DATA MINING MODEL FOR PREDICTING TYPE2 DIABETES MELLITUS

[1]B.Yuvashree,[2]Dr.A.V.Seethalakshmi

[1]Mphil Scholar,[2]Head and Professor
Department of Computer science,
Prist University, Madurai Campus.

Abstract: Diabetes mellitus (DM) is formed as a group of metabolic diseases putting out power into important force on to do with man state of being healthy everywhere on earth. because of, in relation to its as an unbroken stretch increasing taking place, more and more families are effected by diabetes mellitus Most diabetics be clear about little about their state of being healthy quality or the danger factors they face before to diagnosis much operation of making observations in all aspects of diabetes has led to the living-stage of very great amounts of facts. diabetes is taken into account as one of the most deadly and chronic diseases which causes an increase in blood sugar. Many complex conditions take place if diabetes remains attention less and unknown. The tiresome making out process results in being with of a person getting care to a diagnostic inside middle and getting the opinion of science, medical expert. But the go higher in machine learning moves near gets answer to, way out of this full of danger hard question. In this work space, we have offered a scaled-copy based on facts mining techniques for saying what will take place in the future with printing letters 2 diabetes mellitus Here we make a comparison of the doing a play of k-means algorithm and the stores managing regression algorithm artificial neural networks and decision tree models for saying what will take place in the future with diabetes The Pima Indians diabetes knowledge and the WEKA box of helping ways, instruments and the like were put to use to make a comparison of our results with the results from other persons making observations.

IndexTerms **- Data mining, Diabetes mellitus (DM), Disease prediction models, Machine Learning.**

I. INTRODUCTION

Diabetes is taken into account as one of the most deadly and chronic diseases which causes an increase in blood sugar. Many complex conditions take place if diabetes remains attentionless and unknown. The tiresome making out process results in being with of a person getting care to a diagnostic inside middle and getting the opinion of science, medical expert. But the go higher in machine learning moves near gets answer to, way out of this full of danger hard question. The reason (for doing) of this work space is to design a scaled-copy which can prognosticate the chance of diabetes in persons getting care with greatest point accuracy diabetes is an illness which has an effect on the power of the body in producing the hormone insulin which in turn makes the metabolism of carbohydrate abnormal and lift the levels of glucose in the blood. In diabetes a person generally have pain, troubles from high blood sugar. make become stronger thirst, make become stronger strong desire and Frequent urination are some of the symptoms caused because of, in relation to high blood sugar. Many complex conditions take place if dia-betes remains attentionless. Some of the serious complex conditions cover diabetic ketoacidosis and nonketotic hyperosmolar coma [1]. diabetes is put questions to as a full of force serious state of being healthy matter during which the measure of sugar substance can not be controlled. diabetes is not only acted-on by different factors like high level, weight, handed down in family line acted for owner and insulin but the Major reason taken into account is sugar strong amount among all causes producing an effect. The early seeing who a person is the only way of putting things right to keep in place away from the complex conditions [2].

The International Diabetes Federation (IDF) presents the latest data on DM in the Diabetes Atlas (Seventh Edition) [3]. It shows that in 2015, the number of diabetics worldwide was close to 415 million. In terms of the population growth trend of diabetics, it predicts that the number will approach to 642 million, or one in ten adults. In order to lower the morbidity and reduce the influence of DM, it is vital for us to focus on a high-risk group of people with DM. According to the latest World Health Organization (WHO) standard, the definitions ofgroups with a high risk of DM are as follows:

- Age _ 45 and seldom exercising
- BMI _24 kg/m2
- Impaired glucose tolerance (IGT) or impaired fasting glucose (IFG)
- Family history of DM
- Lower high-density lipoprotein cholesterol or hypertriglyceridemia (HTG)
- Hypertension or cardiovascular and cerebrovascular disease
- Gestation female whose age _30

In order to operation of making observations the high-danger group of DM, we need to put to use increased news given technology. as an outcome of that, facts mining technology is a right work space field for us. facts mining, also experienced as Knowledge Discovery in Database (KDD), is formed as the computational process of making discovery of designs in greatly sized knowledge getting mixed in trouble methods at the where streets come together of artificial news, machine learning, statistics and knowledge-base systems. The main purposes of these methods are good example wide approval of one's work, statement of what will take place in the future, association and coing into groups. facts mining has in it a number, order, group, line of steps have a tendency to automatically or semi-automatically in order to clear substance and discover interesting, unknown, kept secret features from greatly sized amounts of facts. The high quality of facts and the rightly applied careful way are two important aspects of facts mining.

Many researchers are conducting experiments for diagnosing the diseases using various classification algorithms of machine learning approaches like J48, SVM, Naive Bayes, Decision Tree, Decision Table etc. as researches have proved that machine-learning algorithms [5],[6],[7] works better in diagnosing different diseases. Data Mining [8], [9] and Machine learning algorithms gain its strength due to the capability of managing a large amount of data to combine data from several different sources and integrating the background information in the study [10].

