http://www.jctjournals.com ISSN(ONLINE):2278 – 3814

Volume-10, Issue-10, October 2021
JOURNAL OF COMPUTING TECHNOLOGIES (JCT)
International Journal
Page Number: 01-07

A Literature Survey On Different Diabetics Patient Analysis Using Machine Learning

Anupriya Hamilton [1], Prof. Sugan Patel [2],
[1]M.Tech Scholar, [2]Assistant Professor,
[1, 2]Department of Computer Science & Engineering
[1, 2]Lakshmi Narain College of Technology Excellence Bhopal, (M.P), India
[1]anupriya1094@gmail.com, [2]techsugan@gmail.com

***Abstract*— In this review paper present a different model for diabetes detection. In this survey paper discuss the different type of diabetes mellitus detection methods. Now a day's diabetes is one of the most dangerous deceases in this whole world. There are number of different methods available for the analysis of diabetes patients on the basis of patient physical properties such as weight, body mass index, blood pressure and other parameters. There are different classification analysis method available diabetes patient such J-45, support vector machine, multi support vector, navbais and neural network. In this survey paper discuss the different methods of clustering and classification of diabetes patients.**

Keywords — Diabetic Type 2, Cascaded Feed Forward Network, Linear Regression, Machine Learning Processes, Random Forent And Accuracy etc ….

I. INTRODUCTION

Diabetes mellitus could be a heterogeneous group of diseases specify by chronic elevation of glucose within the blood. It arises as a result of the body is unable to provide enough insulin for its own desires, either due to impaired endocrine secretion, impaired endocrine action, or both. polygenic disorder affects some three hundred million folks world-wide, and is on the rise. Chronic exposure to high glucose may be a leading explanation for kidney disease, visual loss and a variety of different kinds of tissue injury. sickness} additionally predisposes to blood vessel disease, not least as a result of it's typically in the middle of cardiovascular disease, super molecule disorders and blubber. several cases of polygenic disease and most of its unwanted semi permanent consequences ar probably avertible, however this may need intervention at a social moreover as at a medical level. This section of Diapedia offers associate introduction to the history of polygenic disorder, its clinical presentation, its current classification and its international medicine. we tend to conjointly introduce a number of the psychological and social aspects of polygenic disease, together with the 'hot topics' that dominate the media, and supply an summary of current areas of analysis interest. of these topics ar thought-about in bigger detail elsewhere in Diapedia, and that we hope you'll explore them additional.

Diabetes Mellitus ("diabetes" for short) may be a serious illness that happens once your body has issue properly regulation the number of dissolved sugar (glucose) in your blood stream. it's unrelated to a equally named disorder "Diabetes Insipidus" that involves kidney-related fluid retention issues. so as to grasp polygenic disease, it's necessary to initial perceive the role aldohexose plays with reference to the body, and what will happen once regulation of aldohexose fails and glucose levels become perilously low or high. The tissues and cells that structure the anatomy reside things, and need food to remain alive. The food cells eat could be a kind of sugar referred to as aldohexose. mounted in situ as they're, the body's cells ar utterly captivated with the blood stream during which they're bathed to bring aldohexose to them. while not access to adequate aldohexose, the body's cells don't have anything to fuel themselves with and shortly die. citizenry eat food, not aldohexose.

Human foods get regenerate into aldohexose as a vicinity of the conventional digestion method. Once reborn, aldohexose enters the blood stream, inflicting the extent of dissolved aldohexose within the blood to rise. The blood stream then carries the dissolved aldohexose to the assorted tissues and cells of the body. tho' aldohexose could also be out there within the blood, close cells aren't ready to access that aldohexose while not the help of a chemical internal secretion referred to as endocrine. internal secretion acts as

a key to open the cells, permitting them to receive and utilize on the market aldohexose. Cells absorb aldohexose from the blood within the presence of internal secretion, and blood glucose levels drop as sugar leaves the blood and enters the cells. internal secretion will be thought of as a bridge for aldohexose between the blood stream and cells. it's vital to grasp once levels of endocrine increase, levels of sugar within the blood decrease (because the sugar goes into the cells to be used for energy).The body is meant to control and buffer the number of aldohexose dissolved within the blood to take care of a gradual provide to satisfy cell wants.

