An Application of Extreme Learning Machine for Diagnosis of Diabetes Mellitus

Thiyagarajan C, **Dr. K. Anandha Kumar**, **Dr. A. Bharathi**

Assistant professor in Computer Science, Research Scholar, Bharathiar University, PSG College of Arts & Science, Coimbatore, Tamilnadu, India
Department of Computer Applications, Bannari Amman Institute of Technology, Sathyamangalam, Tamilnadu, India
Professor, Department of Information Technology, Bannari Amman Institute of Technology, Sathyamangalam, Tamilnadu, India

Abstract
Diabetes Mellitus is one of the major health issues in all over the world. This is a disease characterized by lack of or ineffective utilization of the body's insulin. To detect the diabetes mellitus, a machine learning based classification model is proposed with optimal learning and precise performance. Several algorithms are used to detect and diagnosis the diabetes mellitus. In this paper Support Vector Machine (SVM) is compared with proposed Extreme Learning Machine (ELM). To create a database to classify and compare the task with the help of the SVM algorithm and this work concentrated on learning method focuses on splitting diabetes disease from high dimensional dataset. The proposed method uses ELM as the classifier for diagnosis of diabetes. Pima Indian Diabetic Dataset of UCI machine learning repository is used to analyses the performance of diabetes data classification. The proposed ELM technique successfully used for diagnosing diabetes disease, it is shown in the experimental results compared with SVM applied on the same database.
Keywords:
Diabetes Mellitus, Machine Learning, Extreme Learning Machine, Support Vector Machine

1. Introduction

DIABETES Mellitus is one of the major problems in health condition. It may occur at retinal area, nerves and kidneys. Due to diabetes mellitus, body doesn't produce and doesn't make enough insulin. This rate is also critical, Diabetes Mellitus being one of the major contributors to the mortality rate. To regulate the amount of sugar in the body, the amount of insulin production want to be high, but in diabetes it is not achieved. It is important issue that early state detection of disease is needed. The Diabetes Mellitus is classified into three types, such as "type 1 DM", which results from the body's failure to produce insulin, and currently requires the person to wear an insulin pump. This is called as "Insulin-Dependent Diabetes Mellitus" (IDDM). If the cell is combined with an absolute insulin deficiency or cells fail to use insulin properly is termed as Type II DM due to insulin resistance. This is also termed as "Non-Insulin Dependent Diabetes Mellitus" (NIDDM) or "adult-onset diabetes".

The gestational diabetes occurs when women pregnant without a previous diagnosis of diabetes develop a high blood glucose level; it is a type of development of type I DM. Other forms of DM include congenital diabetes, it occurs due to genetic defects of insulin secretion or means of cystic fibrosis-related diabetes, steroid diabetes induced by high doses of glucocorticoids, and several forms of monogenic diabetes. The diagnosis and interpretation of the diabetes data is must because major problem occurs due to this data maintenance. The several researches are made with machine learning algorithms such as Bayes network, Multilayer perceptron and decision tree. The classifier is to be designed is cost efficient, accurate, convenient and important.

The Support Vector Machine (SVM) is a method applied in several financial applications recently, mainly in the area of time series prediction and classification. The aim of this work is to develop a classification algorithm for DM diagnosis and treatment using an ELM classifier algorithm and compared with SVM standard method. The rest of the work is organized as follows, section 2 discusses about the previous work in this field, section 3 presents the proposed research methodology, section 4 evaluates the proposed method and finally section 5 provides the conclusion.