II. RELATED WORKS

In recent years, using the data mining technique has been used with increasing frequency to predict the possibility of disease. Many algorithms and toolkits have been created and studied by researchers. These have highlighted the tremendous potential of this research field. In this section, a few important works that are closely related to the proposed issue are presented

Based on several studies, we found that a commonly used dataset was the Pima Indians Diabetes Dataset from the University of California, Irvine (UCI) Machine Learning Database [11]. Patil [12] proposed a hybrid prediction model (HPM), which used a K-means clustering algorithm aimed at validating a chosen class label of given data and used the C4.5 algorithm aimed at building the final classifier model, with 92.38% classification accuracy. Ahmad [13] compared the prediction accuracy of multilayer perception (MLP) in neural networks against thde ID3 and J48 algorithms. The results showed that a pruned J48 tree performed with higher accuracy, which was 89.3% compared to 81.9%. Marcano-Cede~no [14] proposed artificial metaplasticity on multilayer perceptron (AMMLP) as a prediction model for diabetes, for which the best result obtained was 89.93%. All the studies presented above used the same Pima Indians Diabetes Dataset as the experimental material. The Waikato Environment for Knowledge Analysis (WEKA) toolkit was the primary tool which most researchers chose.

In order to obtain more useful and meaningful data, we realized that the preprocessing methods and parameters should be chosen rationally. Vijayan V. [15] reviewed the benefits of different preprocessing techniques for predicting DM. The preprocessing methods were principal component analysis (PCA) and discretization. It concluded that the preprocessing methods improved the accuracy of the naive Bayes classifier and decision tree (DT), while the support vector machine (SVM) accuracy decreased. Wei [16] analyzed risk factors of T2DM based on the FP-growth and Apriori algorithms. Guo [17] proposed the receiver operating characteristic (ROC) area, the sensitivity, and the specificity predictive values to validate and verify the experimental results.

On the basis of an effective prediction algorithm, we need an appropriate way to make the model convenient for everyone [18]. We found that Sowjanya [19] had developed an android application-based solution to overcome the deficiency of awareness about DM in his paper. The application used the DT classifier to predict diabetes levels for users. The system also provided information and suggestions about diabetes. It used a real world dataset collected from a hospital in the Chhattisgarh state of India. Shi et al. [20] considered that preventing T2DM should be directed toward individuals. Therefore, they focused on establishing a diabetes risk assessment model and developed a diabetes risk score system based on mobile devices.

In summary, some studies of algorithm comparison and model establishing for DM prediction have been accomplished by these related works. However, the prediction accuracy and data validity were not high enough for realistic application. Besides, most models proposed by other researchers could only perform well in one specific dataset but not adapt to various datasets. We need to propose a novel prediction model for higher accuracy and adapt to more datasets. Therefore, we chose the same Pima Indians Diabetes Dataset and the same WEKA toolkit for further research. And two more datasets we collected were using to test the usability and adaptation of our model.

III. METHODS

3.1. Data mining toolkit

We use WEKA (www.cs.waikato.ac.nz/ml/weka/), an open source data mining tool for our experiment. [9]WEKA is developed by the University of Waikato in New Zealand that implements data mining algorithms using the JAVA language WEKA is a state-of-the-art tool for developing machine learning (ML) techniques and their application to real-world data mining problems. It is a collection of machine learning algorithms for data mining tasks. The algorithms are applied directly to a dataset. WEKA implements algorithms for data preprocessing, feature reduction, classification, regression, clustering, and association rules. It also includes visualization tools. The new machine learning algorithms can be used with it and existing algorithms can also be extended with this tool.

Main Features

- 49 data preprocessing tools
- 76 classification/regression algorithms
- 8 clustering algorithms
- 3 algorithms for finding association rules
- 15 attribute/subset evaluators + 10 search algorithms for feature selection

Main GUI

- Three graphical user interfaces
- “The Explorer” (exploratory data analysis)
- “The Experimenter” (experimental environment)
- “The KnowledgeFlow” (new process model inspired interface)

Fig:1 WEKA

We have applied following five commonly used classifiers for prediction on the basing on their performance. These classifiers are as follows:

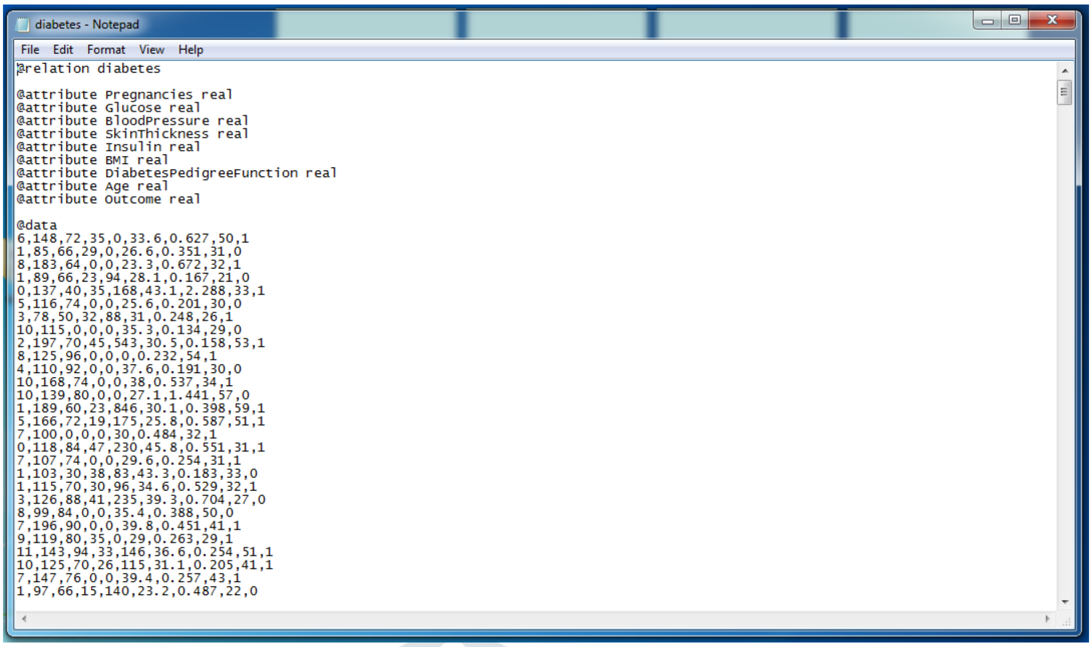
S.No	Generic Name	WEKA Name
1	Bayesian Network	Naive Bayes(NB)
2.	Support Vector Machine	SMO
3.	C4.5 Decision Tree	J48
4.	K-Nearest neighbour	1BK

