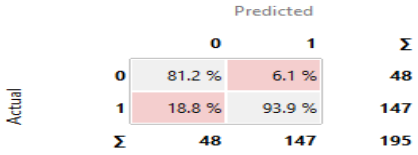


Figure 5:

Fig no 5: In this graph in which showing the prediction result of the random forest in which (0) is reached on 40% and (1) red is effected with disease are 147.

Fig no 6: In this matrix is clearly tells us the 48(0) 18%is non Parkinson and 147(1) 81 and 93%.9 is Parkinson its showing the actual prediction value in matrix forms**.**

b) Logistic regression model

Logistic regression is a machine learning algorithm used for rating problems; it is a prediction analysis algorithm and is based on the concept of possibilities.

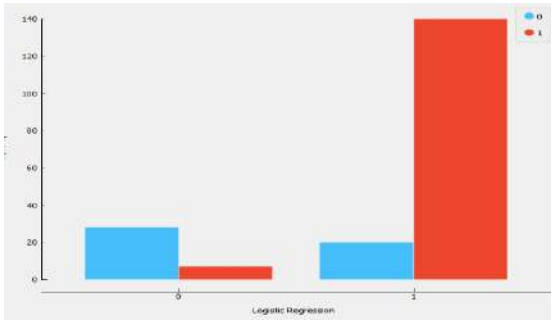


Figure 6:

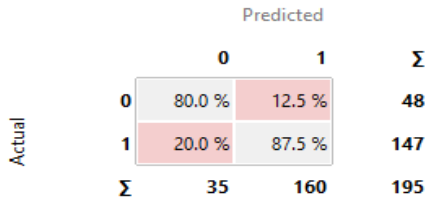


Figure 7:

In fig-7 this graph in which showing the prediction result of the Logistic regression in which (0) is reached on 35% and (1) red is effected with disease is 147.

In this figure 8 in which confusion matrix is and logistic regression in which clearly showing the prediction value is (0) are 35 not effected patients and (1) positive report 160 is Parkinson disease**.**

c) Constant model

This learner produces a model that always predicts the majority for classification tasks mean value for regression tasks.

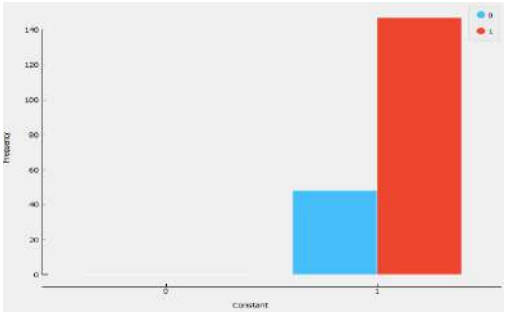


Figure 8:

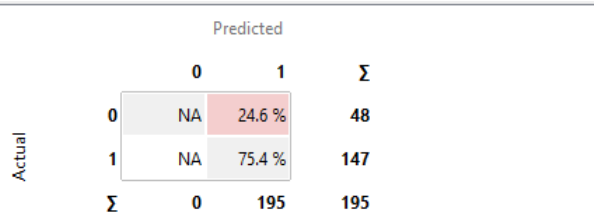


Figure 9:

In fig-9 these graphs is constant model it's totally diff from other three algorithms. In which detect only (1) are Parkinson disease effect patients. In this fig-10 is showing the confusion matrix of constant model showing the predict result its only show the (1) form of results.

d) Naïve bayes model

In machine learning, naïve Bays are a family of simple "probabilistic-classifiers" based on applying Bays'-theorem with strong-independence assumptions between the features.

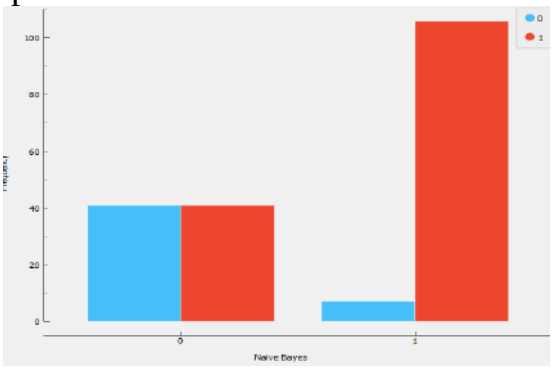


Figure 10:

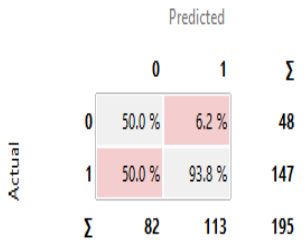


Figure 11:

In this graph is showing the naïve bayes results is gives 48(0) and 147 is (1) zero is negative and 1 is positive report.

1.3 Models comparison

Now I'm showing you the all models results and tells which the model is best predictions results the table is showing bellow.

Model	AUC	CA	F1	Precision	Recall
Random Forest	0.945	0.908	0.908	0.908	0.908
Logistic Regression	0.888	0.862	0.854	0.857	0.862
Constant	0.436	0.754	0.648	0.568	0.754
Naive Bayes	0.879	0.754	0.770	0.830	0.754

Figure 10:This is the comparison table in which clearly showing the accuracy results 1st random forest the classification accuracy is 90% this is highest accuracy result and prediction then other models logistic and naïve byes gives the best result then constant model. This table clearly visible the all models of F1-values precision recall all of this values random forest is top on the list of these algorithms.

results

The detecting pattern is a linear pattern in which the label of the class or target is categorical. This variable has two categories, yes/no /0 or 1. For this dataset on Parkinson's disease, the category will be whether the person is suffering from Parkinson's disease or if the person is a non-Parkinson's disease. Random forest and logistic regression used for the of target variables**.**1st of the" Random forest" is gives the best accuracy of the Parkinson's disease (Accuracy 94.45%**)** And 2nd a logistic regression model it gives less accuracy result then random, (Accuracy**:** 0.888**%)** and 3rd CONSTANT model is (is Accuracy: 0.43%) and 4thNaive bays model is (Accuracy: 0.871794) in all algorithms in which the highest accuracy is the (Random forest AND logistic regression classifier).

conclusion/future work

In this paper we are using machine learning algorithms for Parkinson's disease detecting the patient those who are effect with Substantia Nigra and we find the best accuracy of the state of PD provides the right to the disease. Hence these analysis and results it's completely showing and proves the Parkinson disease effected patient and better understand for dataset. Future its strategies are very technical's and impressive and efficient for the detecting the diseases and most of the people did not know they were effect with Parkinson disease and there brain is slowly work with the time and thy lot of memory and ye didn't what happened to me so the this analysis is work for the every field of the life because this analysis is very strong and accurate for the future.

REFERENCES

[1] Amit Kumar Patra1, R. R., Azian Azamimi Abdullah3 and Satya Ranjan Dash4. (2019). Prediction of Parkinson's disease using Ensemble Machine

Learning classification from acoustic analysis. Journal of Physics: Conference Serie, 1372 (2019) 012041.

[2] Benba, A., Jilbab, A., & Hammouch, A. (2016). Discriminating between patients with Parkinson's and neurological diseases using cepstral analysis. IEEE transactions on neural systems and rehabilitation engineering, 24(10), 1100-1108.

[3] De Vos, M., Prince, J., Buchanan, T., FitzGerald, J. J., & Antoniades, C. A. (2020). Discriminating progressive supranuclear palsy from Parkinson's disease using wearable technology and machine learning. Gait & Posture, 77, 257-263.

[4] Amit Kumar Patra1, R. R., Azian Azamimi Abdullah3 and Satya Ranjan Dash4. (2019). Prediction of Parkinson's disease using Ensemble Machine Learning classification from acoustic analysis. Journal of Physics: Conference Serie, 1372 (2019) 012041.

[5] C. Salvatorea, A. C. b., I. Castiglioni c,∗, F. Gallivanonec, A. Augimeri b, M. Lopez d,, & G. Arabia e, M. M. e., M.C. Gilardi c, A. Quattrone b,. (2014). Machine learning on brain MRI data for differential diagnosis of

[6] Parkinson's disease and Progressive Supranuclear PalsyC. Salvatorea. Journal of Neuroscience Methods 222 (2014, 230-237.

[7] Frid, A., Kantor, A., Svechin, D., & Manevitz, L. M. (2016). Diagnosis of Parkinson's disease from continuous speech using deep convolutional networks without manual selection of features. Paper presented at the 2016 IEEE International Conference on the Science of Electrical Engineering (ICSEE).

[8] Hashmi,S. F. (2013). A Machine Learning Approach Diagnosis of Parkinson's Disease Claremont Colleges Scholarship @ Claremont, 789.

[9] Khan, M. M., Chalup, S. K., & Mendes, A. (2016). Parkinson's disease data classification using evolvable wavelet neural networks. Paper presented at the Australasian Conference on Artificial Life and Computational Intelligence.

[10] Muthumanickam S1,G. J., Eunice Daphne J3(2018). Parkinson's Disease Detection And Classification Using Machine Learning And Deep Learning Algorithms– A Survey

[11] Muthumanickam International Journal of Engineering Science Invention (IJESI) ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726 www.ijesi.org |, |Volume 7 Issue 5 Ver. 1 || May 2018 || PP 56-63

[12] Polat, K., & Nour, M. (2020). Parkinson Disease Classification using one against all based data sampling with the acoustic features from the speech signals. Medical Hypotheses, 109678.