2. Literature Survey

Ewaldet al (2012) investigated 1868 patients diagnosed with diabetes mellitus who had been admitted to our hospital during the last 24 months. The ultimate aim of Aslam et al (2013) is to facilitate the diagnosis of diabetes, a rapidly increasing disease in the world. A Genetic Programming (GP) based method has been used for diabetes classification. GP has

Manuscript received September 5, 2026
Manuscript revised September 20, 2026
https://doi.org/10.22937/IJCSNS.2026.26.9.11

been used to generate new features by making combinations of the existing diabetes features, without prior knowledge of the probability distribution. The goal of Henderson et al (2014) study of patients treated with clozapine was to examine the incidence of treatment-emergent diabetes mellitus in relation to other factors, including weight gain, lipid abnormalities, age, clozapine dose, and treatment with valproate. Canivell and Ramon (2014) discuss the epidemiology, etiopathogenesis, diagnostic criteria and clinical classification of what is currently known as autoimmune diabetes. Klompaset al (2013) extracted 4 years of data from the EHR of a large, multisite, multispecialty ambulatory practice serving ～700,000 patients. They flagged possible cases of diabetes using laboratory test results, diagnosis codes, and prescriptions. They assessed the sensitivity and positive predictive value of novel combinations of these data to classify type 1 versus type 2 diabetes among 210 individuals.

A nation-wide cohort, the Better Diabetes Diagnosis study was used by Ludvigsson et al (2012) to determine serum C-peptide at diagnosis in 2734 children and adolescents. Faust et al (2012) review algorithms used for the extraction of these features from digital fundus images. Furthermore, they discuss systems that use these features to classify individual fundus images. Akram et al (2013) proposed a three-stage system for early detection of MAs using filter banks. In the first stage, the system extracts all possible candidate regions for MAs present in retinal image. Acharya et al (2012) documents a system that can be used for automatic mass screenings of diabetic retinopathy. Four classes are identified: normal retina*,* Non-Proliferative Diabetic Retinopathy (NPDR), Proliferative Diabetic Retinopathy (PDR), *and* Macular Edema *(ME)*. Zürbig et al (2012) evaluated urinary proteome analysis as a tool for prediction of DN. Capillary electrophoresis–coupled mass spectrometry was used to profile the low–molecular weight proteome in urine.

Siwyet al (2014) aimed to validate for the first time the CKD273 classifier in a multicentre (9 different institutions providing samples from 165 T2D patients) prospective setting. Giancardo et al (2012) introduce a new methodology for diagnosis of DME using a novel set of features based on colour, wavelet decomposition and automatic lesion segmentation. The proposed method by Kumari et al (2013) uses Support Vector Machine (SVM), a machine learning method as the classifier for diagnosis of diabetes. The machine learning method focus on classifying diabetes disease from high dimensional medical dataset. Noronha et al (2012) have used discrete wavelet, transform and support vector machine classifier for automated detection of normal and diabetic retinopathy classes. The wavelet-based decomposition was performed up to the second level, and eight energy features were extracted. A new classification scheme based on Support Vector Machine (SVM) selected features to train Rotation Forest (RF) ensemble classifiers is presented by Ozcift (2012) for improving diagnosis of PD.

3. Support Vector Machine Classifier

This section provides the detail description of the proposed methodology of SVM classifier. SVM is a method used in medical diagnosis for classification and regression. IT also simultaneously minimizes the empirical classification error and maximizes the geometric margin. Due to this SVM is also called Maximum Margin Classifiers. Totally SVM is a based on based on guaranteed risk bounds of statistical learning theory i.e. the so called structural risk minimization principle. For set of training and analysis SVM is used in the form of algorithm. The classifier either it may be a binary linear or nonlinear. In addition to that linear form, It efficiently perform a non linear classification. SVMs can efficiently perform non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces. Without knowing the space the classifier is constructing with the control of kernel trick

An SVM model is a represented by points in space, mapped so that the examples of the separate categories are divided by a clear gap. Hence the space becomes wide as much as possible. For example, if a point is initialized then set of points belonging to either one of the two classes, an SVM finds a hyperplane having the largest possible fraction of points of the same class on the same plane. This separating hyperplane is called the Optimal Separating Hyperplane (OSH) that maximizes the distance between the two parallel hyper planes. It also reduce the risk of misclassifying examples of the test dataset.