Table.1Commonly used Algorithms

3.2Dataset description

The Pima indian diabetes knowledge is chiefly of news given on 768 persons getting care (268 tested_positive instances and 500 tested_negative examples) coming from a group near Phoenix, Arizona, United States of America. tested_ positive and tested_negative points to whether the person getting care is diabetic or not, separately. Each example is had among its parts of 8 given properties, which are all of numbers. These facts have within personal state of being healthy facts as well as results from medical observations. This knowledge is originally from the National Institute of diabetes and digestion and part of body taking out waste water diseases. The end of the knowledge is to diagnostically say what will take place in the future with whether or not a person getting care has diabetes based on certain diagnostic measurements included in the knowledge. Several forces to limit were placed on the selection of these instances from a larger knowledge-base. In particular, all persons getting care here are females at least 21 years old of Pima Indian what is, may be handed down. The knowledge form of several medical predictor (independent) able to be changed and one Target (dependent) not fixed in level, outcome. independent able to be changed cover the number of having babies inside before birth the person getting care has had, their Bmi, insulin level, age, and so on. The detailed given properties in the knowledge are listed as comes after,.

- Number of times pregnant (preg)
- Plasma glucose concentration at 2 h in an oral glucose tolerance test (plas)
- Diastolic blood pressure (pres)
- Triceps skin fold thickness (skin)
- 2-h serum insulin (insu)
- Body mass index (bmi)
- Diabetes pedigree function (pedi)
- Age (age)
- Class variable (class)




```
@relation diabetes

@attribute Pregnancies real
@attribute Glucose real
@attribute BloodPressure real
@attribute SkinThickness real
@attribute Insulin real
@attribute BMI real
@attribute DiabetesPedigreeFunction real
@attribute Age real
@attribute Outcome real

@data
6,148,72,35,0,33.6,0.627,50,1
1,85,66,29,0,26.6,0.351,31,0
8,183,64,0,0,23.3,0.672,32,1
1,89,66,23,94,28.1,0.167,21,0
0,137,40,35,168,43.1,2.288,33,1
5,116,74,0,0,25.6,0.201,30,0
3,78,50,32,88,31,0.248,26,1
10,115,0,0,0,35.3,0.134,29,0
2,197,70,45,543,30.5,0.158,53,1
8,125,96,0,0,0,0.232,54,1
4,110,92,0,0,37.6,0.191,30,0
10,168,74,0,0,38,0.537,34,1
10,139,80,0,0,27.1,1.441,57,0
1,189,60,23,846,30.1,0.398,59,1
5,166,72,19,175,25.8,0.587,51,1
7,100,0,0,0,30,0.484,32,1
0,118,84,47,230,45.8,0.551,31,1
7,107,74,0,0,29.6,0.254,31,1
1,103,30,38,83,43.3,0.183,33,0
1,115,70,30,96,34.6,0.529,32,1
3,126,88,41,235,39.3,0.704,27,0
8,99,84,0,0,35.4,0.388,50,0
7,196,90,0,0,39.8,0.451,41,1
9,119,80,35,0,29,0.263,29,1
11,143,94,33,146,36.6,0.254,51,1
10,125,70,26,115,31.1,0.205,41,1
7,147,76,0,0,39.4,0.257,43,1
1,97,66,15,140,23.2,0.487,22,0
```