CUSTOMER SEGMENTATION USING CLUSTERING ALGORITHMS AND MLP IN WEKA

Khadija Tahir Riphah. International University Lahore, Pakistan.

Abstract. *Recently Customer Segmentation and market basket analysis becomes a useful and crucial tool for e-commerce Industry. It is playing important role in online recommendation systems and Local/Global Retailer/Wholesale market. Customer Relationship Management (CRM) have Key market mechanism to manage and develop valuable relationships with long-term buyers, requires developing business intelligence. Specifically, Customer Segmentation is about making groups of customers with shared characteristics i.e. age, spending trends, location. Likewise, clustering means organizing things like so that similar type of things keep on in the same group. Due to the similarity between these two we can apply clustering algorithm. This paper is about applying top machine learning algorithms K-means Clustering, hieratical clustering and popular artificial neural network class Multi neural network for clustering mall customer data and test/compare algorithm efficiency using efficient data mining tool WEKA. It Leads to the ability to develop fast and efficient automation CRM system.*

Keywords: *Customer Segmentation, Weka, Machine Learning, PCA, K-Means Clustering, Centroid based Clustering, Hierarical clustering, Neural Networks*

. INTRODUCTION

In 1956 Smith introduced the concept of 'Customer Segmentation' that is also known as 'Market Segmentation. It stated as that Market segmentation involves viewing a Heterogenous/Divergent market as a number of smaller similar markets in response to differing preferences, idntify to the desires of consumers for more exact satisfaction of their varying wants" [1]. Customer Segmentation helps to learn and understands customer's behaviors to getting successful in customer satisfaction and Best Profit. Basically, one can do Customer segmentation by evaluating distinct types of information about customers.

There are 4 types of features to make customer segmentation, Geography based (Where they live and work), demographic based (age, religion, gender, income, education level etc), psychographic (lifestyle and personality traits) and behavior oriented (spending, consumption, usage and desired benefits) and so on. Clustering refers to dividing a dataset in a grouped set of comparable data.

Clustering algorithms are those machine learning algorithms which use unlabeled data. Centroid based and Hieratical Clustering algorithms are the most popular two types from all other types of clustering algorithms [2].

An ANN is based on a collection of artificial neurons called nodes that model loosely the neurons in a biological brain. As with the synapses in a biological brain, each connection will send a signal to other neurons. An artificial

neuron that receives a signal then processes it, and is able to signal associated neurons [3].

Artificial neural networks (ANNs) are computational networks which are inspired by biology. Among the different forms of ANNs.

In this paper, we used multilayer perceptron (MLPs) with backpropagation method for its parameters to be optimized. MLPs, the ANNs most commonly used for a Various large number of problems, based on a supervised procedure and consist of three layers: input, hidden, and output. [4].

WEKA - an open source software which gives tools to information preprocessing, usage of a few Machine Learning calculations, and representation devices so you can create AI methods and apply them to genuine information mining issues. [5]

LITERATURE REVIEW

(A. S. M. Shahadat Hossain 2017) Used Centroid based and density-based clustering algorithm for customer segmentation. After testing both algorithms concluded that Density based algorithm have extra option as compare to k-means and Density based algorithms also can be used for better customer segmentation.

(S¸ukr ¨ u Ozan ¨ ∗ and Leonardo O. Iheme∗ 2019) Used Multilayer Perception model only on psychographic data for converting manual customer segmentation to artificially Intelligent customer management system

(P. Anitha , Malini M. Patil 2019) The purpose of this approach is to apply business intelligence to identify potential customers by passing on relevant and timely facts to businesses within the retail industry. This review is mainly based on the ideas for using the RFM (Deploy, Frequency and Monetary) version and the K-means algorithm for database segmentation. On the basis of silhouette calculations, a variety of database clusters are set up. Such effects are compared to different parameters according to sales transactions such as sales accuracy, sales frequency and sales volume.

(Tushar Kansal, 2Suraj Bahuguna, 3Vishal Singh, 4Tanupriya Choudhury) used internal clustering validation rather than external clustering validation, which depends on some external data like labels. Internal cluster validation can be used for choosing clustering algorithm which best suits the dataset and can correctly cluster data into its opposite cluster

(K. Maheswari 2019) Used K means clustering and find best possible number of clustering to improve algorithm performance after normalizing data set

(panelRoung-ShiunnWuaPo-HsuanChoub 2011) Introduced a soft clustering method that uses latent mixed-class membership clustering approach to classify online customers based on their purchasing data across categories. For making customer segments they used a technique from latent Dirichlet allocation model.

Problem Statement/Research Questions

Commonly customers are classified using binary classification and labels as ‘Standard’ ‘Premium’. This is the age of technology and competitive market is growing fast. Still some companies have not adopted automation tool for customer

segmentation. How Can a multinational company become more customer focused through a market segment-oriented system?

What algorithm Performance is best for the given dataset?

What is the significance of Multilayers Perception algorithm?

How much similarity found between different types of clustering algorithm?

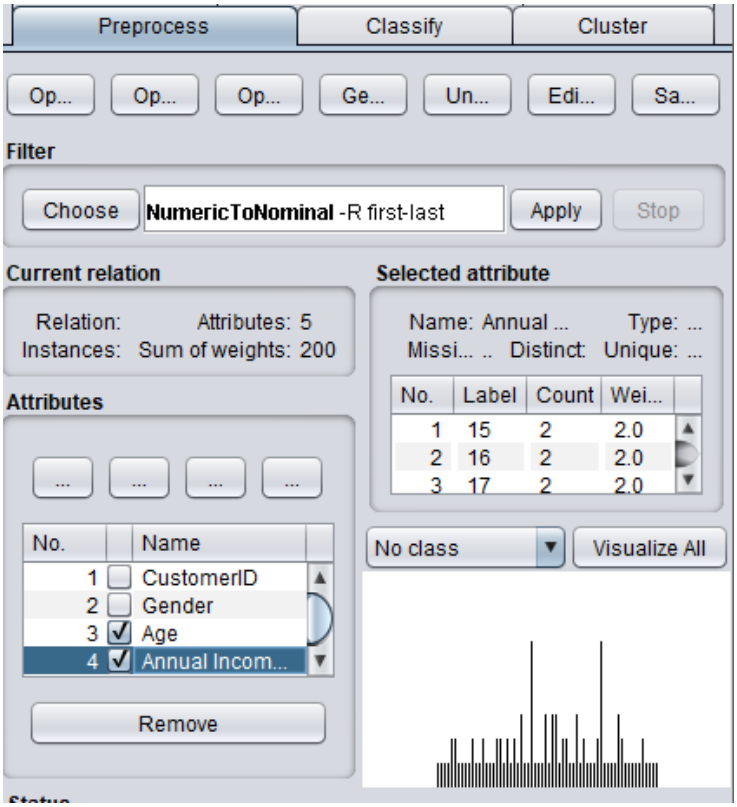
RESEARCH METHODOLOGY

Data Set

Collected Mall customer data set From Kegel. Which contains 200 data values of different customers i.e. age, spending score so on. [7]

Preprocess

Converted numeric data to nominal/ Un labeled data to perform clustering as shown in image below.



Model Selection

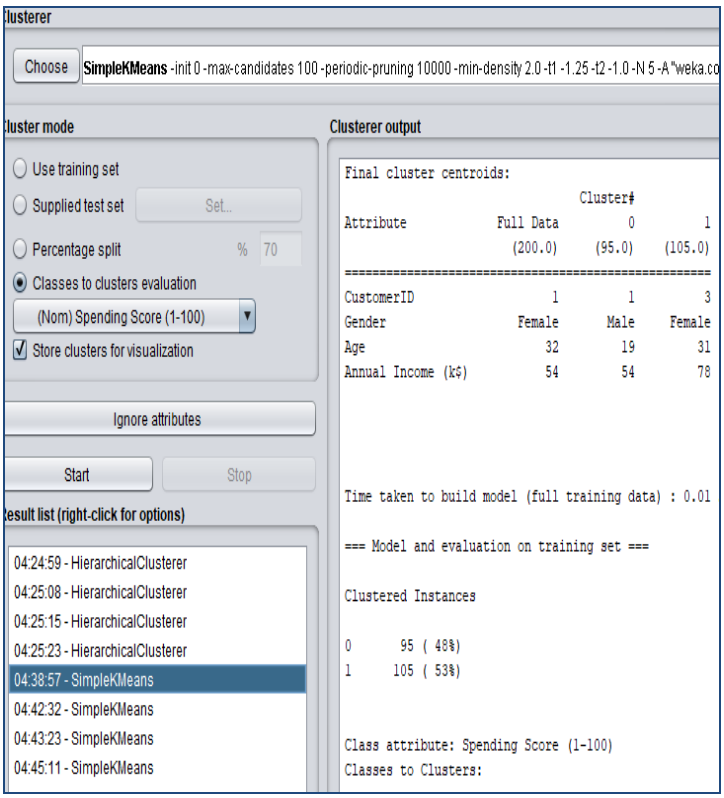
K-means

The algorithm is referred to as k-means, since the letter k represents the number of clusters chosen.

K-means functions as follows:

Firstly, Initializes k number of centroids within the data domain. Cluster data into k number of classes thorough assigning every data point to the nearest centroid Secondly Measure the mean of each cluster of objects and transfer the centroid to that location., then iterate steps 2 and 3 prior to cluster allocation the same points in a sequential iteration.