Given labeled training data as data points of the form

$$M = \{(x_1, y_1), (x_2, y_2), \dots\dots, (x_n, y_n)\} \qquad (4)$$

where y_n=1/-1, a constant that denotes the class to which that point x_n belongs. n=number of data sample. Each $x_n$is a p-dimensional real vector. The SVM classifier first maps the input vectors into a decision value, and then performs the classification using an appropriate threshold value. To view the training data, we divide (or separate) the hyperplane, which can be described as:

Mapping: $w^T.x + b = 0$

Where w is a p-dimensional weight vector and b is a scalar. The vector w points perpendicular to the separating hyperplane. The offset parameter b allows increasing the margin. When the training data are linearly separable, select these hyperplanes so that there are no points between them and then try on maximizing the distance between the hyperplane. Found out that the distance between the hyperplane as 2/ |w|. To minimize |w|, need to ensure that for all i either

$$w.x_i - b \geq 1\ or w.x_i - b \leq -1 \qquad (5)$$

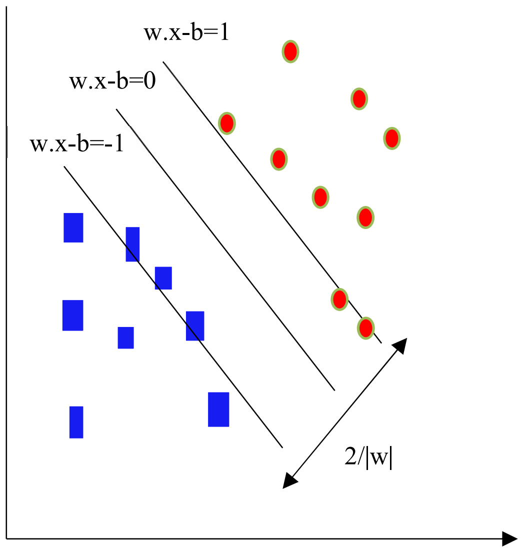

Figure 1: Maximum Margin Hyper Planes for ELM Trained with Samples from Two Classes

4. ELM Classifier

The drawback of the SVM is slow in test phase and size complexity. A new learning algorithm for Single Hidden Layer FeedForward Networks (SLFNs), called Extreme Learning Machine (ELM), which helps in solving regression and classification problems. This algorithm also used to reach good solutions analytically, and its learning speed is faster than other traditional methods.

The input weights and biases are not updated, it determined randomly during training iterations only. For hidden neuron layer the activation function like sine, Gaussian may be selected and for output neurons linear activation functions are used. It is a Multi class classification where number of output neurons will be automatically set equal to number of classes.

The output weights are obtained by using norm Least Squares (LS) and pseudo inverse of a linear system.

For a given N arbitrary input-output relation (x_i,t_i), where $x_i = [x_{i1}, x_{j2}, \ldots, x_{in}]^T \in R^n$ and $t_i = [t_{i1}, t_{j2}, \ldots, t_{in}]^T \in R^m$, a single layered network with $\tilde{N}$ neurons in hidden layer and a given activation function g(x) is modeled as

$$\sum_{i=1}^{\tilde{N}} \beta_i g(w_i.x_j + b_i) = o_j, j = 1, \ldots, N \qquad (1)$$

where $w_i = [\omega_{i1}, \omega_{i2}, \ldots, \omega_{in}]^T$ shows the weight of the i^{th}hidden neuron, and $\beta_i = [\beta_{i1}, \beta_{i2}, \ldots, \beta_{in}]^T$ shows the weight of the i^{th} hidden neuron to the output neurons, and b_i is the bias of the i^{th} hidden neuron.

The SLFNs which is defined in Eq. (1) can be approximated with zero error. In other words, the Eq. (1) can be rewritten in the following form:

$$H\beta = T \qquad (2)$$

After calculation of H, the output weights are determined by using below equation

$$\beta = H^* \times T \qquad (3)$$

TABLE 1
ACCURACY AND EXECUTION TIME

Techniques	Accuracy (%)	Execution time (seconds)
SVM	79	45
ELM	94	21

The SVM models for classification have been developed for the classification of diabetes dataset. The diagnostic performance of the developed models is evaluated using classification accuracy and execution time.

