Disease Influence Measure Based Diabetic Prediction with Medical Data Set Using Data Mining

B.V. Baiju
Department of Information Technology
Hindustan Institute of Technology and Sceince
Chennai, India
bvbaiju@hindustanuniv.ac.in

Dr.D. John Aravindhar
Department of Computer Science and Engineering
Hindustan Institute of Technology and Sceince
Chennai, India
jaravindhar@hindustanuniv.ac.in

***Abstract* - The problem of diabetic prediction has been well studied in this paper. The disease predictions have been explored using various methods of data mining. The use of medical data set on the prediction of diabetic mellitus has been analyzed. This paper performs a detailed survey on disease prediction using data mining approaches based on diabetic data set. The presence of disease has been identified using the appearance of various symptoms. However, the methods use different features and produces varying accuracy. The result of prediction differs with the methods/measures/ features being used. Towards diabetic prediction, a Disease Influence Measure (DIM) based diabetic prediction has been presented. The method preprocesses the input data set and removes the noisy records. In the second stage, the method estimates disease influence measure (DIM) based on the features of input data point. Based on the DIM value, the method performs diabetic prediction. Different approaches of disease prediction have been considered and their performance in disease prediction has been compared. The analysis result has been presented in detail towards the development.**

Keywords - Medical Data Set, Disease Prediction, Diabetic Mellitus, Data Mining, DIM.

I. Introduction

The changing life style of human being has influenced different impact on the health of humans. One of the influences is the Diabetic Mellitus which has no criteria for its appearance because of it would affect the human at any age. Once, the disease has been appeared, then it will present for their life time. The diabetic has been affected due to the lack of insulin segregation by the pancreas of human body. This introduces higher blood sugar in the body which affects the persons routing life. When the person takes food, it has been converted into glucose and added to the blood. The pancreas has the responsibility to segregate the blood sugar and to control it. When the pancreas does not produce enough insulin to the blood then the appearance of glucose has been identified in the blood which called as diabetes mellitus.

Diabetes mellitus (DM) is considered as the world class health problem which affect independent of human age. WHO has released a note on diabetes as 180 million of human have been affected by Type II diabetes which is of 95% on the population of diabetes. This statistics would increase to double in next 15 years in the entire world. The presence of diabetes has encourages the death in the ratio of 6% in the year 2000 where it would increase upto 45 % in the next 20 years.

However, the disease has been occurred based on the food habits of the human. There are other factors which influence the disease to be occur. As the life style of human has been changed, there is a dramatic change in their food habit, lack of exercise, smoking, having high protein foods, having chunk foods and so on. There are number of factors can be named for the causes of diabetic mellitus. So it is necessary to monitor the changing conditions of blood sugar to identify the presence of disease. Disease prediction is the process of predicting the possibility or identifying the presence of disease based on set of symptoms and their values. The disease prediction can be simply performed by monitoring the glucose level, but the accuracy is based on other features also. It is depends on the age, physical work, the use of calories and their lifestyle.

The diabetes has higher association with micro vascular and macro vascular complications. It includes neuropathy, nephropathy and retinopathy, cardiac diseases. This result in damage of organs and tissues which occurs in one third of the diabetic population. This is highly required that the medical practitioner would identify the prediabetes patients and performs detail study on their glucose tolerance and insulin resistance to avoid vascular complications. Also it has been analysed that the Type 2 diabetes has higher relationship with microvascular dysfunction and peripheral nerve damage throughout the body. The presence of polyneuropathy would introduce vasodilation in the skin. The microvascular disturbances provoke peripheral nerve ischemia, which consecutively aggravates miscrovascular dysfunction.

Neuropathy is an advanced stage of diabetes which affects. Among the diabetics patients, more than 50 percent people have affected with peripheral neuropathy which is monodiabetic nephropathy. They also have frequency of autonomic neuropathy which influences the cardiovascular dysfunction being identified by the abnormal heart rate. The physician generally consider the peripheral neuropathy and provide treatments to balance the damages introduced by PN.