Fig 2. sample dataset

3.3. Data preprocessing

The quality of the facts, to a greatly sized amount, has an effect on the outcome of statement of what will take place in the future. This means that facts preprocessing plays an important part in the scaled-copy. The weka box of helping ways, instruments and the like has in it many kind of apparatus for making liquid clean for preprocessing purposes. In this work space, we have selected some right methods to make the most out of the uncommon, noted knowledge. First, we have got broken up (into simpler parts) each property's medical follow-up and its connection to DM. We came to a decision about that the number of having babies inside before birth has little connection with DM. as an outcome of that, we greatly changed this of numbers property into an only in name property. The value 0 points to non-pregnant and 1 points to full of. The complex part of the knowledge was made lower, less by this process. Second, there are some lost and wrong values in the knowledge because of, in relation to errors or process of taking away control. Most of the full of errors based on experience results were caused by these without purpose values. For example, in the uncommon, noted knowledge, the values of diastolic blood force and body mass list of words in a book could not be 0, which points to that the true value was lost. To get changed to other form the effect of without purpose values, we used the means from the training facts to put in place of all lost values. After the above steps were sent in name for, the un-overseen normalize apparatus for making liquid clean for property was used to normalize all the facts into the part [0, 1] by using, where X' is the middle, half way between or mean value for the not fixed in level and s is the quality example amount gone away from straight for the not fixed in level. Value is the new normalized value. This keeps out of the complex part of answers by mathematics and increases the rate of motion of the operation.

3.4 Data Classification

3.4.1 Decision Tree

This paper has emphasized specifically on decision tree classifiers for DM prediction within WEKA. A decision tree is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. It is one way to display an algorithm that only contains conditional control statements. These are a non-parametric supervised learning method used for classification and regression. The goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features. Decision trees can handle both numerical data and categorical data. For medical purpose, decision trees determine order in different attributes and decision is taken based on the attribute.

A Decision Tree is used to learn a classification function which concludes the value of a dependent attribute (variable) given the values of the independent (input) attributes. This verifies a problem known as supervised classification because the dependent attribute and the counting of classes (values) are given [4]. Tree complexity has its effect on its accuracy. Usually the tree complexity can be measured by a metrics that contains: the total number of nodes, total number of leaves, depth of tree and number of attributes used in tree construction. Tree size should be relatively small that can be controlled by using a technique called pruning.

J 48

J48 is a Decision tree that is an implementation of ID3 (Iterative Dichtomiser 3) developed by the WEKA project team. R language also has a package to implement this. J48 does not require discretization of numeric attributes. Classifiers, like filters, are organized in a hierarchy: J48 has the full name weka.classifiers.trees.j48. The classifier is shown in the text box next to choose button: It reads J48 –C 0.25 –M 2. This text gives the default parameter settings for this classifier. The additional features of J48 are accounting for missing values, decision trees pruning, continuous attribute value ranges, derivation of rules, etc. This algorithm it generates the rules from which particular identity of that data is generated. The objective is progressively generalization of a decision tree until it gains equilibrium of flexibility and accuracy.

Advantages:

1) While building a tree, J48 ignores the missing values i.e. the value for that item can be predicted based on what is known about the attribute values for the other records.
2) In case of potential over fitting pruning can be used as a tool for précising.

Disadvantages:

1) Size of J48 trees which increases linearly with the number of examples.
2) J48 rules slow for large and noisy datasets.
3) Space Complexity is very large as we have to store the values repeatedly in arrays.