There are two ways to measure th minimum distance between any two points which are Euclidean distance, Manhattan distance, Chebyshev distance, Minkowski distance etc. [6]



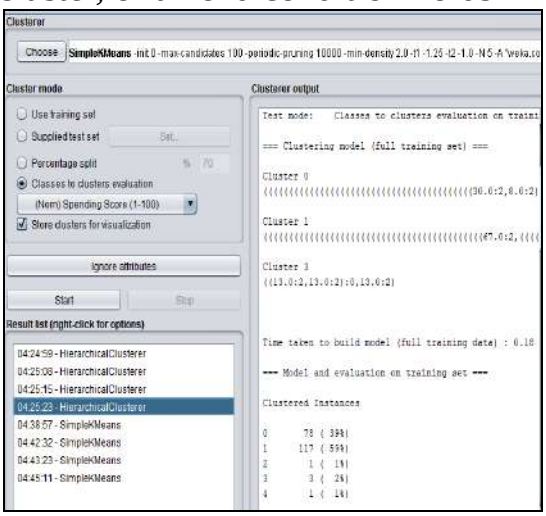
Chosen behavioral data feature i.e. spending store for making data clusters. Applied Euclidian distance method. Selected multiple number of clusters for best results.

Hieratical Clustering

The various hierarchical methods make a progressive decomposition of N cases. It very well may be classified into two classes:

1. top-down (or divisive) approach.
2. base up (or agglomerative) approach

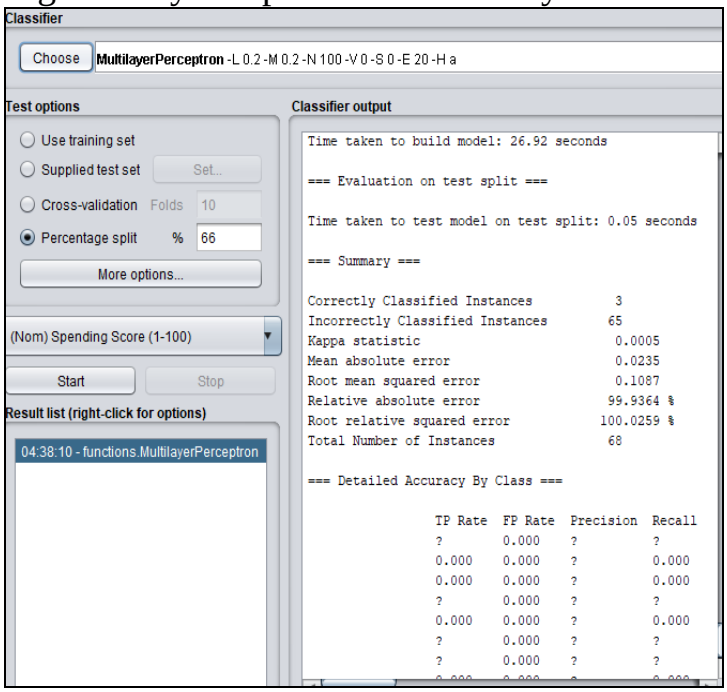
The top-down methodology begins with a single cluster having all the N cases and afterward split into littler bunches in each progressive emphasis,
until in the long run each occasion is in one bunch, or an end condition holds. The bottom up approach begins with each case framing a different group and afterward progressively combines the clusters near each other, until all the groups are converted into a single cluster, or an end condition holds



Hieratical clustering is applied using Euclidian Method. Multiple numbers of clusters chosen and applied

Multilayer Perception Algorithm

In the Multilayer perceptron, there can more than one straight layer (blends of neurons).We can build the quantity of the shrouded layer as much as we need, to make the model progressively complex as indicated by our undertaking.



RESULTS & DISCUSSION

KMeans

Distance Metrics		Number of Clusters			
		2	3	4	5
Euclidian Distance	Time Elapsed	0 seconds	0 seconds	0 seconds	0 seconds
	Sum of squared error	571.0	566.0	559.0	553.0
	No of Iterations	3	3	3	2

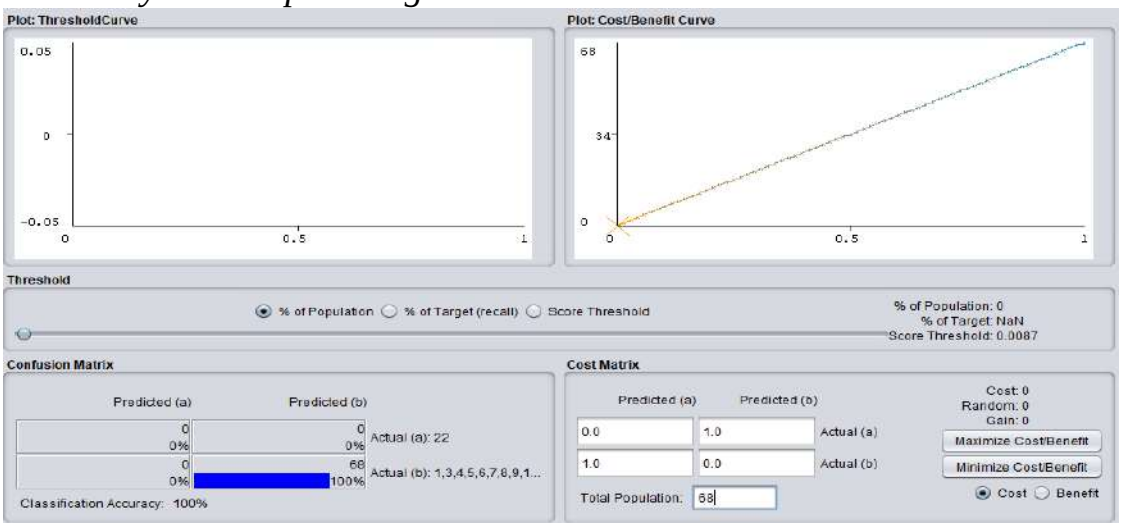
Above table Describe the number of clusters and algorithm computational Results. After applying multiple clusters concluded that 2 number of clusters are best for analysis

Hieratical Clustering

Distance Metrics		Number of Clusters			
		2	3	4	5
Euclidian Distance	Time Elapsed	0.16 seconds	0.14 seconds	0.08 seconds	0.18 seconds
	No of Iterations	2	3	2	2

Above table Describe the number of clusters and algorithm computational Results. After applying multiple clusters concluded that 2 number of clusters are best for analysis

Multilayer Perception Algorithm Results



Selected Behavioral data feature and apply algorithm, selected 5 layers and test results as shown in above Diagram

CONCLUSIONS & RECOMMENDATION

As Conclusion we achieved desired research objective of evaluation and performance analysis of centroid based clustering algorithm and Hieratical clustering. In addition, with Multilayer Perception algorithm a class of Neural Network. After Evaluation we can say that for our data set K-means clustering algorithm perform best. Weka is also can consider for customer segmentation, can be handle by non-technical pe In addition, Neural networks also gives best result but we left some limitations in it. For Future Researchers can explore other model of neural networks i.e. CNN in comparison with Deep Learning algorithm i.e RNN

References

[1] W. R. Smith, "Product differentiation and market segmentation as alternative marketing strategies," Journal of marketing," Journal of Marketing, vol. Vol. 21, Jul., 1956.

[2] A. S. M. S. Hossain, "Customer segmentation using centroid based and density based clustering algorithms," 3rd International Conference on Electrical Information and Communication Technology (EICT), pp. pp. 1-6, 2017.

[3] Wikipedia, "Artificial neural network," n.d.. [Online].

[4] S. I. L. Ozan, "Artificial Neural Networks in Customer Segmentation," 27th Signal Processing and Communications Applications Conference (SIU), pp. 1-4, 2019.

[5] Tutorialspoint, "what_is_weka," n.d.. [Online]. Available: https://www.tutorialspoint.com/weka/what_is_weka.htm.

[6] J. E.-E. I. N. S. Nazzal, " Multilayer Perceptron Neural Network (MLPs) For Analyzing the Properties of Jordan Oil Shale," World Applied Sciences Journal, 2008.

[7] Kaggel, "mall customer dataset," Keggal, [Online]. Available: https://www.kaggle.com/shwetabh123/mall-customers.

[8] D. P. K. P. Anuj Sharma, "A Neural Network based Approach for Predicting churn of customer churn," International Journal of Computer Applications, pp. 10-20, A Neural Network based Approach for Predicting.

[9] M. A.-A. M. T. Fahmida Afrin, "Comparative Performance Of Using PCA With KMeans And Fuzzy C Means Clustering For Customer Segmentation," INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH , vol. Volume 4, no. Issue 10, OCTOBER 2015.

COMPARISON OF DIFFERENT CLASSIFICATION ALGORITHMS FOR A DIABETES DATASET USING WEKA

Fakhra Akhtar. Department of Computer Science, Riphah University, Lahore, Pakistan.

Abstract. Data mining relies upon compiling information assortment, warehousing and computer processing. This paper analyses the various algorithms used in machine learning system. It analyses and interprets the results based on the classification of a single data set using various classification techniques. The dataset used in this paper relates to Diabetes. The algorithms used in this paper are logical regression, Random Forest Algorithm and JRip. The data set is taken from UCI ML repository and it relates to the population of effected diabetic patients in India in which mostly population are young females in their early twenties. The dataset contains in total 768 instances and 9 attributes which specify the descriptive properties of diabetic patients. The objective of this paper was to identify diverse classification techniques on a single data set using WEKA tool. it is observed that highest accuracy is 81.81 % and the lowest one is 79.22% The JRip yields highest accuracy followed by logistic regression classification.