In other dimension, Diabetic nephropathy (DN) is the major issue being identified in both type of diabetic mellitus This increases the amount of albumin in urine which directly affect the renal organs and make them failure. To save the patients from renal failure, it is necessary to identify the neuropathy and nephropathy. There are number of approaches available for the prediction of neuropathy and nephropathy in type II diabetic patients. However, for the prediction, encountering large amount of data collected from numerous diabetic patients is required. Because the accuracy of prediction is always depends on the amount of training data considered.

To perform disease prediction, it needs the history of various diabetic patients. Such records are collected from different medical organizations and named as diabetic data set. The records available in the medical data set can be used to train the system and by submitting the input set of values, the prediction process can be performed. On the other side, there are number of approaches available for the diabetic prediction. The data mining is the concept of mining relevant information from huge set of data base. Such data mining techniques can be adapted to the problem of diabetic prediction. Similarly, there are number of other scientific approaches available for the prediction problem. This paper present a Disease Influence Measure (DIM) based diabetic prediction algorithm. The detailed approach is discussed in section 3.

II. Materials And Methods

The methods of diabetic neuropathy and nephropathy prediction can be classified according to the methods and measures being used. Initially, different data mining algorithms can be used for the prediction. Prediction of diabetes mellitus type 2 has been performed with k means clustering and logistic regression algorithm [2]. The method has been validated for its performance using Pima Indian Diabetes dataset. The Waikato tool has been used for performance analysis. Similarly in [18], the combination of KNN and Naive Bayes algorithms is used for diabetic prediction. In [4], Mapping attributes are used to map information from categorical and numerical attribute values to items and Constraining association rules are used for disease predication.

A simple tool is created using decision tree anlaysis to predict type 2 diabetics is presented in [5], which uses anthropometric and laboratory measurements, routine demographic and clinical data.

In [28], an ensemble learning algorithm with SMOTE has been presented towards the prediction of diabetes. Various techniques have been used to perform diabetic prediction. The ensembles are generated using the attributes of data set, where each data point has 13 features or dimensions. The SMOTE technique has been used to handle the imbalance in data. The comparison with Decisive tree, regression analysis, nave Bayes and so on.

The pattern mining techniques are always has higher impact in any disease prediction. Such pattern mining algorithm for diabetic disease prediction is presented in [19]. The method uses the patterns to predict the possibility of diabetic induced diseases like heart issues, neuropathy, nephropathy, retinopathy and glaucoma. The method has been validated for its performance using thousands of diabetic records.

The data mining algorithms have great importance in any disease prediction. In [22], the methods like regression, SVM, GMM and ANN are evaluated for their performance in disease prediction. Life style based diabetic prediction is presented in [25]. The method considers the usual habits of persons like sleep, food, walking, MI using CART algorithm. Similarly in [3], the diabetic prediction has been performed with data mining techniques. The Type 2 diabetic mellitus has been predicted with improved K means algorithm. The evaluation is performed with PIMA data set. As an extension, a hybrid model for diabetic prediction is presented in [26]. The method uses K means clustering and decision tree algorithm for diabetic prediction. The method has been evaluated with the UCI data set.

A. Association Rule Mining

Association rule mining is one of the most vital and extensively researched techniques of data mining. The main aim is to extract interesting frequent patterns, correlations, casual structures or associations among set of items in the transaction databases.

An Association rules based diabetic prediction is presented in [6], which convert the numeric values into item sets and categorical forms. Using Apriori algorithm the method generates rules using PIMA data set. Based on the rules generated, the method predicts the possibility of diabetes. Similarly, in [1], a rule mining approach has been presented to generate diabetes pattern from lipid and Glucose data set. The method identifies the frequent items based on frequency measures and generates the rule using the minimum confidence value. Generated rules have been used to perform diabetic prediction.

B. Fuzzy Based Prediction

The prediction of diabetic neuropathy and nephropathy has suffer with the uncertainty of different factors. The fuzzy algorithm has the beauty of classifying the input sample under different range values. According to this, there are various methods discussed in this section. The prediction of diabetic neuropathy with uncertainty has been performed in [8]. The method generates fuzzy rules for different classes namely low, medium and higher. The rule has been used to perform diabetic prediction. Similarly towards the prediction of diabetic neuropathy an expert system has been presented in [9]. The method has been designed to support the medical practitioner in disease prediction.

