ASIAN JOURNAL OF PHARMACEUTICAL AND CLINICAL RESEARCH

INNOVARE ACADEMIC SCIENCES
Knowledge to Innovation

Vol 10, Issue 9, 2017

Online - 2455-3891
Print - 0974-2441

Review Article

# A NOVEL APPROACH FOR FINDING DIABETIC MELLITUS USING ENSEMBLE MODEL FOR AN OPTIMIZED CLASSIFICATION

**SEKAR KR*, KAMALADEVI M, SETHURAMAN J, RAVICHANDRAN KS**

**Department of CSE and ICT, School of Computing, SASTRA University, Thanjavur, Tamil Nadu, India. Email: sekar_kr@cse.sastra.edu**

*Received: 02 May 2017, Revised and Accepted: 15 June 2017*

**ABSTRACT**

Diabetic mellitus is a chronic disease caused by hyperglycemia which should be treated with high care and medications. The objective of this work is to identify and classify the severity of the diabetic disease using the training data set. This is caused due to the defect in insulin secretion that may affect several organs in the body. Blood pressure and diabetic mellitus are the common twin diseases occurred in about 69.2 million people living in India around 8.7% of the population as per the data resealed in the year 2015. Correct diet, regular exercise will control disease to a great extent. In this research paper the applied methodology is a concurrent classifier for the diabetic mellitus and the results are analyzed with the supervised learning. From the University of California and Irvine repository related attributes for the diabetic mellitus are carefully measured through the ensemble classifier and the results are categorized in the dataset. This work results that boosting can be made to the dataset for obtaining accurate results and classifications. In the conclusion, ensemble methodology is the well proven methodology from the year 1993. For forecasting in "N" number of domains, so for the ensemble classifier produces 93% of the accurate results are made. An audit can be made on the results and suggestions are given to the patients for taking medications with the help of medical practitioners.

**Keywords:** Diabetic mellitus, Boosting, Ensemble classifier, Supervised learning and hyperglycemia.

 DOI: http://dx.doi.org/10.22159/ajpcr.2017.v10i9.19563

## INTRODUCTION

Diabetes mellitus (DM) is a prolonged disease that it should be treated with good medication will provide longevity to the human life. Plenty number of medicines are available in the existing market to cure the disease in Allopathic, Ayurveda and Siddha. It is possible to control the disease, but eradicating the disease to the core is not successful till such time. Heart attack cannot be realized by the diabetic mellitus patients, so taking a remedial measure for the heart attack cannot be sometimes taken at a right moment will cause great disastrous to the gifted human life. Nowadays many decision support system, mining techniques, big data analytics, and association rules helping the medical physicians to identify and predict whether the patient will be affected by this acute disease or not at an early stage itself. In our research work, 48 attributes related to the diabetic mellitus and 30 tuples are taken for the analysis. The cascading methodology provides great results by boosting the dataset to the ensemble model. The same diabetic mellitus dataset can be given to the battery of classifiers, and the results are appreciable to the core. In this paper, not only the classification results but also some recommendation have given with a clean audit about the patient conditions and suggested some medicine for the remedial cure and care. Diabetes patients should follow proper dosage of drug to maintain blood glucose level. Data mining model using Adaptive Neuro-Fuzzy Inference System (ANFIS) and rough set method is developed to predict the suitable of the drug.

Experimental analysis has been performed for medical records of some patients, and finally, ANFIS give reliable results than rough set methods [1]. Analysis of diabetic treatment for various age groups is predicted using regression based data mining technique. Support vector machine (SVM) algorithm and data set from WHO used for analysis purpose. Only two age group p(o) and p(y) are considered. From the analysis old age people get the treatment immediately and whereas young age people can delay the treatment to avoid side effect [2]. The vigorous forecast model is currently required in all the medicinal areas. Electronic social insurance records are particularly helping them for good forecast because of information exactness. Appropriated information can likewise be taken effectively through the above said framework. Information security is constantly kept up by the store individuals to ensure the delicate information toward the patients. This strategy is connected in sort 2 diabetes for finding the precise expectation [3].

