A Case Study on Soft Computing Techniques Used for Diabetes Mellitus

Saliman Rani*
Department of Computer Science Engineering
Guru Jambheshwar University of Science &
Technology, Hisar
Haryana, India

Dharminder Kumar
Chairperson of CSE Deptt.
Guru Jambheshwar University of Science &
Technology, Hisar
Haryana, India

Abstract—Diabetes is a major health problem in world. And it increase the risk of many other disease like- kidney disease, blindness, nerve damage, blood vessel damage and it contributes to heart disease. Several studies have reported that up to 50% of all patients with diabetes are undiagnosed. Although these patients are not aware of their disease, they are at risk for the micro and macro vascular complications. Early treatment of diabetes and associated cardiovascular risk factors may reduce the occurrence of these complications. Therefore early detection of diabetes will be of importance in reducing the burden of complications of diabetes. The early detection and prediction can give a warning at a stage where some medications and preventive action will help the patient to increase the span of his healthy life. Using various Soft Computing Techniques for medical diagnosis of diseases has recently become widespread. These techniques help physicians as a diagnosis assistant. Classification models are one of the most widely used groups of data mining tools that greatly help physicians to improve their prognosis, diagnosis or treatment planning procedures. Classification accuracy is one of the most important features in order to choose the appropriate classification model; hence, the researches directed at improving upon the effectiveness of these models have never stopped. A design of classifier for the detection of Diabetes Mellitus with optimal cost and precise performance is the need of the age.

Keywords— Diabetes Mellitus; Soft Computing; Classification; Performance Analysis.

I. INTRODUCTION

Diabetes Mellitus is a group of metabolic diseases in which a person has high blood sugar, either because the pancreas does not produce enough insulin, or because cells don't respond to the insulin that is produced. Due to high blood sugar this produces the classical symptoms of polyuria (frequent urination), polydipsia (increased thirst) and polyphagia (increased hunger). There are 3 major types of Diabetes Mellitus, "Type I Diabetes Mellitus", in which body does not produce insulin.This form was previously referred to as"insulin-dependent diabetes mellitus". The second type of Diabetes Mellitus is called "Type II Diabetes Mellitus" which results from insulin resistance, a condition in which cells fail to use insulin properly, sometimes combined with an absolute insulin deficiency. This type also named as "Non insulin-dependent diabetes mellitus" (NIDDM) or "adult-onset diabetes". Finally, "gestational diabetes" occurs when pregnant women without a previous diagnosis of diabetes develop a high blood glucose level, It may precede development of type I DM. Other forms of DM include congenital diabetes, which is due to genetic defects of insulin secretion, cystic fibrosis-related diabetes, steroid diabetes induced by high doses of glucocorticoids, and several forms of monogenic diabetes [1], [2], [3]. All types of Diabetes Mellitus have something in common. Normally, your body breaks down the sugars and carbohydrates you eat into a special sugar called glucose. Glucose fuels the cells in your body. But the cells need insulin, a hormone, in your bloodstream in order to take in the glucose and use it for energy. All types of Diabetes Mellitus have been treatable since insulin became available in 1921. Both type I & II are chronic conditions that cannot be cured. pancreas transplants have been tried with limited success in type I Diabetes Mellitus, gastric bypass surgery has been successful in many with morbid obesity and type II Diabetes Mellitus. Gestational DM usually resolves after delivery [5]. Untreated Diabetes Mellitus can cause many complications. Acute complications include diabetic ketoacidosis and non ketotic hyperosmolarcoma. Series long term complications include cardiocascular disease, chronic renal failure, and diabetic retinopathy. Adequate treatment of the disease is very important, as well as blood pressure control and lifestyle factors such as stopping smoking and maintaining a healthy body weight. Since the cells can't take in the glucose, it builds up in your blood. High levels of blood glucose can damage the tiny blood vessels in your kidneys, heart, eyes or nervous system. that's why diabetes can eventually cause heart disease, stroke, kidney disease, blindness and nerve damage to nerves in the feet (especially if left untreated) [4], [6].

Now days Classification systems have been widely utilized in medical domain to explore patient's data and extract a predictive model. They help doctors to improve their treatment procedures. In recent years, many studies have been performed in the diagnosis of diabetic disease literature.

II.   SOFT COMPUTING TECHNIQUE USED IN DIABETIC MELLITUS

Soft computing method is very well suited to analyze medical data so as to find out the risks strategy.
Soft computing technique such as fuzzy logic, ANN, SVM and other have been proposed and successfully employed for developing intelligent systems.