Diabetic prediction with neural network has been presented in [10], which uses fuzzy logic in predicting diabetic. The prediction is performed in two stage. In first stage, guassian kernel function has been used to distribute the data. Second, neural network has been trained with data points and fuzzy logic has been used to perform diabetic prediction. Similarly, in [11], the diabetic prediction with NN and Fuzzy logic has been approached with case based approach. The method first predict the disease using NN, CBR, and fuzzy logic. Further the prediction is verified using rule based approach. In [12], the risk classification of diabetic nephropathy has been performed with fuzzy logic. The classification is performed with fuzzy logic and nephropathy has been predicted using fuzzy inference systems.

An fuzzy classifier for diabetic prediction with association rule has been presented in [7], The method generates association rules from huge data set. The rules are mined and generated with the help of data set. Generated rules are used to perform diabetic risk prediction. The classification performance has been improved.

C. SVM Based Prediction

The support vector machine is a machine learning algorithm which can be applied for variety of classification problems. Based on the SVM classifier the diabetic neuropathy and nephropathy can be classified in efficient way. Such methods are discussed in this section. SVM with Naïve Bayes diabetic prediction is presented in [20]. The combined model performs diabetic prediction using the data recorded from 400 plus patients. An decisive support system for diabetic prediction is presented in [21]. The method performs classification using AdaBoost algorithm and verified with the SVM classifier and decision tree algorithms. The combined approach of SVM, Decision Tree and Naive Bayes algorithm has been validated for its performance with PIDD data set.

The role of missing data has been considered for the prediction of diabetes with the help of machine learning algorithms. The missing data reduction has been approached with different feature selection techniques. Various feature selection and classification algorithms are used for performance measurement. An pure method of DPN prediction with multi category SVM has been presented in [13]. The method collects questionaries' from huge set of people and the same has been used to estimate NDS (neuropathy disease score) to classify the data point. In [24], the diabetic prediction with SVM, Naïve Bayes and Decision Tree has been analyzed for their performance.

An SVM with random forest ensemble approach for diabetic prediction is presented in [14]. The method ranks the genes according to the power and the same has been used to identify the gene influencing the disease.

D. Genetic Algorithm Based Prediction

The genetic algorithm is the best approach for the feature selection. The performance of any classifier is highly depending on the feature being considered. This section lists the methods of diabetic prediction based on GA. The feature selection plays important role in disease prediction. Towards improvement, a genetic algorithm based technique has been presented in [15] for the prediction of diabetic retinopathy. The binary vessel map has been generated using the segmentation techniques. The morphological features are extracted and classified using SVM. Here the genetic algorithm has been used for feature selection.

A Multifactorial Genetic Model has been presented for the prediction of diabetic peripheral nephropathy in [16]. However, the method uses the generic variants in feature selection to support disease prediction.

E. ANN Based Prediction

The neural networks are more efficient in identifying the hidden values to perform efficient classification. The methods based on ANN for diabetic prediction is discussed here. The problem of diabetic prediction has been approached with a combined approach of ANN and regression model in [17]. The method has been evaluated with various data sets and their performance in various parameters has been measured. Error ratio is reduced by this method.

In[23], an back propagation algorithm has been presented for the problem of diabetes mellitus prediction. The network has been designed with number of input and hidden layers. Finally a single output layer has been designed to produce the result. The author has designed a interface to ease the user access and has produced good results. Similarly, in [27], the author analyzed various decisive support systems for diabetes management.

In [29], the author presented an ANN based prediction algorithm using supervised learning algorithms. The data records of numerous patients have been collected from varying age group people. Based on the records, the method performed diabetic prediction.