5. Experimental Results

The experimental results are carried out using PIMA dataset. Pima Indian Diabetes Data (PIDD) set is available publicly from the machine learning database at UCI (Smith *et al.,* (1988)). This dataset consist only females at the age of 21 of Pima Indian heritage living near Phoenix, Arizona. This data set is extracted from a larger database originally owned by the National Institute of Diabetes and Digestive and Kidney Diseases.

Our proposed method is evaluated using accuracy and execution time in PIMA dataset. Table 1 shows the accuracy and execution time for the SVM and the proposed ELM classifier. The proposed method ELM achieves better accuracy 94% with less execution time as 21 seconds than the preceding method.

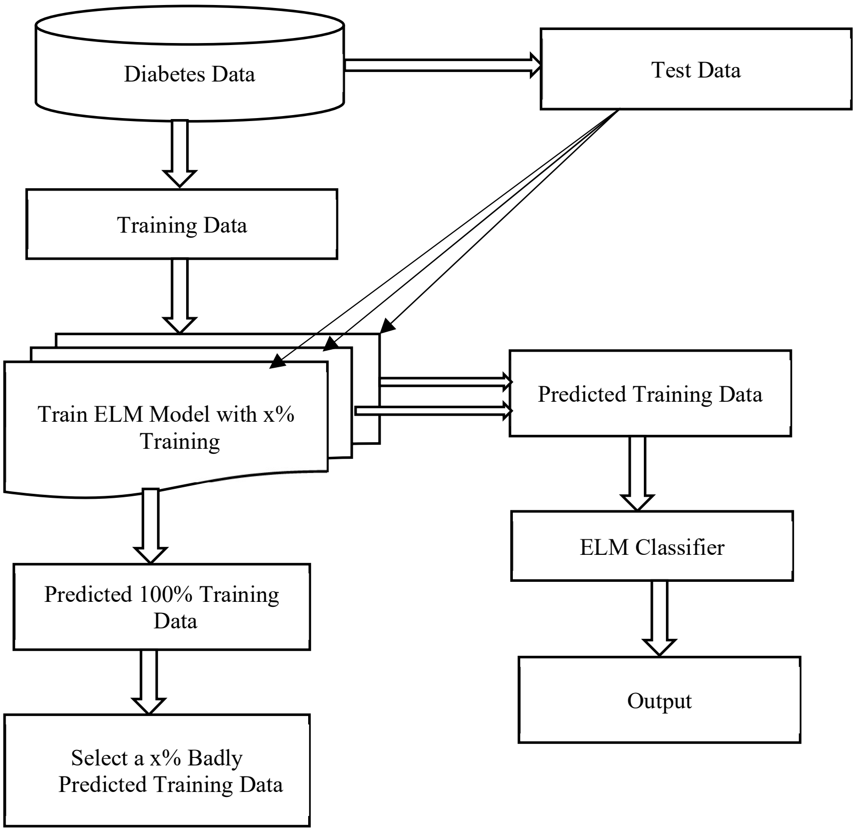

Figure 2: Architecture of the Proposed System

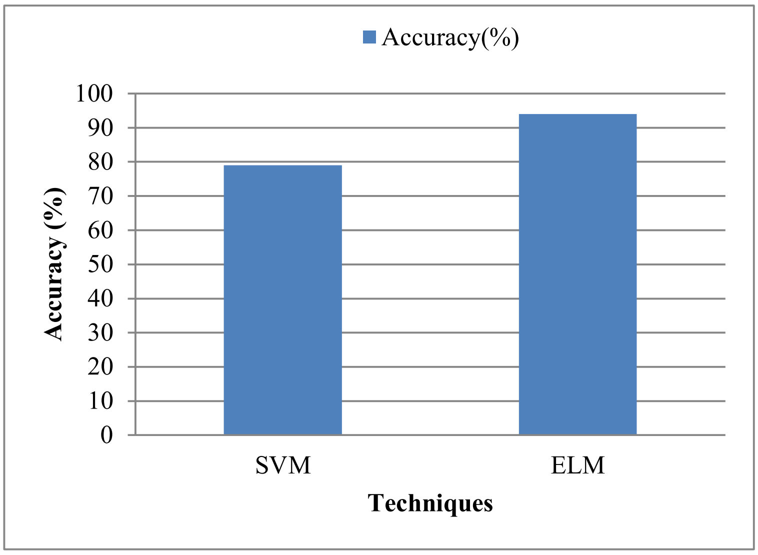

Figure 3: Accuracy

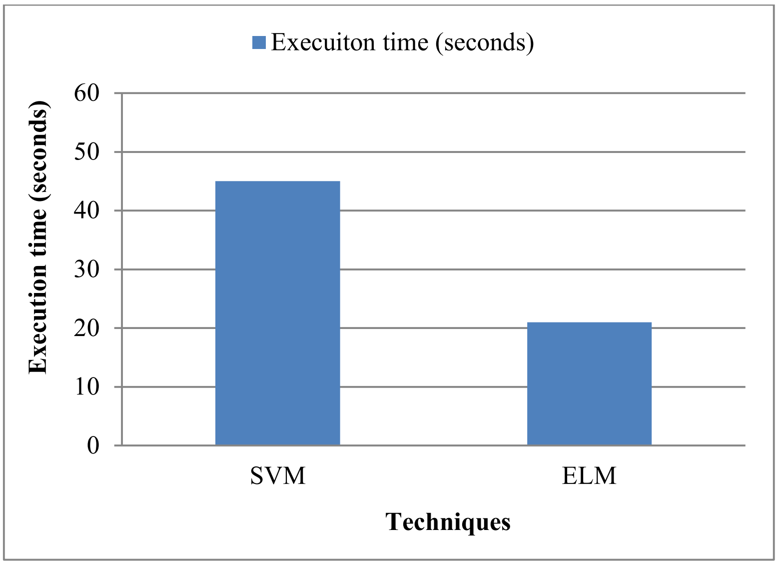

Figure 4: Execution Time

Figure 3 and 4 shows the accuracy and execution time for both SVM and the proposed ELM method. It is clear from the figure that proposed method provides better accuracy than the preceding method. Execution time is also very low when compared with the SVM.

6. Conclusion

In this work, diabetes classification using extreme learning machine is proposed. All the patients' data are classified by ELM. The Overall classification accuracy of the ELM is computed by averaging the classification accuracies. The