DM is one of the major causes of deteriorating human beings, economic and social factors. Prevention is better than the cure is more appropriate in terms with patients suffering from aliments from diabetes. N number of predictions that have been made through single classifier and nowadays with cascading classifier highly pronounced as an ensemble. Above said methodology provides a greater precision than the legacy classifier [4]. Diabetes is broadly classified into three types such as Type I diabetes, Type II diabetes, and Gestational diabetes. Inadequate productions of insulin lead to Type I diabetes, people can have the longest life with periodic test and good planning food. Regular insulin injection and pills than managing the glucose level control of the patient. Type II DM is a black box model which hits only adults up to 90-95%. Rules based classifiers are employed for the diagnosis but yet find better accuracy. The new classification model has been proposed and noted as recursive rule extraction which is a white box in nature, provides greater accuracy. J48 graft is the classifier having the algorithm re-rules extraction [5]. Characterization method recognizes phenotypes in organized therapeutic records. Highlight space asses by the classifier for different diseases. SVM is utilized for test reason with single element space and two. Of course, a semantically educated component does not factually enhance execution for these models [6].

The A1C test is used for this purpose. Type II diabetic book keeping around 80% of grown ups rather 26 million grown ups influenced by these illnesses. Heavy complication and life-threatening issues are there. Gestational diabetes can be identified through the breathing problem and fetus having abnormal growth. Diabetes is one of the major health problems in all over the world. Using prediction model, DM can predicated through different symptoms and attributes. Different type of classifier is used to predict the diseases at an early stage. Performance measure such as accuracy, sensitivity, and specificity are analyzed. Preprocessed data have high performance than noisy data [7].

Data mining and warehousing is the emerging tool for the business analytics to remove the uncertainty in the future. Information is available in bountiful in all domains and knowledge is the vital part to get the inference. The number of patients suffers from diabetes has increased in future. To prevent the phenomenon, we propose an application using the fuzzy hierarchical model for early detection DM. The performance of the application is compared with the medical decision support system and the outcome has got good accuracy. More than 80% value has matched with medical doctor decision. Proposed method meets the effectiveness and efficiency of early detection of DM [8]. For identifying a diabetic patient with severity and non-severity through the classifier methods using Fuzzy set model is taken into account for classification. Data and the dataset semantics were taken from University of California and Irvine (UCI) to Pima Indians diabetes (PID) repository. All the diagnosis expert systems provide a greater accuracy using classifiers with their high precision values. Preprocessing and fuzzification are the two major steps to be caring out before deploying into the classifier [9].

In data mining, classification clustering and association are very much be appreciated not only in business world but also lifesaving medical science. Clustering is non-supervised learning, there we can find and predict the behavior. Association frequent item set and basket product analysis can be possible in wide range for the enterprising market in terms with reorder of commodities. In this paper proposed methodologies in terms with classification. It provides greater vision and predictability about the DM. Prism classifier tree induction classifier and recursive classifier are deployed for this DM treatment prediction. Classifier predicts class label as long term medication, short-term medication and no medication through 36 vital attributes.