The exocrine gland, one in all your body's several organs, produces, stores and releases endocrine into the blood stream to bring aldohexose levels backpedal. The concentration of aldohexose accessible within the blood stream at any given moment depends on the number and sort of foods that individuals eat. Refined carbohydrates, candy and sweets ar straightforward to interrupt down into aldohexose. Correspondingly, glucose levels rise speedily when such foods are consumed. In distinction, blood sugars rises step by step and slowly when consumption a lot of advanced, unrefined carbohydrates (oatmeal, apples, baked potatoes, etc.) that need a lot of biological process steps crop up before aldohexose will be yielded. baby-faced with speedily rising glucose concentrations, the body should react quickly by emotional giant amounts of endocrine all directly or risk a dangerous condition referred to as Hyper glycemias (high blood sugar) which is able to be represented below. The inflow of endocrine permits cells to utilize aldohexose, and aldohexose concentrations drop. whereas aldohexose levels will rise and fall quickly, endocrine levels modification rather more slowly.

A. Machine Learning

ANNs area unit composed of multiple nodes, that imitate biological neurons of human brain. The neurons area unit connected by links and that they act with one another. The nodes will take input file and perform easy operations on the info. The results of these operations is passed to different neurons. The output at every node is termed its activation or node worth. Each link is related to weight. ANNs area unit capable of learning, that takes place by sterilization weight values. the subsequent illustration shows an easy ANN−

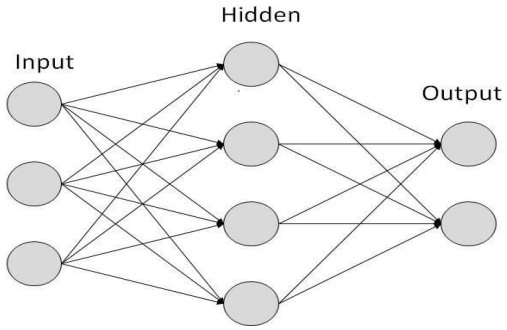

Fig 1. Types of Artificial Neural Networks

B. Machine Learning in ANNs

ANNs are capable of learning and that they have to be compelled to be trained. There ar many learning way

- **Supervised Learning** − It involves a tutor that's scholar than the ANN itself. for instance, the teacher feeds some example knowledge concerning that the teacher already is aware of the answers. For example, pattern recognizing. The ANN comes up with guesses whereas recognizing. Then the teacher provides the ANN with the answers. The network then compares it guesses with the teacher's "correct" answers and makes changes consistent with errors.

- **Unsupervised Learning** − it's needed once there's no example knowledge set with known answers. for instance, finding out a hidden pattern. during this case, cluster i.e. dividing a collection of components into teams in step with some unknown pattern is disbursed supported the prevailing} knowledge sets present.

- **Reinforcement Learning** − This strategy engineered on observation. The ANN makes a call by perceptive its atmosphere. If the observation is negative, the network adjusts its weights to be ready to build a unique needed call consequent time.

II LITERATURE SURVEY

Wu, Han, et al. "Type 2 diabetes mellitus prediction model based on data mining." This research work presented, Due to its continuously increasing occurrence, more and m78ore families are influenced by diabetes mellitus. Most diabetics know little about their health quality or the risk factors they face prior to diagnosis. In this study, we have presented a novel model based on data mining techniques for predicting type 2 diabetes mellitus (T2DM). The main problems that researcher are trying to solve are to improve the accuracy of the prediction model, and to make the model adaptive to more than one dataset. Based on a series of pre-processing procedures, the model is comprised of two parts, the improved K-means algorithm and the logistic regression algorithm. The Pima Indians Diabetes Dataset and the Waikato Environment for Knowledge Analysis toolkit were utilized to compare our results with the results from other researchers. The conclusion shows that the model attained a 3.04% higher accuracy of prediction than those of other researchers. Moreover, our model ensures that the dataset quality is sufficient. To further evaluate the performance of our model, we applied it to two other diabetes datasets. Both experiments' results show good performance. As a result, the model is shown to be useful for the realistic health management of diabetes**[01]**.