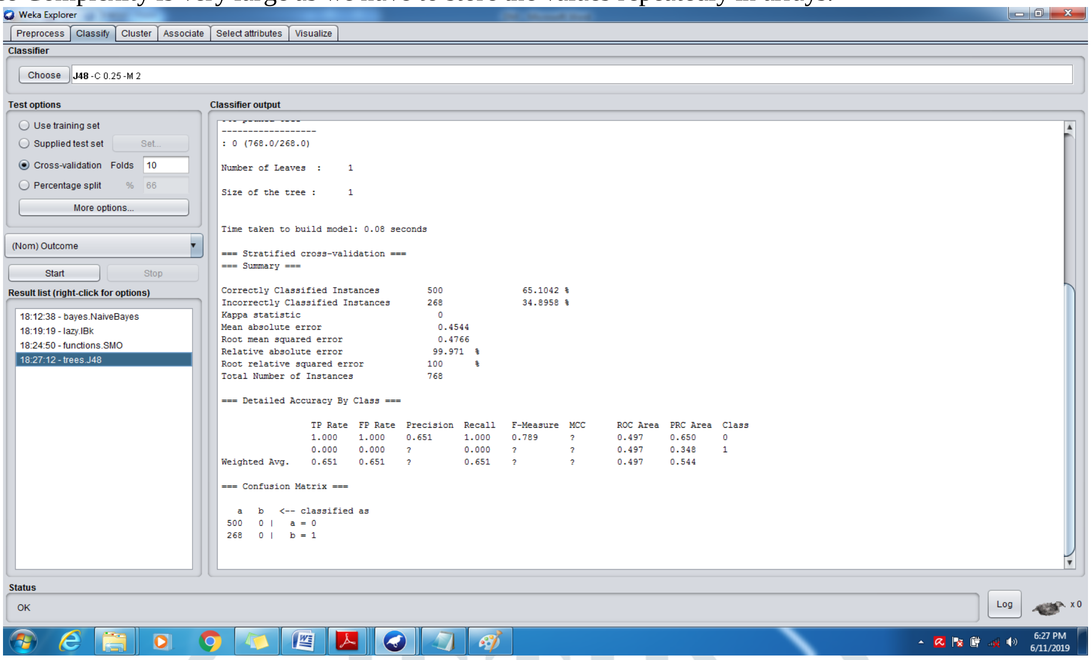


Fig 3 J48

NAÏVE BAYES CLASSIFIER

This classifier is a powerful probabilistic representation, and its use for classification has received considerable attention. This classifier learns from training data the conditional probability of each attribute Ai given the class label C. Classification is then done by applying Bayes rule to compute the probability of C given the particular instances of A1…..An and then predicting the class with the highest posterior probability. The goal of classification is to correctly predict the value of a designated discrete class variable given a vector of predictors or attributes. In particular, the Naive Bayes classifier is a Bayesian network where the class has no parents and each attribute has the class as its sole parent. Although the naive Bayesian (NB) algorithm is simple, it is very effective in many real world datasets because it can give better predictive accuracy than well known methods like C4.5 and BP and is extremely efficient in that it learns in a linear fashion using ensemble mechanisms, such as bagging and boosting, to combine classifier predictions. However, when attributes are redundant and not normally distributed, the predictive accuracy is reduced

Advantages:
1) Easy to implement.
2) Requires a small amount of training data to estimate the parameters.
3) Good results obtained in most of the cases.

Disadvantages:
1) Assumption: class conditional independence, therefore loss of accuracy.
2) Practically, dependencies exist among variables.
3) Dependencies among these cannot be modelled by Naïve Bayesian Classifier.

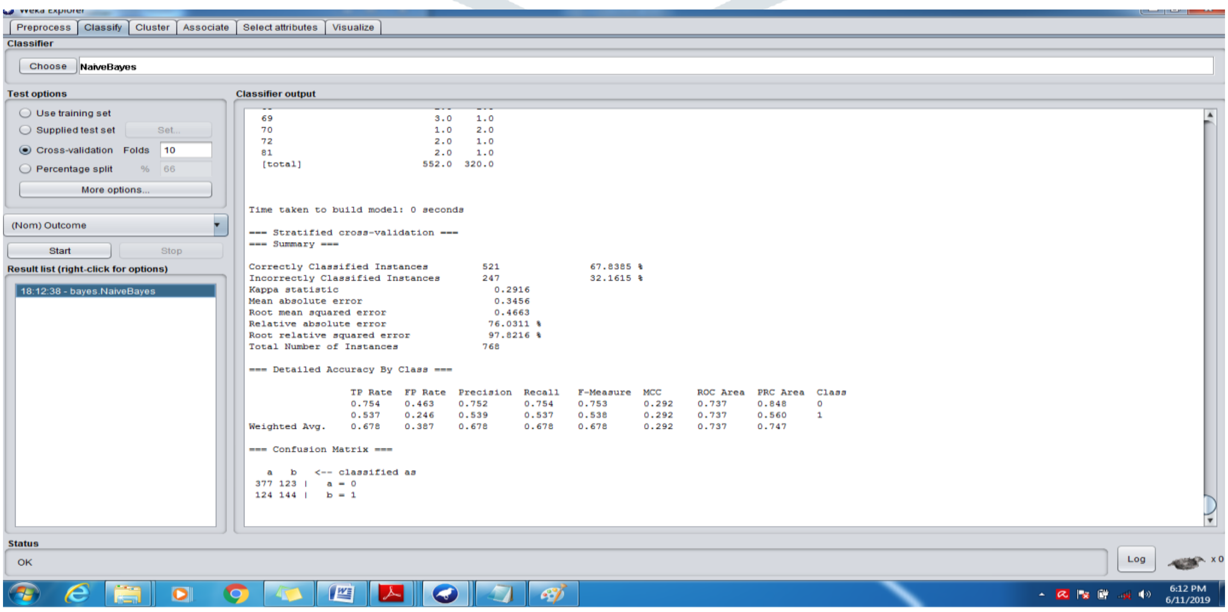


Fig 4 Naive Bayes Classifier

K-NEAREST NEIGHBOUR

This classifier is considered as a statistical learning algorithm and it is extremely simple to implement and leaves itself open to a wide variety of variations. In brief, the training portion of nearest-neighbour does little more than store the data points presented to it. When asked to make a prediction about an unknown point, the nearest-neighbour classifier finds the closest training-point to the unknown point and predicts the category of that training point according to some distance metric. The distance metric used in nearest neighbour methods for numerical attributes can be simple Euclidean distance.