performance parameters such as the classification accuracy and execution time of the ELM have found to be high thus making it a good option for the classification process. From the experiment results, it is concluded that, results of the proposed system can perform better than the SVM. Though, the obtained results shall be comparable with those of the SVMs, but with lesser complexity of the propose system, and minimum compromise with the quality, such a trade-off is acceptable. In future the performance of ELM classifier can be improved by feature subset selection process.

References

[1] Ewald, N., C. Kaufmann, A. Raspe, H. U. Kloer, R. G. Bretzel, and P. D. Hardt. "Prevalence of diabetes mellitus secondary to pancreatic diseases (type 3c)." Diabetes/metabolism research and reviews 28, no. 4 (2012): 338-342.

[2] Aslam, Muhammad Waqar, Zhechen Zhu, and Asoke Kumar Nandi. "Feature generation using genetic programming with comparative partner selection for diabetes classification." Expert Systems with Applications 40, no. 13 (2013): 5402-5412.

[3] Henderson, David C., Enrico Cagliero, Colin Gray, Rima A. Nasrallah, Doug L. Hayden, David A. Schoenfeld, and Donald C. Goff. "Clozapine, diabetes mellitus, weight gain, and lipid abnormalities: a five-year naturalistic study."American Journal of Psychiatry (2014).

[4] Canivell, Silvia, and Ramon Gomis. "Diagnosis and classification of autoimmune diabetes mellitus." Autoimmunity reviews 13, no. 4 (2014): 403-407.