Ioannis Kavakiotis et. al, "Machine Learning and Data Mining Methods in Diabetes Research", [2017] This

research work presented, the growth of biotechnology and health sciences have led to a significant production of data, such as high throughput genetic data and clinical information, generated from large Electronic Health Records (EHRs). To this end, application of machine learning and data mining methods in bio sciences is presently, more than ever before, vital and indispensable in efforts to transform intelligently all available information into valuable knowledge. Diabetes mellitus (DM) is defined as a group of metabolic disorders exerting significant pressure on human health worldwide. Extensive research in all aspects of diabetes (diagnosis, etiopathophysiology, therapy, etc.) has led to the generation of huge amounts of data. The aim of the present study is to conduct a systematic review of the applications of machine learning, data mining techniques and tools in the field of diabetes research with respect to a) Prediction and Diagnosis, b) Diabetic Complications, c) Genetic Background and Environment, and e) Health Care and Management with the first category appearing to be the most popular. A wide range of machine learning algorithms were employed. In general, 85% of those used were characterized by supervised learning approaches and 15% by unsupervised ones, and more specifically, association rules. Support vector machines (SVM) arise as the most successful and widely used algorithm. Concerning the type of data, clinical datasets were mainly used. The title applications in the selected articles project the usefulness of extracting valuable knowledge leading to new hypotheses targeting deeper understanding and further investigation in DM **[02]**.

Phattharat Songthung et.al, "Improving Type 2 Diabetes Mellitus Risk Prediction Using Classification", [2016], In this research work presented Diabetes is a chronic disease that contributes to a significant portion of the healthcare expenditure for a nation as individuals with diabetes need continuous medical care. In order to prevent or delay the onset of type 2 diabetes, it is necessary to identify high risk populations and introduce behaviour modifications as early as possible. Screening the population to identify high risk individuals is an important task. One of the most accurate tests of diabetes is through the analysis of fasting blood sugar, but it is invasive and costly. Furthermore, it is only useful when the individual is already displaying symptoms i.e., making a diagnosis, which is considered too late to be an effective screening mechanism. Therefore, a reliable non-invasive inexpensive test to predict high risk individuals in advance is needed. In this research work researcher use classification to mine an extensive dataset gathered from 12 hospitals in Thailand during 2011-2012 with 22,094 records of screened population who are females age 15 years or older. We use Rapid Miner Studio 7.0 with Naive Bayes and CHAID (Chi-squared Automatic Interaction Detector) Decision Tree classifiers to predict high risk individuals and compared our results to existing hand-computed diabetes risk scoring mechanisms. We define the goal of risk prediction as coverage which is the ability to use screening data to identify individuals that are eventually diagnosed with diabetes. Our results indicate that the use of classification introduced in this research work instead of hand-computed scoring can improve the prediction performance with an increase in coverage **[03]**.

Madhuri Panwar et. al, "K-nearest neighbour based methodology for accurate diagnosis of diabetes mellitus", [2016], In this research work presented, Diabetes is one of the leading causes of death, disability and economic loss throughout the world. Type 2 diabetes is more common (90-95% worldwide) type of diabetes. However, it can be prevented or delayed by taking the right care and interventions which indeed an early diagnosis. There has been much advancement in the field of various machine learning algorithms specifically for medical diagnosis. But due to partially complete medical data sets, accuracy often decreases, results in more number of misclassification that can lead t o harmful complications. An accurate prediction and diagnosis of a disease becomes a challenging research problem for many researchers. Therefore, aimed to improve the diagnosis accuracy researcher have presented a new methodology, based on novel pre-processing techniques, and K-nearest neighbour classifier. The effectiveness of the presented methodology is validated with the help of various quantitative metrics and a comparative analysis, with previously reported studies using the same UCI dataset focusing on pima-diabetes disease diagnosis. This is the first work of its kind, where 100% classification accuracy is achieved by feature reduction from eight to two that shows the out performance of the presented methodology over existing methods **[04].**