KEYWORDS: DATA MINING, WEKA, ALGORITHMS

1. INTRODUCTION

Today in the era of digitalization, data relating to health requires analytical methodology in identifying vital information that are used for the purpose of decision making. It helps in detection of various diseases relating to respiration, cardiac, neuro and diabetes, and assist the paramedic staff to make effective healthcare policies in order to maintain health profiles for patients. There are difficulties and challenges in evaluating the large quantity of data. Healthcare sector data requires accuracy in diagnosis, management and treatment of diseases. Data mining applications in health and medic field could have a great potential in the medical industry. This research basically deals with the analysis and application of various classification algorithms on a single data set.

The first phase of the paper explains about literature review regarding to the defined dataset i.e. diabetic patients, it specifies all the previous research work on diabetes and its impact on the health of various patients. The second phase deals with the methodology data analysis tool and classification of algorithms. The third part discusses the research results based on the classification algorithms used in the Weka Tool. Fourth phase deals with Discussion, conclusion and references.

PROBLEM STATEMENT

The problem of this research paper describes the fact that Nowadays it is necessary to diagnose diabetes as it can occur at any stage, once diagnosed its treatment undergoes various procedures and diagnoses.

RESEARCH QUESTION

How to analyze the numeric diabetes dataset?

How to analyze and interpret diabetes data for treatment.

LITERATURE REVIEW

(1) Care, D. (2006) Defines Diabetes as a group of metabolic sicknesses described by hyperglycemia coming about because of imperfections in insulin emission, insulin activity, or both. The interminable hyperglycemia of diabetes is related with long term harm, and damage of various organs. (2) American Diabetes Association. (2017) Describe the fact that Diabetes can be classified into the following general categories: Type one diabetes because of immune system b-cell devastation, typically prompting supreme insulin inadequacy Type two diabetes because of a dynamic loss of b-cell insulin discharge much of the time on the foundation of insulin opposition Gestational diabetes mellitus (GDM) diabetes analyzed in the second or third trimester of pregnancy before incubation both are heterogeneous (3) Scheen, A. J. (2003). In their research describe Type two diabetes mellitus as a heterogeneous syndrome characterized by abnormalities in carbohydrate and fat metabolism. They described the fact that the causes of type two diabetes are multifactorial and include both genetic and environmental elements that affect beta-cell function and tissue. (4) Coppola, A. et al (2016). Describe the fact that the management of type two diabetes mellitus includes ability and empowerment of the patient to change lifestyle, it was inferred that the exploration has not explained the perfect qualities of a far reaching understanding instruction program in clinical practice. Their work is expected to address open inquiries with respect to the sort of topics, among instructor and patient, foundation of teachers, utilization of new innovations. (5) Braves, et al (2010). The aim of this study was to determine the impact of Ramadan-focused education on weight and hypoglycemic episodes during Ramadan in a Type 2 diabetic Muslim population taking oral glucose-lowering agents. They retrospectively analyzed two groups. Results concluded the fact that Ramadan-focused education in diabetes can empower patients to change their lifestyle during Ramadan.

METHODOLOGY

Artificial intelligence is considered as a use of computerized reasoning (AI) that gives frameworks the capacity to naturally take in and improve as a matter of fact without being customized. It centers around the advancement of PC programs that can get to information and use it for themselves in this research logistic regression, random forest and JRip are used.

Logistic Regression.

Logistic regression was developed by statisticians to describe properties of population growth in particular sector (6) it is used primarily with dichotomous dependent variables It gives a diagram of the calculated logistic regression model and talk about the fundamental similitudes and contrasts between strategic and direct relapse and use information from an expected examination to tell the best way to decipher a calculated relapse investigation.

Random Forest

The Random forest are considered as a learning strategy for characterization, relapse and different assignments that work by building a mass of choice trees at

preparing time. The Random Forest Classifier is characterized as set of choice trees from haphazardly chosen subset of preparing set (7) It alternatively creates two extra snippets of data: a proportion of the significance of the indicator factors, and a proportion of the inner structure of the information.

JRIP

JRIP is an algorithm used as a representative technique in attempts to describe the problems of decision tree learning. It was introduced by Quinlan (1987) in the context of decision tree learning. It is considered and appears to be a very simple, almost trivial, algorithm for pruning, there are many different algorithms that go under the same name. No consensus exists whether rep is a bottom-up algorithm or an iterative method.

The Data

Data used in this research paper relates to diabetes contains in total 768 instances and 9 attributes which specify the descriptive properties of diabetic patients. This data will be helpful for the researcher to make a decision whether a patient is effected by diabetes or not. In order to analysis and investigate the performance on the selected classification algorithms namely logistic regression, JRIP and random forest. we used the WEKA tool to produce the result. the 80% of data used for training and 20% of data used testing.

WEKA

For the purpose of analyzing the data WEKA is used basically The Waikato Environment for Knowledge Analysis (WEKA) occurred through the apparent requirement for a unified work bench that would permit specialists simple access to excel in class methods in AI. It was imagined that WEKA would give a tool compartment of learning calculations, yet in addition a system inside which specialists could execute new calculations without being worried about supporting framework for information control and plan assessment.

RESULTS

To analysis and investigate the performance on the selected classification algorithms namely logistic regression , and random forest. we used the WEKA tool to produce the result. the 80% of data used for training and 20% of data used testing. In WEKA, all data is considered as examples and structures are defined as attributes. The results are shown into several sub items for easier interpretation in the first phase, correct and incorrect classified instances will be segregated in numeric and percentage value. Kappa statistic, mean absolute error and root mean squared error will be in numeric only. The results of the model are shown in Tables 1 and 2 below. Table 1 mainly summarizes the result based on accuracy and time taken for the interpretation of the model. Meanwhile, Table 2 shows the result based on error during execution of the model whereas Figures 1 and 2 are the visualizations of the dataset model results.

1.1 Table 1 Model result of each algorithm

Algorithm (Total Instances, 154)	Correctly Classified Instances % (value)	Incorrectly Classified Instances % (Value)	Time Taken (seconds)	Kappa Statistic
Logistic	81.1688	18.8312	0.02	0.5384

	(125)	(29)		
JRip	81.8182 (126)	18.1818 (28)	0	0.541
Random forest	79.2208 (122)	20.7792 (32)	0.02	0.4993

In Table 1, it is observed that highest accuracy is 81.81 % and lowest one is 79.22 %. The JRip produces highest accuracy followed by logistic regression classification. An average of 154 illustrations out of total 768 illustrations are found to be correctly classified.

Figure 1 graphical visualization of diabetes model

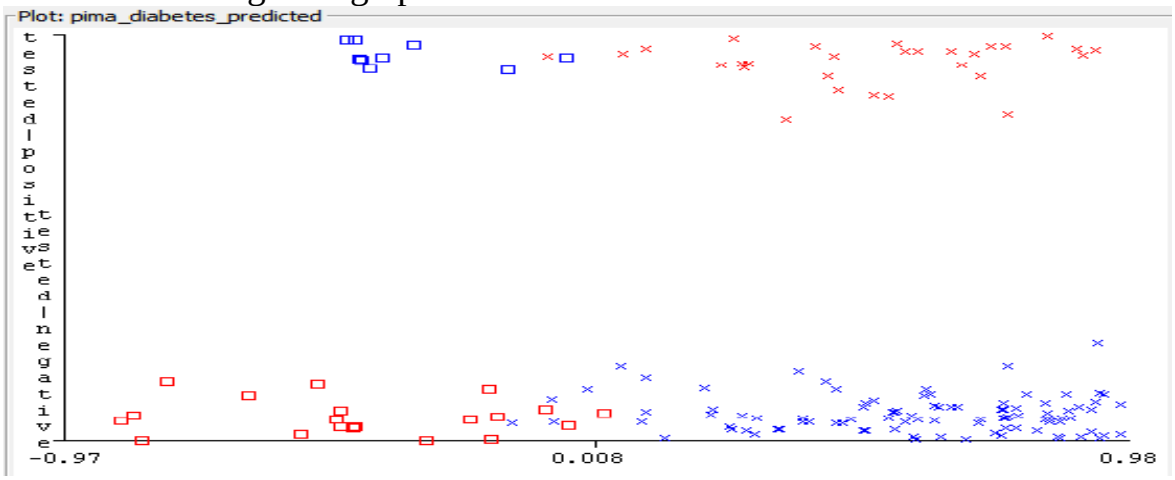


Figure 1 shows the graphical representation of the diabetes data as per algorithms

a. Table 2 Training and simulation errors

Algorithm (Total Instances, 154)	Mean Absolute Error	Root Mean Squared Error	Relative Absolute Error (%)	Root Relative Squared Error (%)
logistic	0.2942	0.3768	65.6566	80.6233
JRip	0.3251	0.3886	72.5638	83.1398
Random forest	0.2991	0.3869	66.7536	82.7709

In table 2 the differences of error are shown It is discovered that the highest error is found in JRip which is 0.388 whereas logistic regression has a lowest error rate of 0.376. Therefore, an algorithm which has a lower error rate will be preferred.