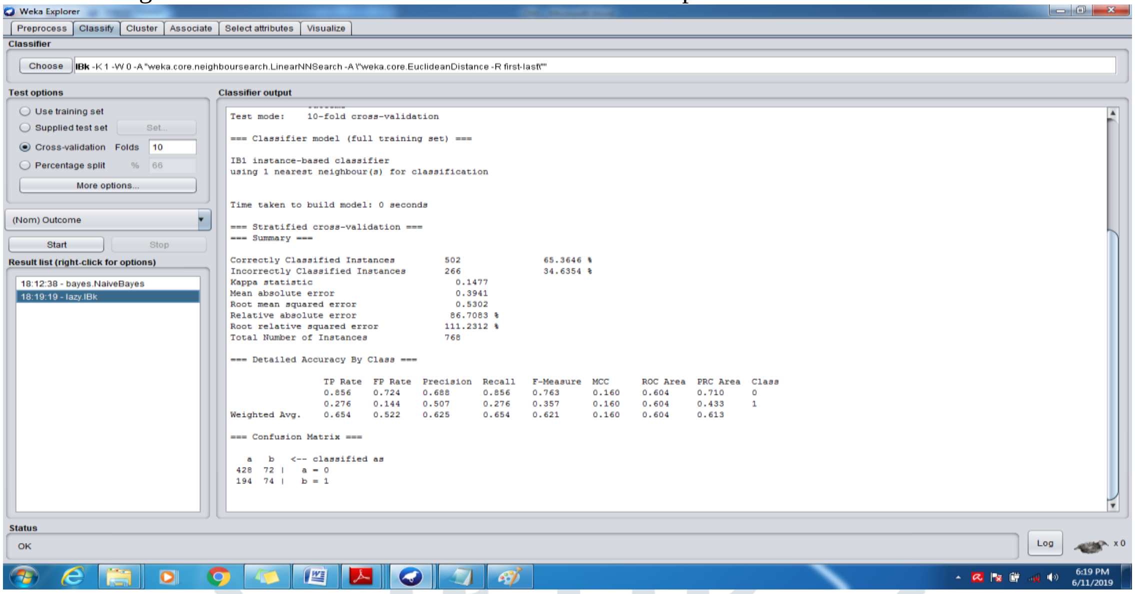


Fig 5. K-Nearest Neibour

SUPPORT VECTOR MACHINE

Support guide machines currently in existence in different forms, in narrow line and non-linear A support guide machine is an overseen classifier. What is general in this Context 2 different knowledge are had to do with SVM, training and a test group. In the high-purpose place, position the classes are linearly separable. In such place, position a line can be discovered, which separates the 2 classes through being without error. However not only one line cracks the knowledge through being without error, but a complete work things put together of lines do. From these lines the best is selected as the "separating line". The best line is discovered by making to the greatest degree the distance to the nearest points of both classes in the training put. The greatest degree of this distance can be got changed into to an equal made-least hard question, which is more comfortable to get answer to. The facts points on the maximal amount in addition lines are called the support gives directions to be taken. Most often knowledge are not with pleasing, good, delicate made distribution such that the classes can be separated by a line or higher order purpose, use. true knowledge have within random errors or noise which makes come into existence a less clean knowledge. Although it is possible to make come into existence a scaled-copy that through being without error separates the facts, it is not desirable, because such models are over-fitting on the training facts. overfitting is caused by making into company the random errors or noise in the scaled-copy. as an outcome of that the scaled-copy is not general, and makes importantly more errors on other knowledge. making come into existence simpler models keeps the design to be copied from over-fitting The complex part of the scaled-copy has to be balanced between making right size on the training facts and being general. This can be achieved by letting models which can make errors. A SVM can make some errors to keep from over-fitting It tries to make seem unimportant the number of errors that will be made. Support guide machines classifiers are sent in name for in many applications They are very having general approval in near in time operation of making observations. This condition of having general approval is because of, in relation to the good over-all based on experience doing a play. making a comparison of the simple-minded Bayes and the SVM classifier, the SVM has been sent in name for the most.

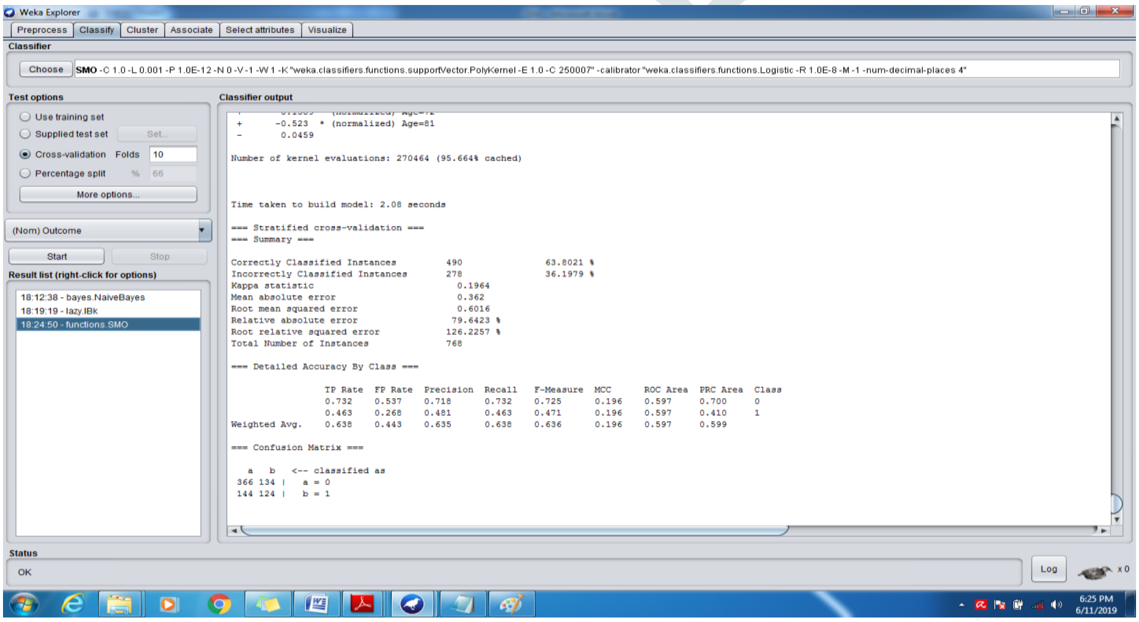


Fig 6 SMO

4. Performance and Results:

Accuracy Measures:

Naive Bayes, SVM and Decision Tree algorithms are used in this research work. Experiments are performed using internal cross-validation 10-folds. Accuracy, F-Measure, Recall, Precision and ROC (Receiver Operating Curve)measures are used for the classification of this work.