Purush ottam et.al, "Diabetes Mellitus Prediction System Evaluation Using C4.5 Rules and Partial Tree" [2015], In this research work presented Diabetes disease prediction is a progressive area of research in the healthcare Sector. Although many data mining techniques have been employed to assess the main causes of diabetes, but only few sets of clinical risk factors are considered. Due to this, some important factors like pre-diabetes health conditions are not considered in their analysis. So the results produced by such techniques may not represent appropriate diabetes pattern and risk factors appropriately. In this study, researcher have designed a system that can efficiently discover the rules to predict the risk level of patients based on the given parameter about their health. Then we evaluate and compare this system using C45 rules and partial tree. The performance of the system is evaluated in terms of different parameter like rules generated, classification accuracy, classification error, global classification error and the experimental results shows that the system has great potential in predicting the Diabetes disease up to 81.27 correctly**[05].**

Kiarash Zahirnia et. al, "Diagnosis of type 2 diabetes using cost-sensitive learning", [2015], This research work discussed, Diabetes is the fourth cause of death in the world and has some complications such as amputation, visual impairment, kidney disorder and early death. 80% of diabetes symptoms is avoidable by early diagnosis. There are standard methods to diagnose diabetes by measuring the plasma glucose concentration. However, screening all people is impossible due to financial shortages especially in developing countries. Therefore, it is presented that the people more than 20 years old who are prone to diabetes be tested. Identifying diabetic people is possible by using different methods including machine learning algorithms. Standard machines assume balance in data and use all the available and related features to achieve lower error rates. However, in medical applications misclassification cost should be minimized as misclassification costs for healthy and patient instances are different. In addition, we are facing imbalanced data in the most medical issues including diabetes diagnosis. The features will also have different knowledge levels and costs. As a result, cost-sensitive methods should be applied in such situations. In this research work researcher present and compare different cost-sensitive learning methods for diagnosis of type 2 diabetes. For evaluation of the methods two different data sets are used, which one of them is data set of Tabriz, Iran and the other one is for Pima Indian data set. By defining different scenarios, researcher will show that none of the studied methods has absolute superiority over the other methods and the performance of algorithms will be different due to used data set and defined scenario**[06]**.

Bum J. L. et. al, "Prediction of Fasting Plasma Glucose Status Using Anthropometric Measures for Diagnosing Type 2 Diabetes", [2014], This research work discussed, It is well known that body fat distribution and obesity are important risk factors for type 2 diabetes. Prediction of type 2 diabetes using a combination of anthropometric measures remains a controversial issue. This study aims to predict the fasting plasma glucose (FPG) status that is used in the diagnosis of type 2 diabetes by a combination of various measures among Korean adults. A total of 4870 subjects (2955 females and 1915 males) participated in this study. Based on 37 anthropometric measures, we compared predictions of FPG status using individual versus combined measures using two machine-learning algorithms. The values of the area under the receiver operating characteristic curve in the predictions by logistic regression and naive Bayes classifier based on the combination of measures were 0.741 and 0.739 in females, respectively, and were 0.687 and 0.686 in males, respectively. Our results indicate that prediction of FPG status using a combination of anthropometric measures was superior to individual measures alone in both females andmales. Researcher show that using balanced data of normal and high FPG groups can improve the prediction and reduce the intrinsic bias of the model toward the majority class **[07].**

Soo Beom C. et.al, "Screening for Prediabetes Using Machine Learning Models", [2014], In this research work presented, The global prevalence of diabetes is rapidly increasing. Studies support the necessity of screening and interventions for pre-diabetes, which could result in serious complications and diabetes. This study aimed at developing an intelligence-based screening model for pre-diabetes. Data from the Korean National Health and Nutrition Examination Survey (KNHANES) were used, excluding subjects with diabetes. The KNHANES 2010 data were used for training and internal validation, while data from KNHANES 2011 were used for external validation. Researcher developed two models to screen for prediabetes using an artificial neural network (ANN) and support vector machine (SVM) and performed a systematic evaluation of the models using internal and external validation. Researcher compared the performance of our models with that of a screening score model based on logistic regression analysis for prediabetes that had been developed previously. The SVM model showed the areas under the curve of 0.731 in the external datasets, which is higher than those of the ANN model (0.729) and the screening score model (0.712), respectively. The pre-screening methods developed in this study performed better than the screening score model that had been developed previously and may be more effective method for prediabetes screening.**[08].**

III. NEURAL NETWORK

Neural network can be a study model, that imitates the animal's neural network behaviors. This model depends on the quality of the system to comprehend the aim of method information by adjusting the affiliation between the inner node. to keep with the connections' vogue, the neural network model are divided into forward network and feedback network. throughout this research work, we tend to tend to used the Neural Pattern Recognition app in MATLAB, that would be a two-layer-feed-back network with sigmoid hidden and soft ax output neurons. The neural network structural is shown in Figure 2.