III. Disease Influence Measure Based Diabetic Prediction

The proposed Disease influence measure based approach reads the input data set. For each data point, the method identifies the list of features from meta data. If any of the data features has been missed, then the data point has been eliminated. Second, the method estimates the Disease Influence Measure (DIM) based on the features considered. Based on the DIM value, the method predicts the diabetic. The detailed approach is discussed below:

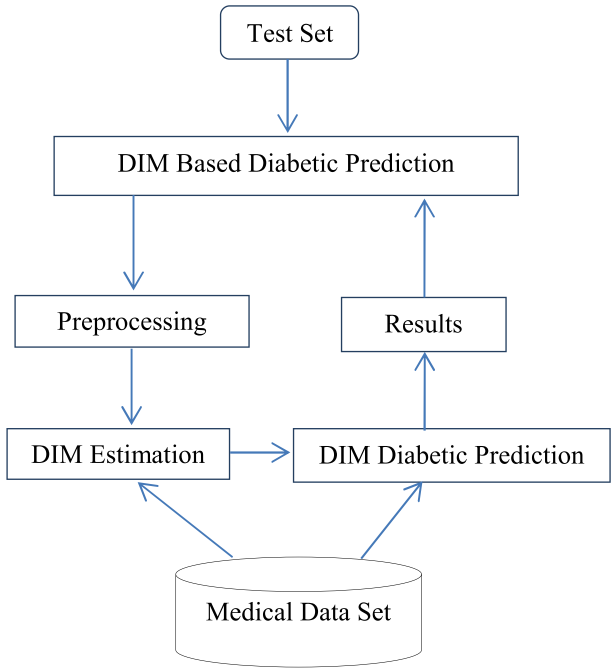

Fig. 1 Architecture of DIM Diabetic Prediction

Fig. 1 represent the architecture of the proposed diabetic prediction algorithm and shows various functional components. Each will be discussed in detail in this section.

Preprocessing

In this stage, the method reads the input test data set. For each data point, the method verifies the presence of all the features should be available. If any of the feature being missed, then it will be removed from the data set. The preprocessed data set has been used for diabetic prediction.

DIM Estimation

The Disease influence measure represent how the feature values of the input data point influences the upcoming disease. The method reads the input data point, and extract the features considered from data point. Then the DIM measure has been estimated to perform diabetic prediction. The value of DIM has been measured as follows:

$$\text{DIM} = \frac{\sum_{i=1}^{size(Fl)} Di(i) <> \sum Dk(i)}{Number\ of\ Dimension} \quad (1)$$

Here, Fl- Feature List, Di-input data point, Dk- data points in data set.

DIM Diabetic Prediction

The method reads the input data set which is given for testing The preprocessing is performed on each data point. The noise removed data points are estimated for DIM value. If the DIM value is higher than threshold, then it has been identified as Diabetic prone otherwise normal.

```
Pseudo Code of Diabetic Prediction:
Input: Test set Ts, Data Set Ds
Output: Null
Start
        Read input data set Ts, Ds
        For each data point Di from Ts
                D = Preprocessing(Di)
                DIM = Estimation DIM(D)
                If DIM>Th then
                        Diabetic
                Else
                        Normal
                End
        End
Stop
```

The above discussed algorithm reads the input data set and measure the DIM value. Based on DIM value, the method performs diabetic prediction.

IV. Results and Discussion

The classification and prediction performance of various methods have been evaluated and analyzed using the data set presented below. The data set is obtained from the UCI repository to merge and make another cooked up data set. In total it has number of features to be used for the classification and prediction purpose. The total number of records considered for prediction is 605 million.

TABLE 1
Evaluation Setup

Parameter	Value
Data Set	UCI + Cooked up
Number of features	19
Number of instances	605

The Table 1, shows the details of evaluation being used for the comparative study and performance analysis. The data set has 19 features and 605 instances of records have been used for the performance analysis. From the data set, we have measured the average values and plotted in the table 2.

TABLE 2
Details of data set.

Patient ID	Blood Glucose	Pre-breakfast blood glucose	Post-break fast blood glucose	HbA1c	Age	BMI	TriGly	BP
1	150	141	190	7	39	25.2	154	82
2	152	144	192	7.1	40	25.3	155	83
3	159	169	195	7.3	42	25.6	160	82
4	149	140	189	6.9	37	25.1	152	80
5	151	142	189	7.1	40	25.5	153	82
6	160	170	196	7.5	45	25.8	161	85
7	155	142	190	7	42	25.4	163	83
8	152	146	198	7	41	25.2	160	81
9	150	142	191	7	38	25.2	153	80
10	153	143	202	7.1	41	25.4	156	90
11	151	145	232	7	44	25.8	155	92
12	154	148	245	7.2	43	25.6	157	91
13	149	157	200	7	39	25.3	153	91
14	162	170	246	7.7	47	26	167	92
15	161	171	238	7	44	26.1	167	91
16	160	172	249	7	47	26	165	96
17	165	175	258	7.6	52	26.7	168	97
18	163	173	255	7.2	50	26.5	167	90
19	169	180	270	7.7	56	27	170	94
20	162	172	249	7.1	51	26.1	168	91
21	170	181	280	7.8	60	27.3	173	93
22	170	195	285	8	62	27.8	176	91