Soft computing differs from conventional (hard) computing in that, unlike hard computing, it is tolerant of imprecision, uncertainty, partial truth and approximation. In effect, the role model for soft computing is the human mind. The guiding principle of soft computing is: "Exploit the tolerance for imprecision, uncertainty, partial truth, and approximation to achieve tractability, robustness and low solution cost". The clinical data may consists of missing , incorrect and sometimes incomplete values set so using soft computing is the better alternative to handle such data. The principal constituents of soft computing are fuzzy logic, neural computing, evolutionary computation and probabilistic reasoning. The principal constituent methodologies in soft computing are complementary rather than competitive. Fuzzy logic handles imprecision, neural computing deals with learning, evolutionary computation is for optimization and probabilistic reasoning handles uncertainty.

In this part, an attempt is made to compile the work done by various scientists and researchers on predictive models applied to clinical data; [7] Developed temporal abstraction using inference graph for knowledge based temporal abstraction framework. Inference graph are a form of knowledge base which consist of transition rules and inference rules. [8] The back-propagation (BP) algorithm (Rumelhart, Hinton, & Williams, 1986) is widely recognized as a powerful tool for training of the MLNNs. [9] presented a clinical decision support system based on On-Line Analytical Processing and data mining. It discovers hidden patterns in the data. [10] describe Expert systems and different artificial intelligence techniques for classification systems in medical diagnosis is increasing gradually. As for other clinical diagnosis problems, classification systems have been used for diabetes diagnosis problem (Polat & Gunes, 2007). A modular clinical decision support system is discussed in [11]. [12] This paper presents a comparative study on Pima Indian diabetes disease diagnostic by using multilayer neural network which was trained by LM algorithm and probabilistic neural network. [13] Used Bayesian analysis to make meaningful diagnostic and treatment decisions in an evidence-based validated framework. [14] MI-MCS-FWSVM method obtains 93.58% accuracy on UCI dataset. [15] Presented a new expert system for diabetes disease diagnosis. They used modified spline smooth support vector machine. [16] describe feature subset selection problem. It is a task of identifying and selecting a useful subset of pattern-representing features from a larger set of features. [17] describe detection of diabetes using genetic programming. In [18] a clinical decision support system for metabolism synthesis is developed using a group of C language integrated production system. [19] describe general regression neural network (GRNN) for diabetes mellitus. [20] proposed decision support system for breast cancer. They have used soft computing approach. [21] used Fuzzy logic for representation of clinical guidelines for automated decision support. [22] describe adaptive neuro-fuzzy inference system (ANFIS). The main objective of this study is classification of diabetic diseases based on the symptoms. The dataset for the training of an ANFIS is collected from the various physicians for the classification of diabetic diseases. [23] describe Classification Of Diabetes Disease Using Support Vector Machine. [24] used recurrent artificial neural network to predict blood glucose sugar and then used neuro-fuzzy expert system to offer short-term therapeutic adviceIn reference. In [25] the author has proposed a neural network based predictive control structure and applied to treat the problem of diabetes management. The control approach is to predict future plant behavior; hence it specifies accurate control actions necessary to stabilize slow process systems such as physiological systems. [26] Suggested a novel method of Thyroid disease diagnosis by using fuzzy cognitive map based decision support system. The system is structured by the combined knowledge from experts and generated fuzzy rules from data. A multi agent system for Clinical Decision Support is presented in [27]. It uses supervised learning techniques for supporting clinical decisions. [28] Presented a data mining method on the study of medical information. They used improved apriori algorithm. [29] Proposes intelligible support vector machines for diagnosis of diabetes mellitus. Recently After reviewing the literature cited above it is identified that using the hybrid of soft computing and data mining techniques for preprocessing of clinical data and prediction can result in an effective prediction model.

III.   PERFORMANCE ANALYSIS

The performance of the proposed predictive model is measured using the following criteria.
Accuracy: The accuracy of a predictive model is the probability of it correctly classifying records in the test dataset.

Accuracy = (TP+TN) / (TP+TN+FP+FN)
Where TP is the number of true positive
      TN is the number of true negative
      FP is the number of false positive
      FN is the number of false negative

Area under receiver-operator characteristics curve (ROC): It characterizes the relationship between sensitivity and (1-specificity). The sensitivity of a test is the probability (0-100%) that a test is positive for the patients with diabetes. The specificity is the probability that a test is negative for patients without diabetes. The curve that has a larger area under curve is better than the one that has a smaller area under curve.