## RELATED WORKS

Massive health records find difficulty in terms with complexity. Predictive analysis algorithm was employed with Map-reduce provides good results in acute DM [10]. Type 2 DM caused because of combination of environmental life factots.Hence cascading classifiers wer employed for the same dataset. The fusion method brings accuracy to the results [11]. The text mining technique can be used to discover new knowledge of diabetes from more than thousands of records from web. Predictive and descriptions are the two approaches used. Descriptive approaches identify explicit reference for diagnosis and treatment of diabetes. Predictive approach predicts the diabetic disease status when the evidence was not explicitly asserted. The finding is compared with medical records [12]. In the previous couple of years, numerous calculations have been created for the computerized identification of a particular pathology, normally diabetic retinopathy, and utilizing eye fundamental photography. The calculation does not concentrate on individual pictures: It considers examination records comprising numerous photos of every retina, together with relevant data about the patient. The primary curiosity is that the substance of examination records (pictures and setting) is portrayed at different levels of spatial and lexical granularity [13]. HBA1c test is more generally used for identifying Type II diabetes. HbA1c test use common cut-off value is 6.5% for the presence of diabetes. However, this common cut-off value is inconsistent for large trials. For large data sets, some biomarker data can be added with a data mining algorithm for estimating the optimal cut-off value of HBA1c. Investigating the extra bookmaker will improve the precision of diagnosing the Sort II diabetes [14]. Detection algorithm for various diseases is highly sensitive and reportedly may false positive that lead to class imbalance problem. To reduce the false positive rate ensemble based adaptive oversampling algorithm is used and it gives better results in AUC and geometric mean [15]. In the major metropolitan city Type-II diabetic test was conducted, and random sample people were tested for the above. It is very extensive test and cost effective to maximize the health benefit [16]. Due to metabolic disorder, the diabetic mellitus was found, and insulin secretions were affected for the patients. Ayurveda treatment is the best one to treat diabetes. Good food, habits, and exercise will reduce and maintain the good health [17]. Type II DM is the major cause due to diet and lifestyle. This is projected up to 23.9 million in the world population by the year 2030. To decrease incretin level dipeptidyl peptidase (DPP-4) an enzyme plays a key role. Metformin is more effective than other drugs [18]. Nutraceuticals is a supplementary which acts as an alternative drug of choice. So many chemical agents are available to treat diabetic patients. Nutrients and plants are the synthetic agents to provide a good result in diabetes [19].

## APPLIED METHODOLOGIES

Methodology 1: Decision tree induction classification for the dataset applied in WEKA tool out of "N" classifiers in the existing market here three main classifiers are employed for the prediction. How the three classifiers playing the dominant roles are adhered. Qualities of decision tree follow easy interpretation and explanation can be seen evidently. Outliers are handled by the decision tree without having any trouble provides the possible outcomes. Feature selection can be made through this classifier. Dataset preparation is simple in this classifier. Non-relationship attributes will affect the outcome result.

## TRAINED DATASET

**Methodology 2: Naive Bayes theorem applied in R-tool**

Bayesian classification has got brilliant qualities. (i) Training the dataset is enough to predict the result, (ii) it is in a semi-supervised learning too, (iii) in the embarrassing situation it can predict an optimal result, and (iv) very time consuming compared with other new classifiers. Expert knowledge and intuition are the inbuilt quality for the Bayes classifier. Example 1:

```
> data2 <- read.csv("C:/Users/Welcome/Desktop/final_test1.csv")
> data2
             race gender age weight Admission..type discharge_disposition_id
1 AfricanAmerican   Male  68   0.04              0.06                      0.08
  medical_specialty Diabetic.Type maxFasting.sugar maxPost.Prandial.glucose
1                 0.05                      0.05                      0.03
  HBA1.LEVEL metformin repaglinide nateglinide chlorpropamide glimepiride
1       0.04         0           0           0             0           0
  acetohexamide glipizide glyburide tolbutamide pioglitazone rosiglitazone
1             0         1         0           0            0            0
  acarbose miglitol troglitazone tolazamide examide citoglipton insulin
1        0        0            0          0       0          1       2
  glyburide.metformin glipizide.metformin glimepiride.pioglitazone
1                   0                   0                      0
  metformin.rosiglitazone metformin.pioglitazone diabetesMed readmitted total
1                       0                      0           1          0  5.38
  Class
1    NA
>
> pred <- predict(classifier, data2)
> pred
[1] No MED
```

```
> data2 <- read.csv("C:/Users/Welcome/Desktop/final_test1.csv")
> data2
             race gender age weight Admission..type discharge_disposition_id
1 Caucasian   Male  51   0.05              0.04                      0.03
  medical_specialty Diabetic.Type maxFasting.sugar maxPost.Prandial.glucose
1                 0.05                      0.04                      0.05
  HBA1.LEVEL metformin repaglinide nateglinide chlorpropamide glimepiride
1       0.04         0           1           0             0           0
  acetohexamide glipizide glyburide tolbutamide pioglitazone rosiglitazone
1             0         1         0           0            0            0
  acarbose miglitol troglitazone tolazamide examide citoglipton insulin
1        0        0            0          0       0          1       1
  glyburide.metformin glipizide.metformin glimepiride.pioglitazone
1                   0                   0                      0
  metformin.rosiglitazone metformin.pioglitazone diabetesMed readmitted total
1                       0                      0           1          0  5.35
  Class
1    NA
> pred <- predict(classifier, data2)
> pred
[1] Long MED
```

