Deep Belief Neural Network Model for Prediction of Diabetes Mellitus

P. Prabhu[1]
[1]Assistant Professor in Information Technology
Directorate of Distance Education
Department of Computer Applications
Alagappa University, Karaikudi, Tamilnadu, India
e-mail: pprabhu70@gmail.com,
prabhup@alagappauniversity.ac.in

S. Selvabharathi[2]
[2]Assistant Professor(Part-Time)
Department of Electronics and Communication
Engineering
Alagappa Chettiar College of Engineering and
Technology, Karaikudi, Tamilnadu, India
e-mail: selvabharathime@gmail.com

Abstract— **Diabetes Mellitus is metabolic chronic disease in which blood glucose levels are too high. In India nearly 8.7% of population suffers from diabetes in age range from 20 to 70. Unidentified and untreated diabetes leads to so many health difficulties such as damage of heart, kidneys, eyes, nerves and blood vessels. There are already several methods exists to support clinical decision making but still need improvements to solve the issues and challenges. In this research work, deep belief network model is designed for providing computational intelligence for prediction of patient affected by diabetes mellitus with maximum accuracy. Pima Indians Diabetes Dataset is used to analyze and experiment this prediction model. Firstly, the dataset is pre-processed by applying normalization technique. Secondly, the prediction model using deep belief neural network is designed. At the end, an experimental results proved that the comparison of overall performance of deep belief networks method is better than familiar classifiers namely naïve Bayes, Decision Tree,. Logistic Regression (LR), Random Forest (RF) and Support Vector Machine (SVM).**

Keywords- Deep belief model, Predictive Analytics, Diabetes, Prediction Model, Machine Learning, Classification.

I. INTRODUCTION

Diabetes mellitus is a chronic disease which has various challenges and issues in human health. The most common types of diabetes mellitus are Type 1, Type 2, Prediabetes and Gestational diabetes; In type-1 diabetes is a chronic condition; patient's immune system mistakenly attacks and destroys the beta cells in patient's pancreas that produce insulin. In type-2 diabetes is a chronic condition in which insulin production decreases in patients body which lead to high blood sugar. Recent studies have shown that the 80% of type-2 diabetes can be prevented by early detection. Prediabetes is a condition in which blood sugar is high, but not high enough to be type 2 diabetes. Gestational diabetes affects pregnant women with high blood sugar.

According to the WHO (World Health Organization) the number of children having type 1 diabetes is very high as mentioned in the motivation. Hence it can be said that Diabetes is a serious chronic disease[11].

It is very important to establish predictive models using those risk factors for interventions relating to the development of diabetes. Previous studies have suggested that anthropometric measurement and adipocyte size can serve as predictors of diabetes incidence using traditional statistical methods [6-10].

Data mining is a step in Knowledge Discovery process consists of various techniques to extract intelligence from the data. In recent years, Artificial Neural Networks play vial role in prediction of disease. Deep Learning architecture consists of many levels used to learn functions. It has many non-linear operations with many hidden layers. Restricted Boltzmann Machines (RBM) is energy based probabilistic model, which is a restricted version of Boltzmann machines (BM) that is a log-linear Markov Random Field.

The purpose of this research work was to develop deep learning based Deep Belief Network intelligent model predict diabetes. This work also compares the experimental results of this model with conventional models.

The remaining of the research work is organized as follows; Section 2 describes the literature related to models and algorithms of the prediction of diabetes, Section 3 describes the methodology used in the prediction model, section 4 presents experimental dataset and discussion on experimental evaluation results and section 5 determines the conclusion of the research work.

II. RELATED WORKS

Research in any field needs review of literature. This section discuss about various literature related to deep belief networks and diabetes prediction.

Deepti Sisodia et.al., [3] experimentally analysed classification algorithms namely Decision Tree, SVM and Naive Bayes to detect diabetes at an early stage. The performances of algorithms are evaluated using Weka tool and compared using various statistical measures namely precision, recall and f measure.