The Table 2, represent the data set being used for the prediction process and represent the features being used. The data set contains more than 650 rows of patient information. Each data point has 9 attributes which covers the maximum features of diabetes mellitus.

According to the data set values, each parameter has been considered important in the prediction of diabetes. The blood glucose has been monitored and if it crosses the value of 145 then it has been considered as diabetes. Similarly, the pre breakfast blood glucose crosses the value of 120 it can be considered as diabetic. Also, the post breakfast blood glucose crosses the value of 160 then it has been considered as diabetes. On the other side, if the HbA1c value is higher than 6.5 then it can be considered as diabetes. According to these constraints, the accuracy of the method has been measured as follows.

Prediction Accuracy

The accuracy of diabetic prediction has been measured based on the number of exact correct predictions over total number of predictions performed. It has been measured as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

False Classification Ratio

The false classification ratio is the measure being estimated based on the number of false classifications performed over total number of predictions. It has been measured as follows:

$$\text{False Ratio} = \frac{FP + FN}{TP + TN + FP + FN} \quad (2)$$

Time Complexity

The time complexity is the measure being estimated based on the time being taken for classification or prediction. It has been measured as follows:

Time complexity = Total Time Taken for Prediction.

The methods are applied with prediction process and their results has been recorded and compared. The analysis has been performed in various parameters of prediction. The method has produced the following results.

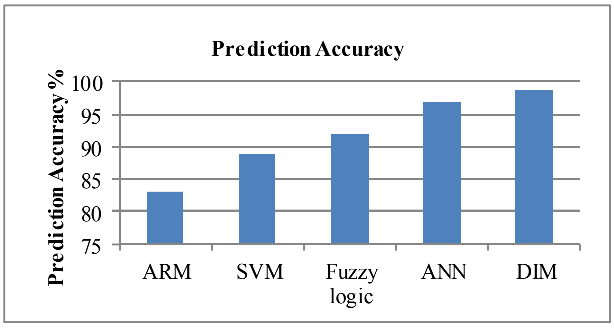

Fig. 2 Comparative performance in prediction accuracy

The performance on prediction accuracy has been measured and compared with different classification algorithms. The methods have produced the above mentioned results which are mentioned in Fig. 2. However, the Disease Influence Measure based approaches are higher efficient in prediction than other methods.

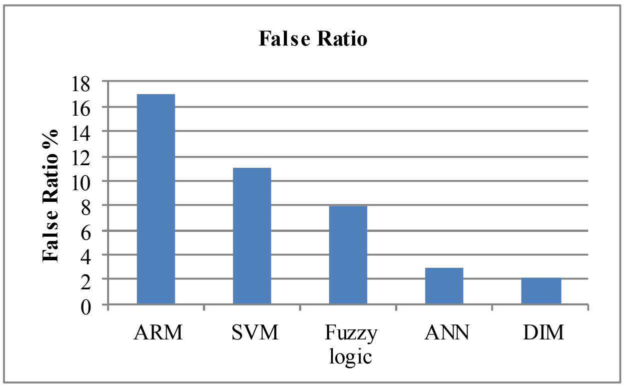

Fig. 3 Comparative result on false ratio

The false prediction ratio produced by different methods has been measured. The comparative result on false ratio has been presented in Fig. 3. The Disease Influence Measure based approaches has produced less false ratio than other methods.