Discussion

Based on results shown in Table 1, it is observed that the highest accuracy is 81.81 % and the lowest one is 79.22 %. The JRip yields highest accuracy followed by logistic regression classification. An average of 154 instances out of total 768 instances are found to be correctly classified. The total time required to build the model is also a crucial parameter in comparing the classification algorithm. Both random forest and logistic regression took 0.02 s while as compared to JRip it took almost 0s shortest time as compared to others. The normal Kappa score from the chose calculation is around 0.541 Based on the Kappa Statistic models, the precision of this arrangement reasons for existing is considerable. In table 2 the distinctions of errors are appeared. It is discovered that the highest error is found in

JRip which is 0.388 whereas logistic regression has a lowest error rate of 0.376. Therefore, an algorithm which has a lower error rate will be preferred based on more powerful classification capability in terms of medical and healthcare fields.

CONCLUSION AND RECOMMENDATION

The results show the fact that highest accuracy is 81.81 % and the lowest one is 79.22% The JRip yields highest accuracy followed by logistic regression classification. It is very important to have an Awareness and proper knowledge of various aspects of Diabetes the recommendation of this paper describes the fact that precautionary measures should be adopted for the prevention and control of the disease.

REFERENCES

1) Care, D. (2006). Diagnosis and classification of diabetes mellitus. Diabetes care.

2) American Diabetes Association. (2017). Classification and diagnosis of diabetes.

3) Scheen, A. J. (2003). Pathophysiology of type 2 diabetes. Acta Clinica Belgica, 58(6),

4) Coppola, A., at el (2016). The role of patient education in the prevention and management of type 2 diabetes: an overview. Endocrine,

5) Bravis, V., at el (2010). Ramadan Education and Awareness in Diabetes (READ) programme for Muslims with Type 2 diabetes who fast during Ramadan.

6) Wright, R. E. (1995). Logistic regression.

7) Liaw, A., & Wiener, M. (2002). Classification and regression by random Forest. R news,

8) Mohamed, W. N. H. W., Salleh, M. N. M., & Omar, A. H. (2012, November). A comparative study of reduced error pruning method in decision tree algorithms.

9) Elomaa, T., & Kaariainen, M. (2001). An analysis of reduced error pruning. Journal of Artificial Intelligence Research,

10) Markov, Z., & Russell, I. (2006). An introduction to the WEKA data mining system.

ТАБИИЙ ФАНЛАРНИ ЎРГАНИШДА ИНФОРМАТИКАНИНГ РОЛИ

Комилжонов Н.А., Медатов А.А. Андижон давлат университети, Ўзбекистон.

Аннотация. *Маъруза тезисларида замонавий ахборот технологияларини табиий фанлар муаммоларни ўрганишда тадбиқ қилиниши кўриб чиқилган.*

Калит сўзлар: *Оқсил, ДНК, биоинформатика, ген.*

Аннотация*. В тезисах доклада рассматрены применение современных информационных технологий при изучении проблем естествознания.*

Ключевые слова*: Углевод, ДНК, биоинформатика, ген*

Annotation*. The thesis discusses the use of modern information technologies in the study of natural science problems.*

Keywords: *Carbohydrate, DNA, bioinformatics, gene*

Ҳозирги вақтда мавжуд бўлган барча фанларни информатикасиз тасаввур қилиб бўлмайди. Ўтган асрнинг 60-70 йилларидан бошлаб ЭҲМ лари биологияда ишлатила бошлади. Вақт ўтиши билан замонавий компьютерлар пайдо бўлди ва параллел равишда биологияда кўплаб олинган экспериментал маълумотлар тўпланди.

Жорий аср бошларида биомедицина, биоинформатика фанлари ривожланди. Биологик тажрибалардан олинаётган маълумотлар сиғими шу даражада кўпайиб кетдики, уларни таҳлил қилиш ва ёдда ушлаб туриш кишилар учун мушкул ишга айланди. Маълумотлар хажмининг бу даражада ўсиши маълумотларни сақлаш заруриятини қелтириб чиқарди.

Оқсилларнинг дастлабки аниқланган таркиби атлас китоб кўринишида чоп қилинди. Кейинги амалга оширилган ишларда кўплаб оқсилларнинг таркиби ўрганилаб уларни китоб кўринишида сақлашни имкони бўлмай қолди. Уларни ўрганиш инсон ақли чегарасидан чиқди. Бу ишларни амалага ошириш учун махсус дастурлар керак бўлди. Ўтган асрнинг сўнги ўн йиллигида турли тирик организмлар одамдан тортиб турли микробларгача бўлган ораликда тирик организм генларининг тузилиши ўрганилди. Инсон генини ўрганиш натижасида олинган натижа шуни кўрсатадики, олинган натижа 15000 та томдан иборат китобга сиғиши аниқлади. Бунча мураккаб маълумотлар билан ишлаш учун инсонларга компьютер технологияларига суянмасдан иложи йуқ бўлиб қолди.

Илм фан ривожланиши ва шу билан бирга ахборот технологияларининг имкониятларини ортиши биология фанининг муаммоларини ўрганишда кенг имкониятлар яратди. “Биоинформатика” фанининг вазифаларига турли оқсил ва нуклейин кислоталарни таҳлил қилиш учун ҳисоблаш алгоритмларни яратиш, маълумотларни системалаштириш кабилар киради. Биоинформатика

фани информатикадан ташқари амалий математика, биохимия, биофизика, статистика ва бошқа аниқ фанларга ҳам суяниб иш кўради.

Биоинформаткани ҳозирги вақтга келиб молекуляр биологиядан деярли алоҳида ажралиб чиққан йўналиш сифатида қараш мумкин. Биоинформатика фани ҳозирги кунда биология фанининг геномика, биологик ривожланишнинг компьютер моделлаштирилиши, генлари компьютерлар ёрдамида таҳлил қилиш ва генлар популяциясини моделлаштириш каби йўналишларида ўз ўрнини топган. Агар биоинформатика фанини асосий йўналишларига алоҳида тўхталадиган бўлсак уларни қуйидаги асосий бўлимлар ташкил қилади.

Генлар тузилишидаги кетма-кетликни ўрганиш. Турли организмлар ДНК стуктураларини ўрганиш ва улар ҳақидаги маълумотларни таҳлил қилишни оддий ҳолда имкони йўқ. Бунинг учун албатта махсус дастурлардан фойдаланиш керак бўади. Турли генларни ўрганишда уларни фрагментларга бўлиниб ўрганилади ва уларни йиғиб бир бутун ҳолга келтириш мураккаб ҳисобланади. Уларнинг тузилиш моделини яратиш компьютер дастурлари орқали амалга оширилади. Бу нарса биоинформатика фанини олдида турган муҳим муамолардан бири ҳисобланади. Турларни эволюция жараёнини ўрганувчи биология турли организмларни эволюция жарёнида генларни ўзгаришини ўрганиш билан бир қаторда уларни ўзаро солиштириш ишларини комплекс ўрганишни амалга оширади. Уларни ровожланиш моделларини яратади.

Биологик хилма-хилликни баҳолаш. Экосистеманининг биологик хилма хиллиги ўша белгиланган муҳитнинг генетик умумийлигини беради. Бу битта заррача ёки бутун бошли биологик объект бўлиши мумкин. Махсус дастурлар ёрдамида муҳитни визуаллаштириш, таҳлил қилиш ва бошқа одамларга етказиш мумкин.

Эксперимантал йўл билан олинган маълумотлар биоинформатика усулларидан фойдаланиб ўрганилади. Бунда олинган маълумотлар аввал олинган ва маълумотлар базасига киритилган қийматлар билан солиштирлади, хулоса қилинади. Агар олинган қийматлар маълумотлар банкидан топилмаса маълумот маълумотлар банкига киритилади. Мавжуд маълумотлар банкининг вазифаларига маълумотларни сақлаш, системалаштириш, уларни янгилаш ва улардан фойдаланиш имкониятларини бериш киради. Юқорида кўрилган вазифалар уз навбатида имконияти жихатдан ўта юқори бўлган компьютерларни талаб қилади.

Биоинформатиканинг яна бир асосий имкониятларидан бири турли вирус ва микробларга қарши дори ишлаб чиқаришда фойдаланилади. Ўрганилган микроб ёки вируснинг тузилиши ундаги мавжуд карбоксил ва аминогруппаларнинг қандай кетма кетликда жойланишини моделлаштириш унга қарши қандай структурага эга бўлган дори уни зарарсизлантиришини ўрганиш имкониятини беради. Мисол тариқасида Auto Dok дастурини кўрадиган бўлсак унинг ёрдамида маълум структурага эга бўлган дори аниқланган 3D структурали оқсил молекуласи билан қандай

таъсирлашишини кўриш имкониятини беради. Булардан ташқари нуклеотид ва аминокислотлар кетма кетлигини солиштириш учун ишлаб чиқилган ACT – (Artemis Comparison Tool), Arlequin, Bio Edit каби дастурларни мисол қилишимиз мумкин.

Хулоса қилиб айтганда ҳозирги замон информатика фани имкониятларини турли фанлар билан биргаликда ишлатилиши қўлланилаётган фаннинг имкониятларини ортишига, улардаги муаммоларни тез ва аниқлик даражаси юқори бўлган ҳолда хал бўлишига олиб келади. Бу эса ўз навбатида информатика фанининг ҳам қирраларини ортишига ва уни янада фан сифатида янада юксак ривожланишига сабаб бўлади.