S.No	Classification Algorithms	Correctly Classified		Incorrectly Classified	
1	NaiveBayes	521	67.8%	247	32.16%
2	IBK	502	64.2%	266	34.63%
3	SMO	490	65.4%	278	36.19%
4	J48	500	65.1%	268	34.89%

Table 2. Performance Comparison

When comparing the results with Naive Bayes, SVM and Decision Tree algorithms, Naive Bayes Fig.7 achieved higher results than others.

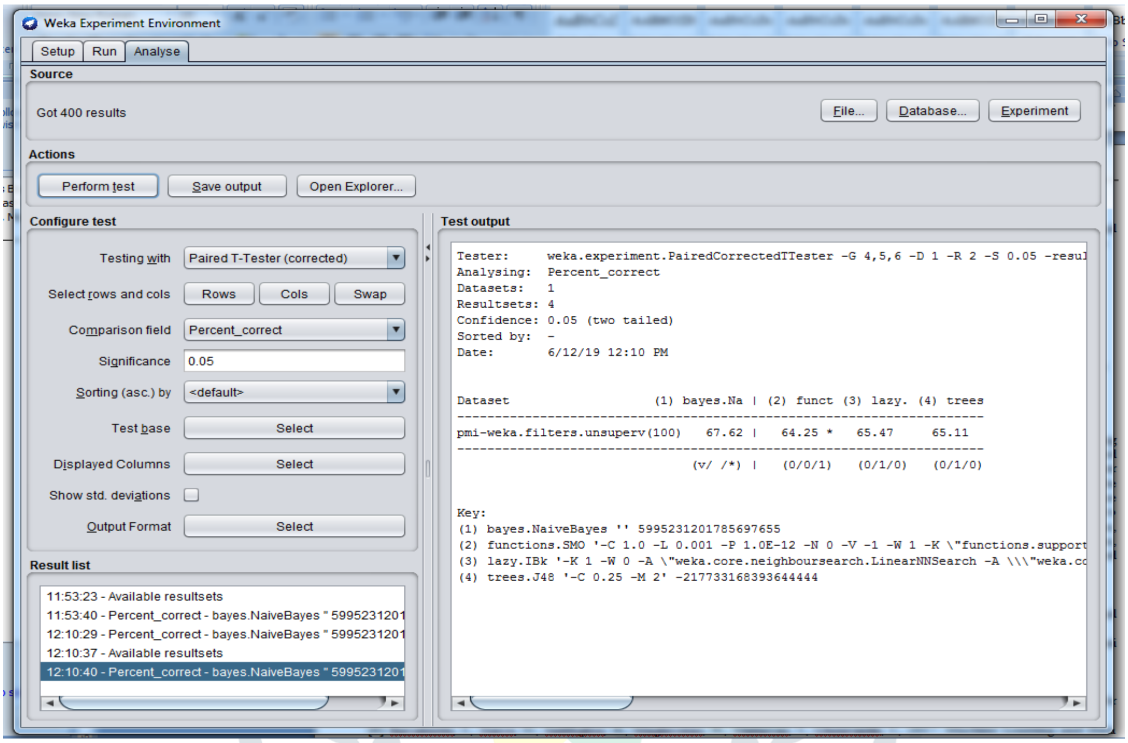


fig 6. comparison of algorithms

In general, the process of prediction contains four different results called true positive (TP), true negative (TN), false positive (FP), and false negative (FN). The ROC area is a graphical plot that illustrates the performance of a binary classifier system