**Methodology 3: SVM applied R-tool**

SVM has a machine learning classifier provides always high accuracy results because of the linear and non-linear segmentation. Overfitting can be handled very gently by this classification. Normally, the classifier used in text mining, opinion mining and sentiment analysis. But in the class of supervised learning, it always works better. Hence N-dimensional spaces were deployed to predict the accurate result up to 93%.

>model<- train_model (container, "SVM", kernel = "linear", cost = 1)

>predMatrix<- create_matrix (prediction Data, original Matrix=dt Matrix)

| | | | | | | |
|---|---|---|---|---|---|---|
| Caucasian | Female | [50-60) | 0.04 | 0.06 | 0.05 | 0.04 |
| 0.05 | 0.04 | 0.04 | 0.04 | 0 | 0 | |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 1 | 0 | |
| 0 | 0 | 2 | 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 5.36 | | | |

## RESULTS

**SVM_LABEL SVM_PROB**

*Short MED 0.7963774*

Three different classifiers such as decision tree, Bayesian, and SVM are applied to the same dataset contains 38 columns and 25 tuple as training set (Table 1). All the classifier provides a good precision of the

**Table 1: Trained data set for diabetes from UCI after feature selection**

| Race | Gender | Age | Weight | Adm-type | DC-Id | med-Spec | Dia-typ | maxFast-Su | Max Po-Pran glu | HBA1-Lvl | Metformin | Repaglinide |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| African American | Female | 20-30 | 0.03 | 0.02 | 0.03 | 0.01 | 0.02 | 0.03 | 0.03 | 0.04 | 0 | 0 |
| Caucasian | Male | 50-60 | 0.05 | 0.04 | 0.03 | 0.05 | 0.05 | 0.04 | 0.05 | 0.04 | 1 | 0 |
| Caucasian | Female | 80-90 | 0.04 | 0.02 | 0.03 | 0.05 | 0.05 | 0.04 | 0.05 | 0.04 | 1 | 1 |
| Caucasian | Female | 90-100 | 0.03 | 0.02 | 0.08 | 0.04 | 0.05 | 0.05 | 0.06 | 0.04 | 1 | 0 |
| African American | Male | 60-70 | 0.04 | 0.06 | 0.08 | 0.05 | 0.05 | 0.03 | 0.03 | 0.04 | 0 | 0 |
| African American | Female | 50-60 | 0.04 | 0.06 | 0.03 | 0.03 | 0.05 | 0.03 | 0.03 | 0.04 | 0 | 0 |
| African American | Female | 60-70 | 0.04 | 0.06 | 0.03 | 0.05 | 0.05 | 0.04 | 0.04 | 0.04 | 0 | 0 |
| African American | Male | 10-20 | 0.04 | 0.04 | 0.03 | 0.04 | 0.03 | 0.03 | 0.04 | 0.04 | 1 | 0 |
| Caucasian | Female | 60-70 | 0.04 | 0.02 | 0.03 | 0.05 | 0.05 | 0.04 | 0.05 | 0.04 | 1 | 0 |
| Caucasian | Female | 20-30 | 0.04 | 0.02 | 0.03 | 0.03 | 0.02 | 0.02 | 0.04 | 0.04 | 1 | 0 |
| Caucasian | Male | 50-60 | 0.04 | 0.02 | 0.03 | 0.01 | 0.05 | 0.04 | 0.05 | 0.05 | 0 | 1 |
| Caucasian | Female | 20-30 | 0.05 | 0.04 | 0.03 | 0.05 | 0.02 | 0.03 | 0.03 | 0.03 | 1 | 0 |
| African American | Male | 10-20 | 0.02 | 0.04 | 0.03 | 0.03 | 0.03 | 0.02 | 0.03 | 0.03 | 0 | 0 |
| Caucasian | Female | 70-80 | 0.04 | 0.02 | 0.05 | 0.04 | 0.05 | 0.04 | 0.04 | 0.04 | 0 | 0 |
| Caucasian | Male | 50-60 | 0.05 | 0.04 | 0.03 | 0.05 | 0.05 | 0.04 | 0.04 | 0.04 | 0 | 0 |
| Caucasian | Female | 30-40 | 0.04 | 0.04 | 0.03 | 0.03 | 0.02 | 0.03 | 0.03 | 0.03 | 1 | 0 |
| African American | Male | 70-80 | 0.04 | 0.06 | 0.08 | 0.04 | 0.05 | 0.04 | 0.05 | 0.04 | 0 | 0 |
| African American | Female | 10-20 | 0.04 | 0.04 | 0.03 | 0.03 | 0.03 | 0.02 | 0.03 | 0.03 | 0 | 0 |