IV. CONCLUSIONS

Diabetes, Breast cancer and Heart Disease are the leading causes of death worldwide and the early prediction of these diseases are very important task. The computer aided disease prediction system helps the physician as a tool for disease diagnosis. From the analysis it is concluded that, soft computing intelligence techniques plays a major role in disease classification and prediction. Neural Network and Bayesian Classifier are widely used today for their great accuracy, coverage area and less time requirement. The classification accuracy can be improved by reduction in features. In spite of lots of others hybridized intelligence techniques, expert system, optimization techniques are under research to improve the accuracy, these techniques are most commonly used for classification with acceptable accuracy for medical diagnosis and various algorithms are under research which will improve the efficiency of these approaches.

REFERENCES

[1]   Centers for disease control and prevention, national diabetes, fact sheet, www.cdc.gov/diabetes (2011).

[2]   World health organization, diabetes center fact sheet n312,www.who.int/mediacentre/factsheets/fs312/en (2011).

[3]   Kurt George Matthew Mayer Alberti and PZ Zimmet. Defini-tion, diagnosis and classification of diabetes mellitus and its complications. part 1: diagnosis and classification of diabetes mellitus. provisional report of a who consultation. Diabetic medicine, 15(7):539–553, 1998.

[4]   SETTING IN DIABETES MELLITUS CARE. Guidelines for improving the care of the older person with diabetes mellitus. Hypertension, 16:2, 2003.

[5]   A Keech, RJ Simes, P Barter, J Best, R Scott, Marja-Riitta Taskinen, P Forder, A Pillai, T Davis, P Glasziou, et al. Effects of long-term fenofibrate therapy on cardiovascular events in 9795 people with type 2 diabetes mellitus (the field study): randomised controlled trial.Lancet, 366(9500):1849– 1861, 2005.

[6]   Jaakko Tuomilehto, Jaana Lindstr¨om, Johan G Eriks-son, Timo T Valle, Helena H¨am¨al¨ainen, Pirjo Ilanne-Parikka, Sirkka Kein¨anen-Kiukaanniemi, Mauri Laakso, Anne Louheranta, Merja Rastas, et al. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. New England Journal of Medicine, 344(18):1343–1350, 2001.

[7]   Pham Van Chung and Duong Tuan Anh,"Applying Temporal Abstraction in Clinical Databases", *International Conference on Research,Innovation and Vision for the Future*, 2007 IEEE, pp.192-199, 5-9 March 2007. doi: 10.1109/RIVF.2007.36915

[8]   Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart & J. L. McClelland (Eds.). Parallel distributed processing: Explorations in the microstructure of cognition (Vol. 1, pp. 318–362). Cambridge, MA: MIT Press.

[9]   Chua Sook Ling and Sellappan Palaniappan,"Clinical Decision Support System Using OLAP Mining", *in the International Journal of computer Science and Network Security*, Vol.8 No.9, pp. 290-296, September 2008

[10]   Polat, K., Gunes, S., & Aslan, A. (2008). A cascade learning system for classification of diabetes disease: Generalized discriminant analysis and

         least square support vector machine. Expert Systems with Applications, 34(1), 214–221.

[11]   Fran Wu, Mitch Williams, Peter Kazanzides, Ken Brady and Jim Fackler,"A modular Clinical Decision Support System Clinical prototype extensible into multiple clinical settings," *3rd International Conference on Pervasive Computing Technologies for Healthcare*,2009. PervasiveHealth 2009., pp.1-4, 1-3 April 2009 doi:10.4108/ICST.PERVASIVEHEALTH2009.6078

[12]   Hasan Temurtas, Nejat Yumusak , Nejat Yumusak , Feyzullah Temurtas, "A comparative study on diabetes disease diagnosis using neural networks", Conference on Expert Systems with Applications ,2009 Elsevier Ltd. doi:10.1016/j.eswa.2008.10.032

[13]   Howard Robin., John. S. Eberhardt III., W. D. Muller., R. Clark.,J. Kam.,"Classification of Pathology Data Using a Probabilistic (Bayesian) Model", *in proceedings of 18th International Conference on Systems Engineering,* 2005.,pp. 286- 291, 16-18 Aug.2005. doi: 10.1109/ICSENG.2005.22

[14]   Davar Giveki, Hamid Salimi, GholamReza Bahmanyar, Younes Khademian, "Automatic Detection of Diabetes Diagnosis using Feature Weighted Support Vector Machines based on Mutual Information and Modified Cuckoo Search".