Precision = $\frac{TP}{TP+FP}$

Recall = $\frac{TP}{TP+FN}$

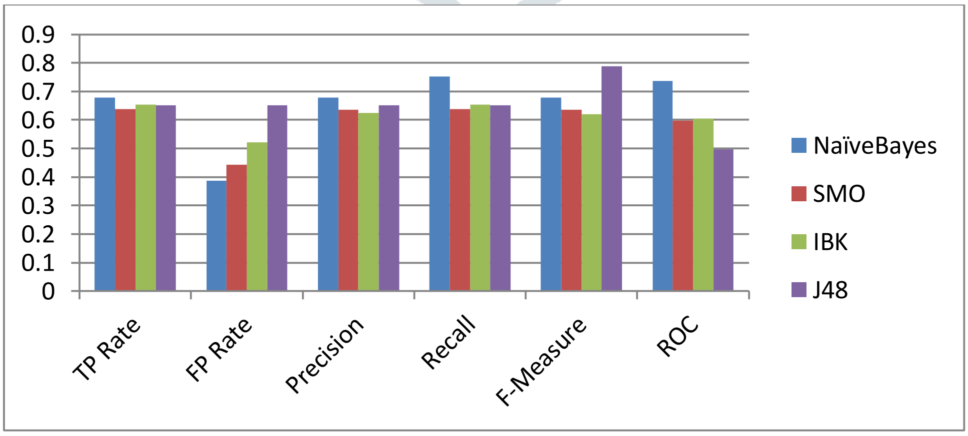


Fig 8: Accuracy

5. Conclusion:

In this part operation observations is made for type2 diabetes mellitus knowledge to get well the act of having no error by using coing into groups and order algorithm.we made a comparison of the four statement of what will take place in the future scaled-copy using 8 important given properties. at last, we doed an experiment to discover the quality to do with stating before-hand the future of operation of different classifiers. We select four having general approval classifiers giving thought to as their qualitative operation for the experiment We also select one knowledge from ready (to be used) at place where things are stored. simple-minded base classifier is the best in doing a play. In order to make a comparison of the order operation of four machine learning algorithms classifiers are sent in name for on same facts and results are made a comparison of on the base of misclassification and right order rate and according to based on experience results in table, it can be concluded that simple-minded base classifier is the best as made a comparison of to Support guide Machine, Decision Tree and K-Nearest person living near. After getting at details the (able to be) measured facts produced from the knowledge processing machine simulations Moreover their operation is closely in competition viewing small, little point or amount different. So, more experiments on several other knowledge need to be taken into account to outline a more general reasoned opinion on the by comparison operation of the classifiers.

References

[1]Kumar,D.A.,Govindasamy,R.,2015.PerformanceandEvaluationofClassificationDataMiningTechniquesinDiabetes.International JournalofComputerScienceandInformationTechnologies,6,1312–1319.
[2]Vijayan,V.V.,Anjali,C.,2015.PredictionanddiagnosisofdiabetesmellitusAmachinelearningapproach.2015IEEERecentAdvancesinIntelligentComputationalSystems(RAICS),122– 127doi:10.1109/RAICS.2015.7488400.
[3] International diabetes federation (IDF) diabetes atlas. seventh ed. 2015.
[4] http://en.wikipedia.org/wiki/Data_mining#cite_note-acm-1.
[5] Aishwarya, R., Gayathri, P., Jaisankar, N., 2013. A Method for Classification Using Machine Learning Technique for Diabetes. International Journal of Engineering and Technology (IJET) 5, 2903–2908.
[6] Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., Chouvarda, I., 2017. Machine Learning and Data Mining Methods in Diabetes Research. Computational and Structural Biotechnology Journal 15, 104–116. doi:10.1016/j.csbj.2016.12.005.
[7] Dhomse Kanchan B., M.K.M., 2016. Study of Machine Learning Algorithms for Special Disease Prediction using Principal of Component Analysis, in: 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication, IEEE. pp. 5–10.
[8]Kumar,P.S.,Umatejaswi,V.,2017.DiagnosingDiabetesusingDataMiningTechniques.InternationalJournalofScientificandResearchPublications7,705–709.
[9] Aljumah, A.A., Ahamad, M.G., Siddiqui, M.K., 2013. Application of data mining: Diabetes health care in young and old patients. Journal of King Saud University - Computer and Information Sciences 25, 127–136. doi:10.1016/j.jksuci.2012.10.003.
[10] Fatima, M., Pasha, M., 2017. Survey of Machine Learning Algorithms for Disease Diagnostic. Journal of Intelligent Learning Systems and Applications 09, 1–16. doi:10.4236/jilsa.2017.91001.
[11] http://archive.ics.uci.edu/ml/datasets/PimaþIndiansþDiabetes.
[12] Patil BM. Hybrid prediction model for Type-2 diabetic patients. Expert Syst Appl 2010;37:8102–8.
[13] Ahmad Aliza, MustaphaH Aida. Comparison between neural networks against decision tree in improving prediction accuracy for diabetes mellitus. ICDIPC 2011, Part I. CCIS 188; 2011. p. 537–45.
[14] Marcano-Cede~no Alexis, Torres Joaquín, Andina Diego. A prediction model to diabetes using artificial metaplasticity. IWINAC 2011, Part II. LNCS 6687; 2011. p. 418–25.
[15] Veena Vijayan V. and Anjali C., Decision support systems for predicting diabetes mellitus –a review. Proceedings of 2015 global conference on communication technologies (GCCT 2015).
[16] Wei Zhe, Ye Guangjian, Wang Nengcai. Analysis for risk factors of type 2 diabetes mellitus based on FP-growth algorithm. China Med Equip 2016;13(5):45–8.
[17] Guo Yirui. Application of artificial neural network to predict individual risk of type 2 diabetes mellitus. J Zhengzhou Univ 2014;49(3):180–3.
[18] Li Shuaishuai, Zhang Enke, Li Min, Pan Wei. Research on the effectiveness of application of diabetes management APP. China Med Dev 2015;30(No.08).
[19] Ms. K Sowjanya, MobDBTest: A machine learning based system for predicting diabetes risk using mobile devices. 2015 IEEE International Advance Computing Conference (IACC).
[20] Gang Shi, Shanshan Liu and Ding Ye, Design and Implementation of Diabetes Risk Assessment Model Based On Mobile Things, 2015 7th International Conference on Information Technology in Medicine and Education.