Micheal Dutt [1] presents a Multi-Layer Feed Forward Neural Networks (MLFNN) model for the prediction of diabetes on publicly available Pima Indian Diabetes (PID) dataset. A several experiments are conducted on this dataset with variation in learning algorithms, activation units, techniques to handle missing data and their impact on classification accuracy have been discussed. The results are tabulated and compared with other machine learning algorithms namely Naive Bayes, Random Forest, and Logistic Regression.

J. Vijayashree, et al., [4]proposed an expert system for diagnosis of diabetic patients using feature reduction

978-1-7281-3663-9/19/$31.00 ©2019 IEEE

techniques namely Recursive feature elimination and principal component analysis for improving the performance. The expert system is built using Deep Neural Networks and Artificial Neural Networks. This study also compares the performance of proposed system using various performance measures namely accuracy, sensitivity and specificity.

Dilip Kumar Choubey[5] proposed a model for diagnosis of diabetic patients by two stage process. In first stage, features are selected using genetic algorithm from diabetes dataset. Second stage, predicts diabetes using multi-layer feed forward neural network. The experimental results are promising when compared with conventional models.

Xue-Hui Meng et al., [6]studied the performance of logistic regression, artificial neural networks (ANNs) and decision tree models for predicting diabetes or pre-diabetes using common risk factors. It helps the future researchers in selecting the optimal predictive models to decrease the incidence of diabetes.

Ramalingaswamy Cheruku[12] proposed Diabetes Classification using Radial Basis Function Network based on cluster validity index which determines optimal number of neurons. This model outperforms well when compared with probabilistic neural network model.

Ciresan, D., et al.[13] described the approach for German traffic sign recognition. DNNs trained on differently pre-processed data into a Multi-Column DNN (MCDNN) supervised way and achieved better human recognition accuracy.

Prabhu and Anbazhagan[14] presented an algorithm for prediction of human opinion about buying products based on the ranking. Neural Network is applied to predict the recommendations for active user. Back propagation neural network classification is used to classify users. This method outperforms in term of various statistical measures when compared with traditional methods.

Prabhu and Anbazhagan[15] proposed hybrid method for generation business intelligence (recommendations) for purchasing items through e-commerce. This method combines frequent item-set mining and clustering techniques to predict recommendations.

III. METHODOLOGY

DBN prediction model is designed in three phases; pre-processing, pre-training DBN followed by fine-tuning back propagation for supervised classification. The Fig. 1. Outline the block diagram of proposed method.

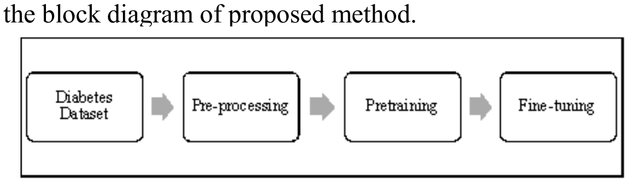

Figure 1 Design of DBN prediction model

The first phase is a pre-processing phase; in this phase diabetes dataset is pre-processed by applying normalization. Diabetes datasets are normalized as data pre-processing for transforming all variable in the same range (0 and 1) using min-max normalization.

Training, validation and test diabetes dataset values are normalized by dividing each value for a sample by the maximum value observed in any sample. Standardization leads to mean value is 0 and standard deviation is 1. Principle component Analysis (PCA) is used to select the relevant features from the training dataset using eigenvectors coefficients and weights. R is correlation matrix, V eigenvector matrix and D diagonal matrix of eigenvalues. The scores are calculated as follows: Z = YV where Z matrix of principal components scores ($n \times m$), Y standardized data matrix ($n \times p$) and V matrix of eigenvectors ($p \times m$).