Адабиётлар

1. *Каменская М.А. Информационная биология / М. А. Каменская. – М.: Академия, 2006. – 368 с.*

2. *Огурцов А.Н. Методы биоинформационного анализа / А.Н. Огурцов. – Х.: НТУ "ХПИ", 2011. – 114 с.*

3. *Основы биоинформатики : учеб. пособие / А. Н. Огурцов. – Х.: НТУ «ХПИ», 2013. – 400 с*

C++ BUILDER DASTURLASH TILIDA ILOVALAR UCHUN ANIMATSIYALAR YARATISH

Matkarimov J.A, Mirzaaxmedov M.K, Madolimov F.E.
Andijon davlat universiteti, O'zbekiston

Annotatsiya: *Ushbu ishda dasturlash mutaxassisliklariga C++ Builder dasturlash tilida ilovalar uchun turli ko'rinishdagi animatsiyalar yaratish jarayoni ko'rsatib o'tilgan.Dasturlarga qo'shimcha tarzda qo'shish uchun harakatlantirish kodlari taqdim etilgan.*

Kalit so'zlar: *C++, Builder, animatsiya, window,inrefeys, timer,OnCreate, OnClick.*

Annotation: *This work includes the process of creating animations for various applications in the C++ program language for programming specialties is shown.Navigation codes are provided to add to the progrms.*

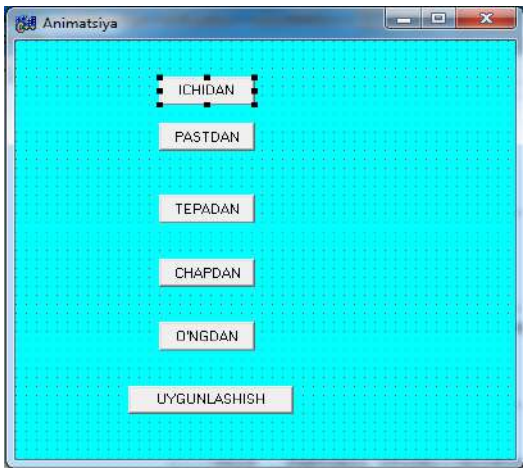
Keywords: *C++, Builder, animation, window,interface, timer,OnCreate, OnClick.*

Аннотация: *В этой статье описывается процесс создания анимации для различных приложений на языке программирования C ++ Builder для специальностей программирования.Добавлены навигационные коды для добавления в программы.*

Ключевые слова: *C++, Builder, animatsiya, window,inrefeys, timer,OnCreate, OnClick.*

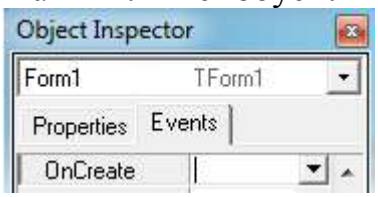
Zamonaviy texnologiyalar, avtomatlashtirish jarayonlari kun sayin jadal suratlarda rivojlanmoqda.Texnik ta'minot bilan birgalikda dasturiy ta'minotga bo'lgan talab ortib bormoqda.Samaradorligi yuqori bo'lgan turli yo'nalishlarga mo'ljallangan ilovalar yaratilib amaliyotga tadbiq eltilyapti.Foydalanuvchilarni ko'pchiligi birinchi navbatda ilova dizayniga(tashqi interfeys) va multimedia imkoniyatiga e'tibor qaratadi.Yuqori imkoniyatga ega dasturlash tillari Pyton,C++,Java,Java script,PHPlar bilan bir qatorda C++ Builder dasturlash tili xam multimediayali ilovalar yaratishga mo'ljallangan.Ma'lumki multimediaya grafika,ovoz,video,animatsiyalardan iborat bo'ladi.Biz ushbu maqolamizda C++ Builder 6 dasturlash tili orqali ilovani asosiy bosh oynalariga chiqish animatsiyalarini hosil qilishni va obyektlarni turli yo'nalishlar bo'yicha harakatlantirishni ko'rsatib o'tganmiz.

Dastlab ilovaning asosiy oynasini shakllantirib olamiz.Formaning o'lchamini Object Inspector oynasidan Height,Width xususiyati,rangini Color xususiyati orqali ixtiyori ko'rinishga keltiramiz. Animatsiyani turli ko'rinishlarda taqdim etish uchun Additional komponentalar palitrasidan BitBtn obyektidan foydalanamiz.



1-rasm.

Animatsiya dastur ishga tushganda avtomatik tarzda ishga tushishini istazangiz Formaning On Create hodisasiga (2-rasm) yoki System komponentalari palitrasidan Timer obyektining On Timer hodisasiga kodlarni yozishingiz mumkin.Timer obyektining Enabled xususiyati true qiymatda bo'lishi kerak.



2-rasm.

```
//-------------------------
#pragma package(smart_init)
#pragma resource "*.dfm"
TForm1 *Form1;
const int tez=3000;
```

3-rasm

1.Asosiy oynaga animatsiya berish turlari.

Global tarzda tez nomli butun toifali o'zgarmasni e'lon qilib olamiz va uning qiymatini 3000 qilib belgilaymiz.Bu yerda bu o'zgarmas qiymatimiz oynaning harakatli hosil bo'lish tezligini bildiradi.(3-rasm)

BitBtn komponentasining xar birining OnClick hodisasiga quyidagi kodlarni yozib chiqamiz.

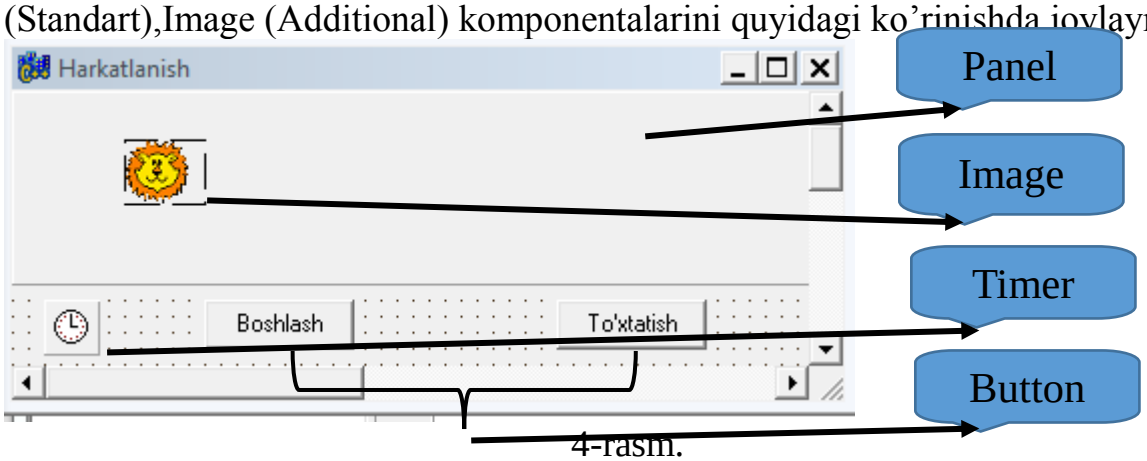
```
void __fastcall TForm1::BitBtn1Click(TObject *Sender)
{
Hide();  //ICHKARIDAN CHIQISH
AnimateWindow(Handle, tez, AW_CENTER );
Show();
}
//---------------------------------------------------------------------------
void __fastcall TForm1::BitBtn2Click(TObject *Sender)
{
Hide();    //PASTDAN CHIQISH
 AnimateWindow(Handle,tez,AW_VER_NEGATIVE);
 Show();
}
//---------------------------------------------------------------------------
void __fastcall TForm1::BitBtn3Click(TObject *Sender)
{
```

```
Hide(); //TEPADANDAN CHIQISH
 AnimateWindow(Handle,tez,AW_VER_POSITIVE);
 Show();
}
//---------------------------------------------------------------------------
void __fastcall TForm1::BitBtn4Click(TObject *Sender)
{
Hide(); //CHAPDAN CHIQISH
AnimateWindow(Handle,tez,AW_HOR_POSITIVE);
Show();
}
//---------------------------------------------------------------------------
void __fastcall TForm1::BitBtn5Click(TObject *Sender)
{
Hide();  //O'NGDAN CHIQISH
AnimateWindow(Handle,tez,AW_HOR_NEGATIVE);
Show();
}
//---------------------------------------------------------------------------
void __fastcall TForm1::BitBtn6Click(TObject *Sender)
{
 Hide();//ORQA FON BILAN UYG'UNLASHIB CHIQISH
AnimateWindow(Handle,tez, AW_BLEND);
Show();
}
```

Bu yerda:

Hide()- funksiyasi oynani yashirib turadi, oynaga berilgan animatsiya funksiyasi faollashgandan so’ng Show() funksiyasi oynani ko’rsatishni boshlaydi.

2.Timer komponentasi orqali obyektlarni harakatlantirish.Obyekt xususiyatiga qarab o’ngdan,chapdan, tepadan,pastdan animatsiyalar berish imkoni mavjud. Buning uchun ilovamizning asosiy form oynasiga Timer(System),2 ta Button,Panel (Standart),Image (Additional) komponentalarini quyidagi ko’rinishda joylaymiz.



4-rasm.

2.1.Panel komponentasining formaga joylaymiz va Object inspector oynasidan Caption xususiyatini o’chiramiz.