[15]   Shanti Wulan Purnami, Jasni Mohamad and A. Emgbong, "A New Expert System for Diabetes Disease Diagnosis using Modified Spline Smooth Support Vector Machine", *International Workshop on biomathematics ,bioinformatics and biostatistic,* 23-26 March 2010 Fukuoka, Japan

[16]   E.P.Ephzibah, "COST EFFECTIVE APPROACH ON FEATURE SELECTION USING GENETIC ALGORITHMS AND FUZZY LOGIC FOR DIABETES DIAGNOSIS", International Journal on Soft Computing (IJSC), Vol.2, No.1, February 2011.

[17]   Muhammad Waqar Aslam and Asoke Kumar Nandi, "DETECTION OF DIABETES USING GENETIC PROGRAMMING", 18th European Signal Processing Conference (EUSIPCO-2010) Aalborg, Denmark, August 23-27, 2010 © EURASIP, 2010 ISSN 2076-1465.

[18]   Qunyi Zhou, "A Clinical Decision Support System for Metabolism Synthesis."*International Conference on Computational Intelligence and Natural Computing ,* 2009. CINC '09 . , vol.2, pp.323-325, 6-7 June 2009 doi: 10.1109/CINC.2009.89

[19]   Kamer Kayaer, " Medical Diagnosis on Pima Indian Diabetes Using General Regression Neural Networks"

[20]   R.R. Jangler, Anupama Shukla and Ritu Tiwari ," Intelligent Decision Support System for Breast Cancer", *The International Conference on Swar Intelligence ICSI 2010*, Part II, LNCS 6146, pp. 351-358, 2010

[21]   J.Warren., G. Beliakov and Van der Zwaag. B., "Fuzzy Logic in Clinical Practice Decision Support Systems,", *in Proceedings of the 33rd Annual Hawai International Conference on System Science-2000.*, pp. 10, 4-7 Jan. 2000 doi: 10.1109/HICSS.2000.926789

[22]   V. Anuja Kumari, R.Chitra, "Classification Of Diabetes Disease Using Support Vector Machine", International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com Vol. 3, Issue 2, March -April 2013, pp.1797-1801

[23]   Kok, "Predicting blood glucose levels of diabetics using artificial neural", Research Assignment for the Master of Science program Computer Science, Delft University of Technology, 2004

[24]   W.A.Sandham, D.J.Hamilton, A.Japp, K.Patterson., "Neural Network and Neuro-Fuzzy Systems For Improving Diabetes Therapy", *in the Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* 1998, vol.3, pp.1438-1441, 29 Oct-1 Nov 1998. doi: 10.1109/IEMBS.1998.747154

[25]   M. F. Alamaireh, "A Predictive Neural Network Control Approach in Diabetes Management by Insulin Administration", *Information and Communication Technologies*, 2006. ICTTA'06, 2nd, vol.1, pp.1618-1623. doi: 10.1109/ ICTTA.2006. 1684626

[26]   E.I. Papageorgiou, N.I.Papandrianos, D. Apostolopoulos and P Vassilakos, "Complementary use of Fuzzy Decision Trees Decision Making in Medical Informatics,", *International Conference on BioMedical Engineering and Informatics,* 2008. BMEI 2008., vol.1, pp.888-892, 27-30 May 2008 ; doi: 10.1109/BMEI.2008.275

[27]   G. Czibula, I.G. Czibula, G.S. Cojocar and A.M. Guran, "IMASC − An Intelligent MultiAgent System for Clinical Decision Support", *First International Conference on Complexity and Intelligence of the Artificial and Natural Complex Systems , Medical Applications of the Complex Systems, Biomedical Computing, 2008*. CANS '08., pp.185-190, 8-10 Nov. 2008 doi: 10.1109/CANS.2008.28

[28]   Xiaofeng Zhao, Liyan Jia, Linping An and Liyan Wang ,"A Data Mining Method on the Study of Medical Information",*2010 international Conference onComputer Application and SystemModeling (ICCASM)*, vol.9, pp. 183-186, 2224 Oct. 2010. doi: 10.1109/ICCASM.2010.5623058

[29]   N. Barakat.,A. P. Bradley Barakat and Mohamad. Nabil H., "Intelligible Support Vector Machines for Diagnosis of Diabetes Mellitus," *IEEE Transaction on Information Technology in Biomedicine* , vol.14, no.4, pp.1114-1120,July2010. doi: 10.1109/TITB.2009.2039485