[5] Klompas, Michael, Emma Eggleston, Jason McVetta, Ross Lazarus, Lingling Li, and Richard Platt. "Automated detection and classification of type 1 versus type 2 diabetes using electronic health record data." Diabetes Care36, no. 4 (2013): 914-921.

[6] Ludvigsson, Johnny, A. Carlsson, Gun Forsander, StenIvarsson, I. Kockum, ÅkeLernmark, Bengt Lindblad, C. Marcus, and Ulf Samuelsson. "C-peptide in the classification of diabetes in children and adolescents." Pediatric diabetes 13, no. 1 (2012): 45-50.

[7] Faust, Oliver, Rajendra Acharya, Eddie Yin-Kwee Ng, Kwan-Hoong Ng, and Jasjit S. Suri. "Algorithms for the automated detection of diabetic retinopathy using digital fundus images: a review." Journal of medical systems 36, no. 1 (2012): 145-157.

[8] Akram, M. Usman, Shehzad Khalid, and Shoab A. Khan. "Identification and classification of microaneurysms for early detection of diabetic retinopathy."Pattern Recognition 46, no. 1 (2013): 107-116.

[9] Acharya, U. Rajendra, Eddie Yin-Kwee Ng, Jen-Hong Tan, S. VinithaSree, and Kwan-Hoong Ng. "An integrated index for the identification of diabetic retinopathy stages using texture parameters." Journal of medical systems36, no. 3 (2012): 2011-2020.

[10] Zürbig, Petra, George Jerums, Peter Hovind, Richard J. MacIsaac, Harald Mischak, Stine E. Nielsen, SiannaPanagiotopoulos, Frederik Persson, and Peter Rossing. "Urinary proteomics for early diagnosis in diabetic nephropathy." Diabetes 61, no. 12 (2012): 3304-3313.
[11] Siwy, Justyna, Joost P. Schanstra, Angel Argiles, Stephan JL Bakker, Joachim Beige, Petr Boucek, Korbinian Brand et al. "Multicentre prospective validation of a urinary peptidome-based classifier for the diagnosis of type 2 diabetic nephropathy." Nephrology Dialysis Transplantation (2014): gfu039.
[12] Giancardo, Luca, Fabrice Meriaudeau, Thomas P. Karnowski, Yaqin Li, Seema Garg, Kenneth W. Tobin, and Edward Chaum. "Exudate-based diabetic macular edema detection in fundus images using publicly available datasets." Medical Image Analysis 16, no. 1 (2012): 216-226.
[13] Kumari, V. Anuja, and R. Chitra. "Classification Of Diabetes Disease Using Support Vector Machine." International Journal of Engineering Research and Applications 3, no. 2 (2013): 1797-1801.
[14] Noronha, Kevin, U. R. Acharya, K. P. Nayak, S. Kamath, and S. V. Bhandary. "Decision support system for diabetes retinopathy using discrete wavelet transform." Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine (2012): 0954411912470240.
[15] Ozcift, Akin. "SVM feature selection based rotation forest ensemble classifiers to improve computer-aided diagnosis of Parkinson disease."Journal of medical systems 36, no. 4 (2012): 2141-2147.
[16] G.-B.Huang, Q.Y.Zhu, and C.K.Siew, Extreme learning machine: a new learning scheme of feed forward neural networks, in: Proceedings of the International Joint Conference on Neural Networks (IJCNN2004),Budapest, 2004, pp. 985– 990
[17] Cortes, C.; Vapnik, V. (1995). "Support-vector networks". Machine Learning. 20 (3): 273–297.
[18] Smith, J.W., Everhart, J.E., Dickson, W.C., Knowler, W.C., & Johannes, R.S. (1988). Using the ADAP learning algorithm to forecast the onset of diabetes mellitus. In Proceedings of the Symposium on Computer Applications and Medical Care, (pp. 261--265). IEEE Computer Society Press.