2.2.Timer komponentasini Enabled xususiyatini False qiymatga o'tkazamiz, interval xususiyatini 50 qiymatga o'gartiramiz va timer hodisasiga ushbu kodni yozamiz.

```
void __fastcall TForm1::Timer1Timer(TObject *Sender)//Timer komponentasiga
{
Image1->Visible=true;//image komponentasini ko'rinish xususiyatiga o'tkazadi
Image1->Left=Image1->Left+3;//image komponentasini harakatlanish qadamini 3 birlikga belgilaydi
if (Image1->Left>Panel1->Width) Image1->Left=0-Image1->Width;//harakatni panel komponentasini eni bo'ylab va qaytib chiqishini belgilaydi

}
//---------------------------------------------------------------------------
```

2.3.Button1 componentamizni caption xususiyatini Boshlash deb nomlaymiz,OnClick hodisasiga quyidagi kodni yozamiz.

```
void __fastcall TForm1::Button1Click(TObject *Sender)//Button1 komponentasiga
{
Timer1->Enabled=true;//harakatni boshlaydi
}
//---------------------------------------------------------------------------
```

2.4.Button2 componentamizni caption xususiyatini To'xtatish deb nomlaymiz,OnClick hodisasiga quyidagi kodni yozamiz.

```
void __fastcall TForm1::Button2Click(TObject *Sender) //Button2 komponentasiga
{
Timer1->Enabled=false;//harakatni to'xtatadi
}
```

Izoh:Image komponentasinig Picture xususiyati orqali ixtiyoriy rasmni istagan o'lchamda forma va panelga moslab olishingiz mumkin.Harakatni tepadan pastga yoki aksini qilmoqchi bo'lsangiz Left funksiyasini Top funsiyasiga ishorasini qarama-qarshisiga o'zgartirasiz.

Adabiyotlar

1. *Sh.A. Nazirov,R.V. Kabulov,M.R. Bahajanov, Q.S. Raxmanov .C va C++ tili. O'quv qollanma//Toshkent.2013.408-420 b.*
2. *Шмидский Я.К. Программирования на языке С++; Самоучитель. Учебное пособие, М.: Диалектика, 2004, 361 b.*

SQL ТИЛИДА ЖАДВАЛ ЯРАТИШ

С.Ҳайдарова Қўқон давлат педагогика институти, Ўзбекистон.

Аннотация: *Ушбу мақолада SQL тилида* ***CREAT TABLE*** *инструкцияси синтаксиси ва ундан фойдаланиб жадвал яратиш усуллари келтирилган.*

Аннотация: *В данной статье приводятся синтаксис инструкции* ***CREAT TABLE*** *и методы создания таблиц на языке SQL с помощью данной инструкции.*

Annotation: *This article describes the syntax of the* ***CREAT TABLE*** *instruction and methods for creating tables in SQL language using this instruction.*

Таянч сўз ва иборалар: *CREAT TABLE оператори, бирламчи калит, бўш ва бўш эмас устунлар.*

Ключевые слова и выражения: *оператор CREAT TABLE, первичный ключ, пустой (не пустой) столбец.*

Keywords and expressions: *CREAT TABLE operator, primary key, empty (not empty) column*

SQL тилида жадвал яратиш учун CREAT TABLE инструкциясидан фойдаланилади. Бу буйруқ қаторларсиз бўш жадвал яратади.

CREATE TABLE буйруғи жадвал номини ва жадвал ўзини маълум тартибда кўрсатилган устунлар номлари кетма – кетлиги таърифи кўринишида аниқлайди. У маълумотлар типлари ва устунлар ўлчовини аниқлайди. Ҳар бир жадвал жуда бўлмаганда битта устунга эга бўлиши керак.

CREATE TABLE буйруғи синтаксиси:

CREATE TABLE <table-name >
(<column name><data type>[(<size>)],
<column name><data type>[(<size>)], ...);

Аргумент қиймати катталиги маълумот турига боғликдир. Агар бу қиймат кўрсатилмаса тизим автоматик қийматни ўрнатади.

Қуйида кўп ишлатиладиган жадваллар билан танишиб чиқайлик. Бу жадваллар Products, Vendors, Customers, Orders ва OrderItems жадваллари бўлиб, улар орасидаги алоқалар 1-расмда тасвирланган. Ҳар бир жадвал учун бирламчи калитлар аниқланган. Мисол учун, Products деб номланган маҳсулотлар жадвалини кўрайлик, бу жадвал маҳсулотлар каталогини ўз ичига олади. Ҳар бир маҳсулот ўзининг уникал, яъни такрорланмайдиган идентификатори (prod_id устуни)га эга ва у vend_id устуни (таъминловчининг уникал идентификатори) билан боғланган. Бу жадвалда prod_id бирламчи калит бўлиб ҳисобланади.

Энди 1-расмда келтирилган Products ва Orders жадвалларни яратишни кўриб чиқайлик.

Аввало Products жадвалини яратишни кўриб чиқамиз:

```
CREATE TABLE Products
    (
  prod_id    char (10)      NOT NULL ,
  vend_id    char (10)      NOT NULL ,
  prod_name  char (255)     NOT NULL ,
  prod_price decimal (8,2)  NOT NULL ,
  prod_desc  text           NULL
) ;
```

Бу ерда

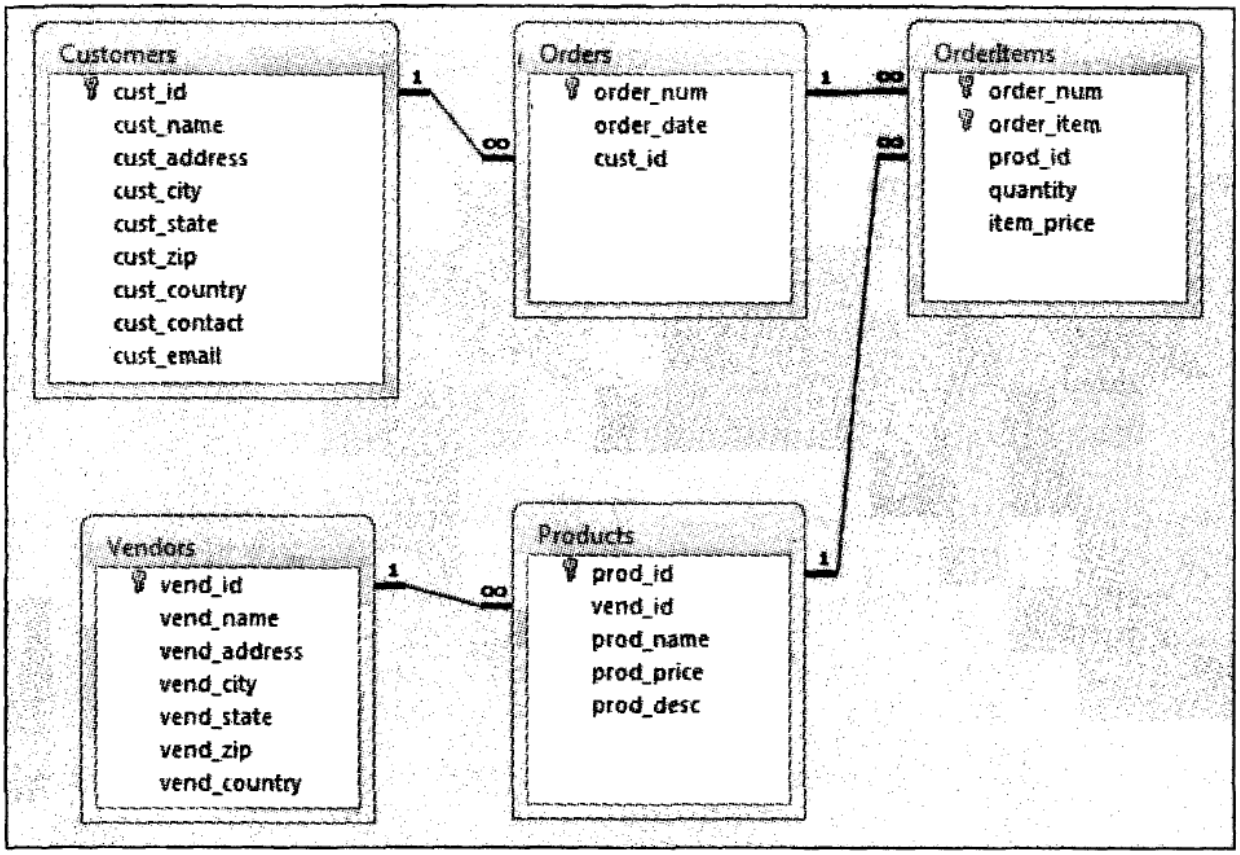
prod_id - маҳсулотнинг уникал идентификатори;

vend_id - таъминловчининг уникал идентификатори;

prod_name – маҳсулотнинг номи;

prod_price – маҳсулотнинг нархи;

prod_desc - маҳсулотнинг тавсифи.



1-расм. Жадваллар орасидаги алоқалар.

Кўриниб турибдики, бунда CREATE TABLE калит сўзидан кейин жадвал номи ёзилади. Жадвални аниқлашда унинг барча устунлари юмалоқ қавсга олинади. Устунлар бир-биридан вергул билан ажратилади. Келтирилган мисолдаги жадвал бешта устундан иборат бўлиб, ҳар бир устун унинг номини бериш билан аниқланади.Ундан сўнг маълумотлар тури аниқланади.

NULL қиймати устунда ҳеч қандай қиймат мавжуд эмаслигини, яъни устун бўш эканлигини, NOT NULL эса бўш эмаслигини билдиради, хуллас, инструкция нуқта вергул билан тугалланади.