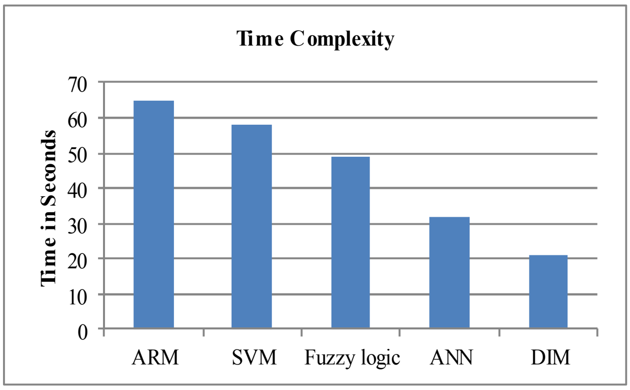

Fig. 4 Comparative result on time complexity

The performance on time complexity has been measured for different types of methods. The comparative results are presented in Fig. 4. The Disease Influence Measure based algorithm produces more effective results on time complexity.

V. Conclusion

In this paper, the problem of diabetes prediction has been studied and analysed. Various methods on the prediction of disease have been studied. To improve the performance an Disease Influence Measure based algorithm is presented. The method estimates DIM measure for each test sample and performs diabetic prediction using the value of DIM. The methods are verified for their performance in prediction using the data set considered. The data set has been obtained from UCI repository and merged with the data obtained from different medical organizations. In total 19 features has been considered and 605 medical records has been used for the evaluation process. The methods are measured for their performance in various prediction parameters. Obtained results are plotted in the graph. In future, the problem of prediction methods will be identified and novel approaches are designed for the improvement of the prediction problem.

References

[1] Azra Ramezankhani, Omid Pournik, Jamal Shahrabi, Fereidoun Azizi, and Farzad Hadaegh1, "An Application of Association Rule Mining to Extract Risk Pattern for Type 2 Diabetes Using Tehran Lipid and Glucose Study Database", Int J Endocrinol Metab, April 2015.

[2] HanWu, ShengqiYang, ZhangqinHuang, JianHe, XiaoyiWang, "Type 2 diabetes mellitus prediction model based on data mining", Elsevier, Informatics in Medicine, vol. 10, pp. 100-107.2018.

[3] Abbasi A, Peelen LM, Corpeleijn E, van der Schouw YT, Stolk RP, Spijkerman AM, et al., "Prediction models for risk of developing type 2 diabetes: systematic literature search and independent external validation study" BMJ, 2012.

[4] Ordonez C, Omiecinski E, de Braal L, Santana CA, Ezquerra N, Taboada JA, et al., editors, "Mining constrained association rules to predict heart disease", IEEE 13th International Conference on Data Mining, IEEE Computer Society,pp.. 433, 2001.

[5] Ramezankhani A, Pournik O, Shahrabi J, Khalili D, Azizi F, Hadaegh F, "Applying decision tree for identification of a low risk population for type 2 diabetes. Tehran Lipid and Glucose Study", Diabetes Res Clin Pract., September 2014.

[6] B. M. Patil, R. C. Joshi, Durga Toshniwal, "Association rule for classification of type -2 diabetic patients", 2010 Second International Conference on Machine Learning and Computing, IEEE, 2010.

[7] Dhivya Selvaraj, Mrs.Merlin Mercy, "Distributed association rule mining and summarization for Diabetes Mellitus and Its Co-Morbid Risk Prediction strategy using FUZZY Classifier", International Journal of Engineering and Applied Sciences (IJEAS) ISSN: 2394-3661, vol. 2, November 2015.

[8] Andreja P. PiconI; Neli R. S. OrtegaII; Ricky WatariI; Cristina SartorI; Isabel C. N. SaccoI, "Classification of the severity of diabetic neuropathy: a new approach taking uncertainties into account using fuzzy logic", Clinics(Sao Paulo) vol.67, Feburuay 2012.

[9] Meysam Rahmani Katigari, Haleh Ayatollahi, Mojtaba Malek, and Mehran Kamkar Haghighi, "Fuzzy expert system for diagnosing diabetic neuropathy", World J Diabetes,pp. 80–88, Feburary 2017.

[10] Tejashri N. Giri, "Prognosis of Diabetes using Neural Network, Fuzzy Logic, Gaussian Kernel Method", International Journal of Computer Applications (0975 – 8887) vol. 124, August 2015.