| Glimepiride | Acetohexamide | Glipizide | Glyburide | Pioglitazone | Rosiglitazone | Acarbose | Miglitol | Troglitazone | Examide | Citoglipton | Insulin | Class |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | Short-Med |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | Short-Med |
| 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | Short-Med |
| 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | Short-Med |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | Short-Med |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 3 | Short-Med |
| 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | No-Med |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Short-Med |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | No-Med |
| 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | Long-Med |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | Long-Med |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | Short-Med |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 3 | Short-Med |
| 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | Short-Med |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | No-Med |
| 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | Long-Med |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | No-Med |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | No-Med |

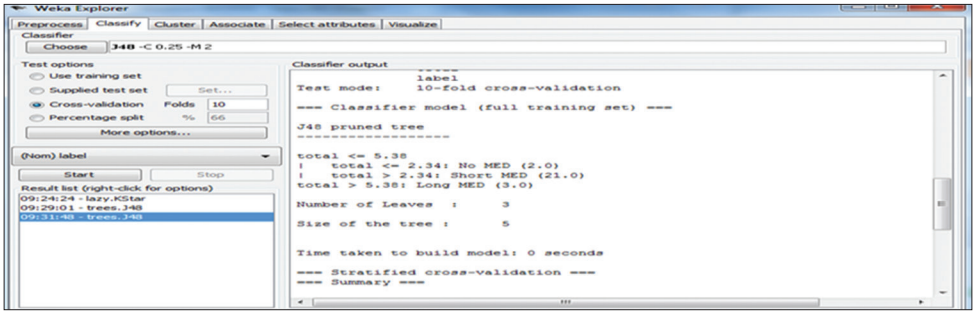


**Fig. 1: Data classification**

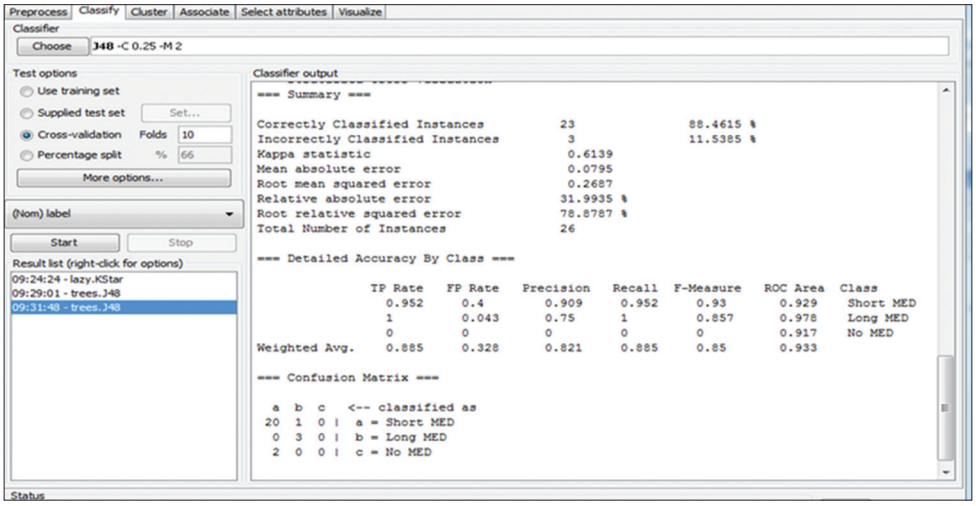


**Fig. 2: Output of the data classification**

answer and identifies a greater relevancy between the classifiers. The incoming pattern to the trained data set with a less entropy provides accurate recommendation it may help the patients to know the peripheral outline for their medication.

The purpose behind applying three classifiers for the same trained dataset is to check the real accuracy of the recommendation. The type of applied methodologies called ensemble model. The variation between the three classifiers is very low. Hence, the recommendation has got a high result. We are recommended the medicine and their dosage. This paper is the eye opener for the research to implement the same concept in the expert system. The results were updated in the methodology itself for the given incoming pattern.