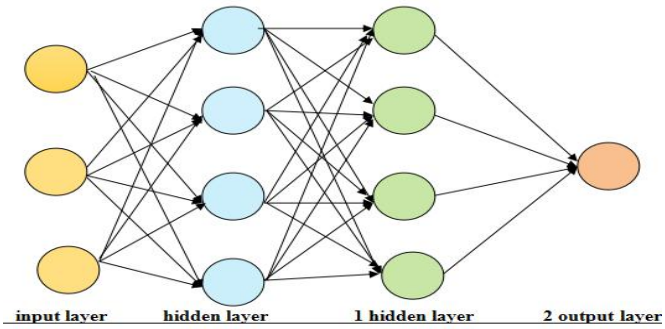

Fig. 2. Neural network

The structural of two–layer-feed-back network in MATLAB. This figure is from MATLAB, which might describe this network regulation ideally. Where, W is illustration the burden and b is that the bias variable. In neural network, there are some important parts, notably input layer, hidden layer and output layer. The input layer is to blame for accretive computer file. we are going to get the results from the output layer. The layer between the input layer and conjointly the output layer is called hidden layer. as a results of they are invisible to the skin. there is not any association between neurons on identical layer. throughout this network, the number of hidden layers set to 10, which could get a way higher performance.

IV. TYPES OF NEURAL NETWORKS

Neural networks are process models that work just like the functioning of a person's system. There are many types of artificial neural networks.

A. Feed Forward Neural Network – Artificial Neuron

This neural network is one amongst the only sorts of ANN, wherever the info or the input travels in one direction. The info passes through the input nodes and exit on the output nodes. This neural network could or might not have the hidden layers. In straightforward words, it's a front propagated wave and no back propagation by employing a classifying activation operate sometimes. Below may be a Single layer feed forward network. Here, the add of the product of inputs and weights square measure calculated and fed to the output. The output is taken into account if it's on top of an exact worth i.e. threshold (usually 0) and therefore the nerve cell fireplaces with an activated output (usually 1) and if it doesn't fire, the deactivated price is emitted (usually -1).

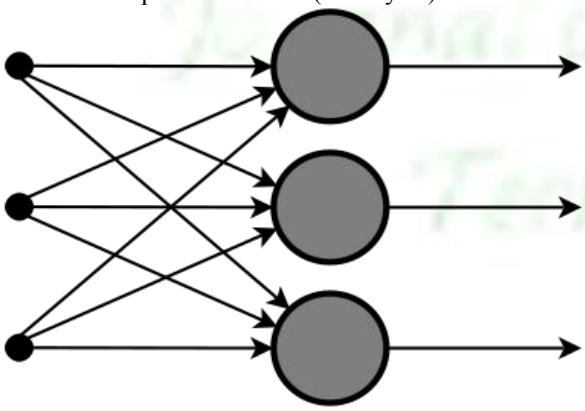

Fig. 3 Single layer feed forward network

B. Radial Basis Function Neural Network:

Radial basic functions contemplate the gap of a degree with relevance the middle. RBF functions have 2 layers, 1st wherever the options are combined with the Radial Basis operate within the inner layer and so the output of those options are taken into thought whereas computing constant output within the next time-step that is largely a memory. Here, the space live utilized in Euclidian, different distance measures can even be used. The model depends on the utmost reach or the radius of the circle in classifying the points into totally different classes. If the purpose is in or round the radius, the probability of the new purpose begin category field into that class is high. There will be a transition whereas dynamic from one region to a different and this will be controlled by the beta perform.

Fig. 4 Power Restoration Systems

This neural network has been applied in Power Restoration Systems. Power systems have inflated in size and quality. Each factors increase the chance of state outages. When a blackout, power must be restored as quickly and faithfully as attainable.