Кейинги жадвал бу буюртмалар жадвали бўлиб, у Orders деб номланади. Бу жадвалда мижозларнинг буюртмалари ҳақидаги маълумотлар сақланади.

Жадвал учта устундан иборат: буюртманинг уникал номери, муддати ва мижоз идентификатори.

Ҳар бир буюртма уникал номер(order_num устуни)га эга. Буюртмалар мос мижозлар билан cust_id устуни орқали боғланган. Бу жадвалда order_num устуни бирламчи калит, cust_id устуни эса ташқи калит бўлиб ҳисобланади.

Orders жадвалини яратишни кўрайлик:

```
CREATE TABLE  Orders
(
   order_num  int          NOT NULL ,
   order_date  datetime   NOT NULL ,
   cust_id    ch  ar(10)     NOT NULL
);
```

Бу ерда

order_num - буюртманинг уникал номери;

order_date - буюртманинг муддати;

cust_id - буюртмани берган мижоз идентификатори.

Адабиётлар

1. *Бен Форта.Освой самостоятельно SQL за 10 минут, 4-е изд.: Пер. с англ.—М .: ООО “И.Д. Вильямс”, 2014. — 288 с .*
2. *Базы данных. Курс лекций по дисциплине: Методические указания по дисциплине базы данных для бакалаврской подготовки студентов по направлению 230100 «Информатика и вычислительная техника»/ А.А. Иптышев, Д.И. Морозов, А.А Городилов, М.А. Иптышев, Красноярск.: СФУ, 2007, 146с.*

АЛГОРИТМ ПОИСКА ПУТЕЙ ДЛЯ АВТОМАТИЗАЦИИ ПРОМЫШЛЕННЫХ РОБОТОВ

Рузиматов С.Ш. - Ташкентский университет информационных технологий имени Мухаммада Ал-Хоразмий, Узбекистан.

***Анотация**: Основным направлением данной статьи является упрощение каналов разработки моделей автоматизации при производстве алгоритмов*

***Ключевые слова**: СП(сигнального пути) , ПР(промышленных роботов), САПР (система автоматизирование промышленных роботов), автоматизация, робот, шаг, алгоритм*

***Annotation**:The main direction of this article is to simplify the channels of development of models of automation in the production of algorithms.*

***Keywords**:: SP (signaling path), PR (industrial robots), ASIR (automation system for industrial robots), automation, robot, step, algorithm*

Нынешний уровень развития информационных технологий требует разработки новых выскоэффективных методов и средств автоматизации программирования роботов. В связи с этим тема данной работы, посвященной автоматизации программирования промышленных роботов (ПР), является весьма актуальной.

В работе используются графовая модель процесса программирования, элементы теории графов и алгоритмов, а также алгоритмический язык С++. Здесь узлам графа приводятся в соответствие программные модули для реализации определенных элементарных операций ПР или процедур, а дугам графа – последующий программный модуль после исходгого. Составленная таким образом графовая модель служит информационной средой для автоматизации процесса составления программ ПР. Составление программы сводится к поиску ориентированного пути в графовой модели. Для поиска ориентированного пути разработан алгоритм поиска специального сигнального пути (СП) в графовой модели. Под термином “специальный сигнальный путь” подразумываются различные маршруты, содержащие или не содержащие отдельные узлы графовой модели.

Разработанный алгоритм поиска путей состоит из 11 шагов. Ниже приводим последовательность шагов данного алгоритма.

1. Установка начальных значений переменных и массивов.
2. Установка начала СП. Для этого в качестве первого элемента СП берется начальный узел.
3. Выбор очередной точки, находящейся от последней точки СП на расстоянии элементарного пути. Если такой точки нет, то осуществляется переход к 8 – шагу.
4. Проверка выбранной точки и:

1. переход к 10 – шагу, если через данную точку ранее прошел СП;
2. переход к 5 – шагу, если данная точка признана конечной точкой СП;
3. переход к 7 – шагу, если через данную точку не прошел СП;

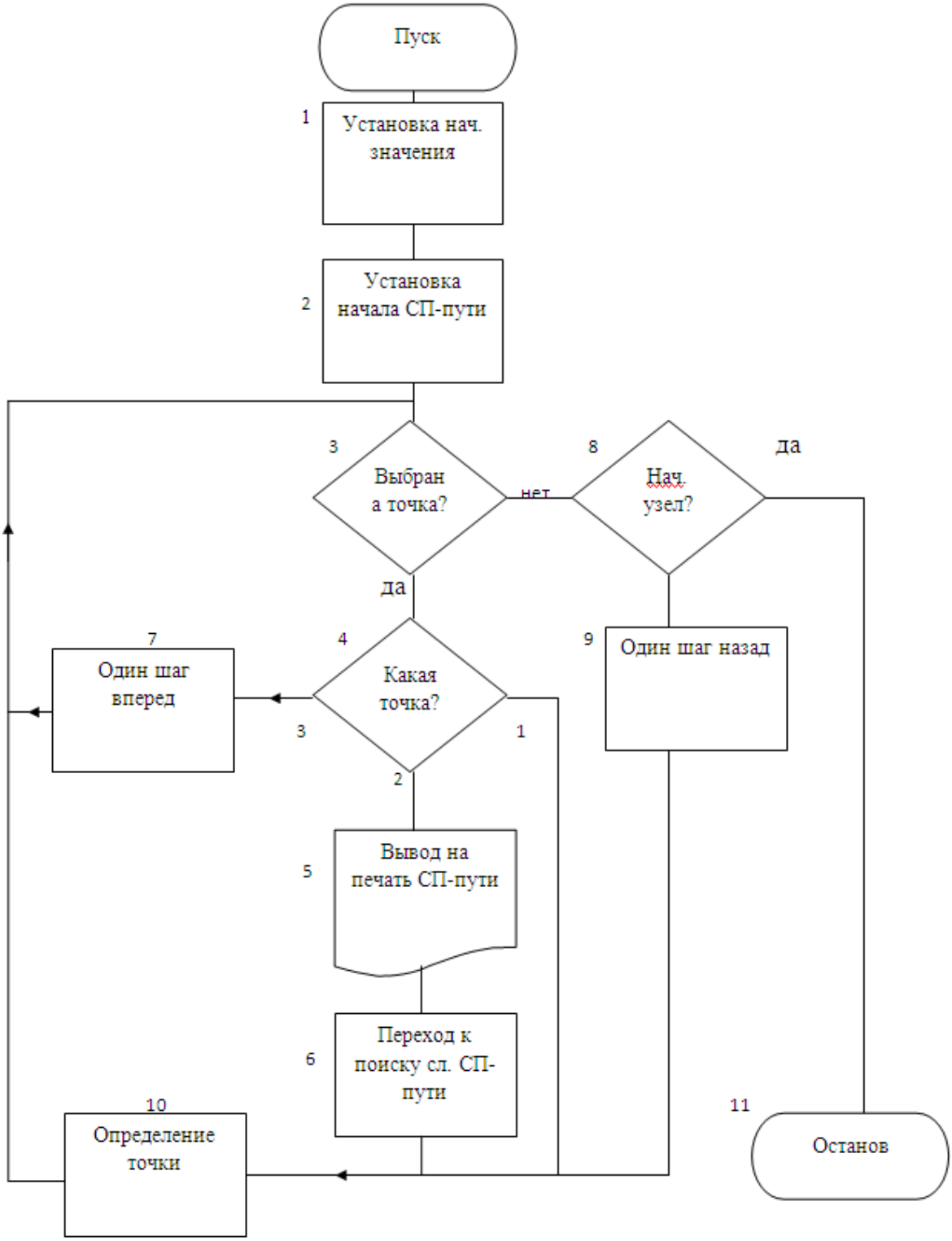


Рис. Структурная схема алгоритма поиска сигнальных путей между прогрмммными модулями в графовой модели

5. Вывод на печать (или запись в специально отведенный массив) сформированного СП.

6. Переход к 10 – шагу для поиска следующего СП.

7. Один шаг вперед. Для этого выбранная точка включается в СП и осуществляется переход к 3 – шагу.

8. Проверка ситуации о том, что при возвращении на один шаг назад не достигается ли начало СП, и осуществляется переход к 11 - шагу, если при возвращении достигается начальный узел.

9. Один шаг назад по СП. Для рассмотрения выбирается предыдущая точка СП.

10. Определение следующей точки, находящейся от рассматриваемой точки на расстоянии элементарного пути и переход к 3 – шагу.

11. Завершение алгоритма.

12. На рис. представлена структурная схема алгоритма нахождения сигнальных путей между заданными узлами графовой модели.

Приведенный алгортм является универсальным, так как его можно использовать для поиска специальных путей (прямых путей, маршрутов, контуров и других структур). Он по сравнению с известными обеспечивает преемственность и высокую скорость, а также простоту реализации на алгоритмических языках. Алгоритм реализован на алгоритмическом языке С++, его можно использовать самостоятельно или в составе существующих программных комплексов и САПР роботов.

Литература

1. *Якубов А. Б. Автоматизация программирования роботов на основе алгоритма поиска специальных путей в графах. //Методы науки. –Уфа: ООО, «Инфинити», №7, 2017, С. 30-31*

2. *Якубов А.Б. Алгоритмы анализа схем оптоэлектронных устройств. - В кн. Теория и машинное проектирование радиоэлектронных цепей и преобразователей информации. – Ташкент: ТашПИ, 1981, вып.322, с. 38-46.*

МУНДАРИЖА