## RESULTS AND DISCUSSIONS

DM trained dataset has taken from the UCI repository has got 78 columns with 348 records using a feature selection methodology; the columns are reduced up to 38 columns and took only 25 records for the investigation. The ensemble model was applied to draw the inference for the incoming pattern. The trained dataset has preprocessed by applying the following methods (i) ordinal data converted to cardinal data using the thresholds. Distributed Methodology was employed to reduce the entropy. Each and every record has taken to identify the class label. Class label and subclass entries were fixed using the threshold called supervised learning. After preprocessed the data three different classifiers were applied to find the recommendation for the incoming patterns. Sample trained dataset has shown in the article, and clear, fast fledged dataset is available in the UCI repository for the reference. No variation found between the three classifiers, so the recommendations for the pattern have got a high precise result (Table 2).

Fig. 1 shows data classification for Bayesian theorem , Fig. 2 shows the results obtained through bayesian, Fig. 3 refers to the tree structure of

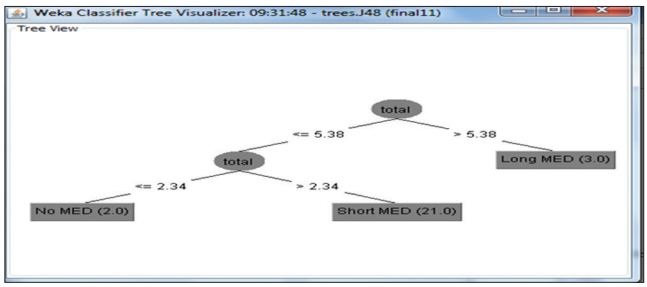


**Fig. 3: Tree visualizer of the data classification**

**Table 2: Comparison about the classifiers**

| Classifier | Performance (%) | Accuracy | Interpretation |
|---|---|---|---|
| Bayesian | 89 | 0.9 | Normal |
| Decision tree | 85 | 0.89 | Normal |
| SVM | 92 | 0.92 | Good |

SVM: Support vector machine

the tree induction classifier and Fig. 4 shows the cost benefit analysis for the data through the baysian theorem shows the clear idea about the classification results.

The above results were obtained through the tools WEKA and r-tools using the factors certainty sensitiveness and interpretation. After the correct audit, some suggestions have made in terms with providing possible medicines in allopathic. Possible treatments for Type II diabetes metformin (Glucophage, Glumetza, others). In general, metformin is the first medication prescribed for Type II diabetes.