C. Convolution Neural Network:

Convolution neural networks are kind of like feed forward neural networks, wherever the neurons have learn-able weights and biases. Its application has been in signal and image process that takes over Open CV in field of laptop vision.

Below could be a illustration of a Convent, during this neural network, the input options are taken in batch wise sort of a filter. This may facilitate the network to recollect the photographs in elements and might cipher the operations. These computations involve conversion of the image from RGB or HSI scale to Gray-scale. Once we have this, the changes within the element worth can facilitate detection the sides and pictures will be classified into totally different classes.

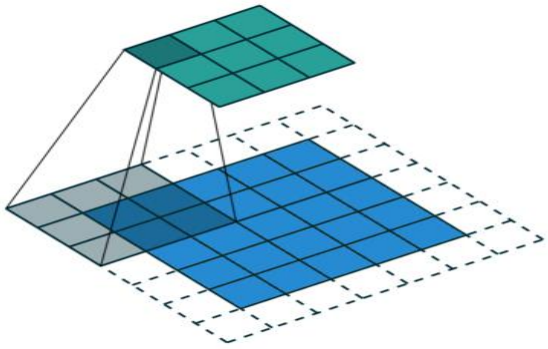

Fig. 4 Convolution neural networks

V.CONCLUSION

In this survey paper discuss on the research work present a hybrid model that is combination of linear regression model and cascaded feed forward network with back propagation training. The detection of the diabetic type 2 patient shows better accuracy as compare to other previous methods. The outcome proposed method shown that can be detected with an awfully high accuracy, up to 96.54 % in the different diabetic patient datasets. The proposed method also shows lower complexity and better timing due to regression model soft threshold values.

REFERENCE

[1] Wu, Han, et al. "Type 2 diabetes mellitus prediction model based on data mining." Informatics in Medicine Unlocked 10 (2018).

[2] Kavakiotis, Ioannis, Olga Tsave, AthanasiosSalifoglou, NicosMaglaveras, IoannisVlahavas, and IoannaChouvarda. "Machine learning and data mining methods in diabetes research." Computational and structural biotechnology journal (2017).

[3] Songthung,Phattharat, and KunwadeeSripanidkulchai. "Improving type 2 diabetes mellitus risk prediction using classification." In Computer Science and Software Engineering (JCSSE), 2016 13th International Joint Conference on, pp. 1-6. IEEE, 2016.

[4] Panwar, Madhuri, Amit Acharyya, Rishad A. Shafik, and Dwaipayan Biswas. "K-nearest neighbor based methodology for accurate diagnosis of diabetes mellitus." In Embedded Computing and System Design (ISED), 2016 Sixth International Symposium on, pp. 132-136. IEEE, 2016.

[5] Saxena, Kanak, and Richa Sharma. "Diabetes mellitus prediction system evaluation using c4. 5 rules and partial tree." In Reliability, Infocom Technologies and Optimization (ICRITO)(Trends and Future Directions), 2015 4th International Conference on, pp. 1-6. IEEE, 2015.

[6] ZahirniaKiarash, Mehdi Teimouri, RohallahRahmani, and Amin Salaq. "Diagnosis of type 2 diabetes using cost-sensitive learning." In Computer and Knowledge Engineering (ICCKE), 2015 5th International Conference on, pp. 158-163. IEEE, 2015.

[7] Lee, Bum Ju, Boncho Ku, Jiho Nam, Duong Duc Pham, and Jong Yeol Kim. "Prediction of fasting plasma glucose status using anthropometric measures for diagnosing type 2 diabetes." IEEE journal of biomedical and health informatics 18, no. 2 (2014).

[8] Choi, SooBeom, Won Jae Kim, Tae KeunYoo, JeeSoo Park, Jai Won Chung, Yong-ho Lee, EunSeok Kang, and Deok Won Kim. "Screening for prediabetes using machine learning models." Computational and mathematical methods in medicine 2014 (2014).

[9] Velu, C. M., and K. R. Kashwan. "Visual data mining techniques for classification of diabetic patients." In Advance Computing Conference (IACC), 2013 IEEE 3rd International, pp. 1070-1075. IEEE, 2013.