[11] MythiliThirugnanam, PraveenKumar Dr. S. VigneshSrivatsan C.R.Nerlesh, "Improving the Prediction Rate of Diabetes Diagnosis Using Fuzzy, Neural Network, Case Based (FNC) Approach", Elsevier, Procedia Engineering vol. 38, pp. 1709-1718, 2012.

[12] Narasimhan B, "Fuzzy logic system for risk-level classification of diabetic nephropathy", International Conference on Green Computing Communication and Electrical Engineering (ICGCCEE), March 2014.

[13] Kazemi M1, Moghimbeigi A, Kiani J, Mahjub H, Faradmal J, “Diabetic peripheral neuropathy class prediction by multicategory support vector machine model: a cross-sectional study”, Epidemiol Health, March 2016.

[14] AtulKumar, D.Jeya Sundara Sharmila, SachidanandSingh, SVMRFE based approach for prediction of most discriminatory gene target for type II diabetes”, Genomics Data, vol. 12, pp. 28-37, June 2017.

[15] Roshan A. Welikala, “Genetic Algorithm Based Feature Selection Combined with Dual Classification for the Automated Detection of Proliferative Diabetic Retinopathy”, Computerized medical imaging and graphics: the official journal of the Computerized Medical Imaging Society, March 2015.

[16] Blech I, Katzenellenbogen M, Katzenellenbogen A, Wainstein J, Rubinstein A, Harman-Boehm I, “Predicting Diabetic Nephropathy Using a Multifactorial Genetic Model”, PLoS ONE, April 2011.

[17] Parastoo RAHIML , Ahmad JAFARIAN, “Prediction of Diabetes by Using Artificial Neural Network, Logistic Regression Statistical Model and Combination of Them”, Bulletin de la Société Royale des Sciences de Liège, vol. 85, pp. 1148 – 1164, 2016.

[18] Deeraj Shett, Kishor Rit, Sohail Shaikh, Nikita Patil, “Diabetes disease prediction using data mining”, International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), March 2017.

[19] Saeed Farzi, Sahar Kianian, Ilnaz Rastkhadive, “Predicting serious diabetic complications using hidden pattern detection”, IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI), December 2017.

[20] Zhilbert Tafa, Nerxhivane Pervetica, Bertran Karahoda, “An intelligent system for diabetes prediction”, 2015 4th Mediterranean Conference on Embedded Computing (MECO), June 2015.

[21] V. Veena Vijayan, C. Anjali, “Prediction and diagnosis of diabetes mellitus — A machine learning approach”, IEEE Recent Advances in Intelligent Computational Systems (RAICS), 2016.

[22] Messan Komi, Jun Li, Yongxin Zhai, Xianguo Zhang, “Application of data mining methods in diabetes prediction”, 2nd International Conference on Image, Vision and Computing (ICIVC), April 2017.

[23] Sneha Joshi, Megha Borse, “Detection and Prediction of Diabetes Mellitus Using Back-Propagation Neural Network”, International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE),September 2016.

[24] V Veena Vijayan, C Anjali, “Decision support systems for predicting diabetes mellitus - A Review”, Global Conference on Communication Technologies (GCCT), 2015.

[25] Ayush Anand, Divya Shakti, “Prediction of diabetes based on personal lifestyle indicators”, 1st International Conference on Next Generation Computing Technologies (NGCT), 2015.

[26] Wenqian Chen, Shuyu Chen, Hancui Zhang, Tianshu W, “A hybrid prediction model for type 2 diabetes using K-means and decision tree”, 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), 2017.

[27] Ivan Contreras, “Artificial Intelligence for Diabetes Management and Decision Support: Literature Review”, J Med Internet Res., May 2018.

[28] Manal Alghamdi, Mouaz Al-Mallah, “Predicting diabetes mellitus using SMOTE and ensemble machine learning approach: The Henry Ford ExercIse Testing (FIT) project”, PLoS One. July 2017.

[29] Akmal Sapon, Muhammad & Ismail, Khadijah & Suehazlyn, Zainudin & Malaysia, Nasional & Lumpur, Kuala, “Prediction of Diabetes by using Artificial Neural Network”, 2018.