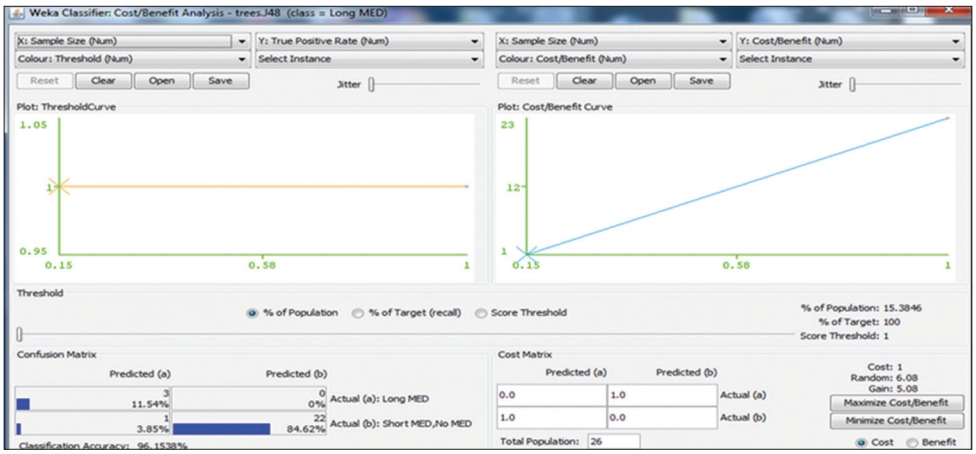


**Fig. 4: Cost Benefit analysis of the data classification**

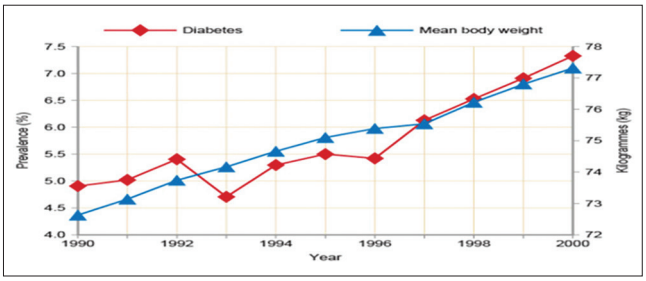


**Fig. 6: Diabetes trend about last decade in terms with mean body weight**

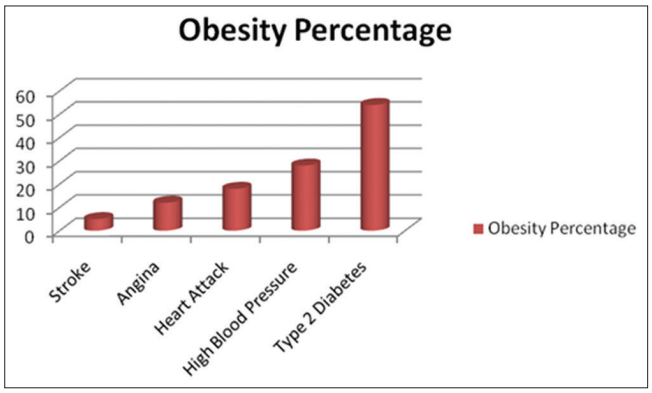


**Fig. 5: Expected body condition toward increasing obesity**

**Table 3: Body condition about increasing percentage in obesity**

| Condition | Obesity percentage |
|---|---|
| Stroke | 5 |
| Angina | 12 |
| Heart attack | 18 |
| High blood pressure | 28 |
| Type 2 diabetes | 54 |

- Sulfonylureas
- DPP-4 inhibitors
- Glucagon-like peptide-1 receptor agonists
- Sodium-glucose co-transporter 2 inhibitors
- Insulin therapy
- Thiazolidinediones
- Meglitinide

All the above said medicines are not to be taken without a consultation with the doctors (medical practitioners) From the Table 3 and Fig. 5 it is inferred that the higher body mass weight leads to disease like stoke, angina, heat attack etc. Fig. 6 shows the 10 years statistics, the people having more mean body weights have the chance of getting diabetes.