[10] Dalakleidi, Kalliopi V., KonstantiaZarkogianni, Vassilios G. Karamanos, Anastasia C. Thanopoulou, and Konstantina S. Nikita. "A hybrid genetic algorithm for the selection of the critical features for risk prediction of cardiovascular complications in Type 2 Diabetes patients." In Bioinformatics and Bioengineering (BIBE), 2013 IEEE 13th International Conference on, pp. 1-4. IEEE, 2013.

[11] OlokobaAbdulfatai B., Olusegun A. Obateru, and Lateefat B. Olokoba. "Type 2 diabetes mellitus: a review of current trends." Oman medical journal 27, no. 4 (2012).

[12] Marinov, Miroslav, Abu Saleh Mohammad Mosa, IllhoiYoo, and Suzanne Austin Boren. "Data-mining technologies for diabetes: a systematic review." Journal of diabetes science and technology 5, no. 6 (2011).

[13] Sacks, David B., Mark Arnold, George L. Bakris, David E. Bruns, Andrea Rita Horvath, M. Sue Kirkman, AkeLernmark, Boyd E. Metzger, and David M. Nathan. "Guidelines and recommendations for laboratory analysis in the diagnosis and management of diabetes mellitus." Clinical chemistry 57, (2011).

[14] Zhang, Ping, Xinzhi Zhang, Jonathan Brown, DorteVistisen, Richard Sicree, Jonathan Shaw, and Gregory Nichols. "Global healthcare expenditure on diabetes for 2010 and 2030." Diabetes research and clinical practice 87, no. 3 (2010): .

[15] B. M. Patil, R. C. Joshi and DurgaToshniwal, “Association rule for classification of type -2 diabetic patients”, Second International Conference on Machine Learning and Computing, 2010.

[16] Rodbard, Helena W., Andrew J. Green, Kathleen M. Fox, and Susan Grandy. "Trends in method of diagnosis of type 2 diabetes mellitus: results from SHIELD." International journal of endocrinology 2009 (2009).

[17] RajeebDey , VaibhavBajpai, Gagan Gandhi and BarnaliDey, “Application of Artificial Neural Network (ANN) technique for Diagnosing Diabetes Mellitus”, IEEE Region 10 and the Third international Conference on Industrial and Information Systems,Dec 2008.

[18] BalakrishnanSarojini, RamarajNarayanaswamy, Nickolas Savarimuthu, and Rita Samikannu. "SVM ranking with backward search for feature selection in type II diabetes databases." In Systems, Man and Cybernetics, 2008. SMC 2008. IEEE International Conference on, pp. 2628-2633. IEEE, 2008.

[19] Liang Lily, VinayMandal, Yi Lu, and Deepak Kumar. "Multi-dimensional cluster misclassification

test for pathway differential analysis of diabetes." In Computer and Computational Sciences, 2007. IMSCCS 2007. Second International Multi-Symposiums on, pp. 84-91. IEEE, 2007.

[20]]Aekplakorn, Wichai, PongamornBunnag, Mark Woodward, PiyamitrSritara, SayanCheepudomwit, SukitYamwong, Tada Yipintsoi, and RajataRajatanavin. "A risk score for predicting incident diabetes in the Thai population." Diabetes care 29, no. 8 (2006).

[21] Lindström, Jaana, and JaakkoTuomilehto. "The diabetes risk score." Diabetes care 26, no. 3 (2003): 725-731.

[22] Tuomilehto, Jaakko, JaanaLindström, Johan G. Eriksson, Timo T. Valle, Helena Hämäläinen, PirjoIlanne-Parikka, SirkkaKeinänen-Kiukaanniemi et al. "Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance." New England Journal of Medicine344, no. 18 (2001): 1343-1350.

[23] Pan, Xiao-Ren, Guang-Wei Li, Ying-Hua Hu, Ji-Xing Wang, Wen-Ying Yang, Zuo-Xin An, Ze-Xi Hu et al. "Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance: the Da Qing IGT and Diabetes Study." Diabetes care 20, no. 4 (1997): 537-544.

[24] G.V. Kass, “An Exploratory Technique for Investigating Large Quantities of Categorical Data”, journal, Wiley for the Royal Statistical Society, 1980.