## CONCLUSION

The proposed ensemble model is predicting the required period of treatment for DM with an accuracy 0.93, which is predominant when the ensemble with SVM. The ensemble model is helpful for doctors treating DM to recommend the treatment and gain the patient trust. Further, researchers can focus on classifying the medicines with their effectiveness in curing the DM.

## REFERENCES

1. Yıldırım EG, Karahoca A, Uçar T. Dosage planning for diabetes patients using data mining methods. Proc Comput Sci 2011;3:1374-80.
2. Aljumah AA, Ahamad MG, Siddiqui MK. Application of data mining: Diabetes health care in young and old patients. J King Saud Univ Comput Inf Sci 2013;25(2):127-36.
3. Li Y, Bai C, Reddy CK. A distributed ensemble approach for mining healthcare data under privacy constraints. Inf Sci 2016;330:245-59.
4. Perveen S, Shahbaz M, Guergachi A, Keshavjee K. Performance analysis of data mining classification techniques to predict diabetes. Proc Comput Sci 2016;82:115-21.
5. Hayashi Y, Yukita S. Rule extraction using recursive-rule extraction algorithm with J48graft combined with sampling selection techniques for the diagnosis of Type 2 diabetes mellitus in the Pima Indian dataset. Inf Med Unlocked 2016;2:92-104.
6. Kotfila C, Uzuner Ö. A systematic comparison of feature space effects on disease classifier performance for phenotype identification of five diseases. J Biomed Inf 2015;58:S92-102.
7. Kandhasamy JP, Balamurali S. Performance analysis of classifier models to predict diabetes mellitus. Proc Comput Sci 2015;47:45-51.
8. Lukmanto RB, Irwansyah E. The early detection of diabetes mellitus (DM) using fuzzy hierarchical model. Proc Comput Sci 2015;59:312-9.
9. Nahato KB, Nehemiah KH, Kannan A. Hybrid approach using fuzzy sets and extreme learning machine for classifying clinical datasets. Inf Med Unlocked 2016;2:1-11.
10. Eswari T, Sampath P, Lavanya S. Predictive methodology for diabetic data analysis in big data. Proc Comput Sci 2015;50:203-8.
11. Zhu J, Xie Q, Zheng K. An improved early detection method of Type-2 diabetes mellitus using multiple classifier system. Inf Sci 2015;292:1-14.
12. Marir F, Said H, Al-Obeidat F. Mining the web and literature to discover new knowledge about diabetes. Proc Comput Sci 2016;83:1256-61.
13. Quellec G, Lamard M, Erginay A, Chabouis A, Massin P, Cochener B, *et al*. Automatic detection of referral patients due to retinal pathologies through data mining. Med Image Anal 2016;29:47-64.
14. Jelinek HF, Stranieri A, Yatsko A, Venkatraman S. Data analytics

identify glycated haemoglobin co-markers for Type 2 diabetes mellitus diagnosis. Comput Biol Med 2016;75:90-7.
15. Ren F, Cao P, Li W, Zhao D, Zaiane O. Ensemble based adaptive over-sampling method for imbalanced data learning in computer aided detection of micro aneurysm. Comput Med Imaging Graph 2016;55:54-67.
16. Hussain M, Naqvi SB, Khan MA, Rizvi M, Alam S, Abbas A, *et al*. Direct cost of treatment of diabetes mellitus Type 2 in Pakistan. Int J Pharm Pharm Sci 2014;6(11):261-4.
17. Srinivas P, Devi KP, Shailaja B. Diabetes mellitus (madhumeha)-an ayurvedic review. Int J Pharm Pharm Sci 2014;6:107-10.
18. Kaganda O, Singh C, Sachdeva K. Recent advancement in treatment of Type-II diabetes mellitus: A review. Int J Pharm Pract Pharm Sci 2015;2(11):4-14.
19. Pandey M, Kumar V. Nutraceutical supplementation for diabetes: A review. Int J Pharm Pharm Sci 2011;3 Suppl 4:33-40.