Second phase is Pretraining, where DBM is constructed by 'stacking' of RBMs and trained. It is a generative graphical model. It is also a multi layered model. This process is used to construct the DBN is named as pre-training of DBN used to model features for classification of datasets. Each layer is trained as RBMs. First step is of training DBN is to train layers sequentially from the bottom visible (observed) layer features. This input layer contains D number of units, where D is input sample dimension. This input layer is fully connected with hidden layers. Each Hidden layer consists of N units of RBM. The output layer consists of one unit which defines the class.

The final phase, called fine tuning is to train (classifier) second layer based on the results from pre-training step. Finally, the entire hidden layers are learned same way till final hidden layer is reached. The Fig. 2. outlines the block diagram of pre-training proposed method.

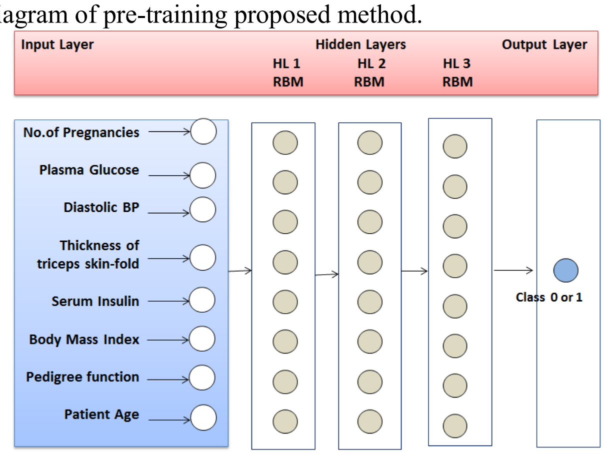

Figure 2. Block diagram of pre-training

The algorithm for DBN model for prediction of diabetes mellitus is shown below:

Algorithm:

Input: Diabetes dataset-Train, Validate, Test
Target dataset-Train, Validate, Test
Output: Assigned class label for each sample.

Begin
Phase 1:
//Pre-processing
Apply min-max Normalization
Feature selection using PCA

R = VDV',
where ,
D is a diagonal matrix whose diagonal elements are the eigenvalues.
V matrix of eigenvectors ($p \times m$)

Z = YV
Where;
Z matrix of principal components scores ($n \times m$)
Y standardized data matrix ($n \times p$) used with the correlation matrix method.
V matrix of eigenvectors ($p \times m$)

Phase 2 :
//Pretraining

a. Construct a fully connected, DBN network, (with D input layer units, H hidden layer units, one output layer unit).
b. Train dataset using Deep Belief Network.
c. Validate dataset using DBN.

Phase 3 :
// Fine-tuning of classification

Apply NN-FF classification on test dataset
Randomly initialize all weights and biases in network

a. Repeat the steps (a-d)
b. Propagate the inputs forward in the network
c. Back-propagate the errors
d. Until termination condition is satisfied

Calculate Validation measures.

End

IV. RESULT AND DISCUSSION

A. Experimental Setup

In this section, the experimental setup for classification of diabetes implemented using DBN is discussed. All our experiments were implemented and simulated using Matlab, Windows 7 Home Basic based PC with Core i3 processor having a speed of 2.40 GHz and 3GB of RAM. Experiments are performed using various accuracy measures like Recall(R), Precision(P) and F1 measure (F).

Recall can be computed using TP and FN.

Recall(R)= TP/(TP+FN)

where;
TP is True Positive and FN is False Negative

Precision can be computed using TP and FP.

Precision(P)= TP/(TP+FP)
where FP is False Positive,

F1 can be computed as a weighted average of recall and precision.

F1= (2 × recall × precision)/(recall + precision)

B. Pima Indian Women Diabetes Dataset (PIDD)

Pima Indian Women Diabetes Dataset (PIDD) obtained from UCI machine repository [2] is used to test the performance of the proposed model as a benchmark. This dataset consists of 768 samples of female patients with 500 samples identified as negative class, 268 are identified as positive. It contains 8 input attributes and one output attribute for target class.

Different values of parameters are tested for performance. The best parameters for classification are identified. The number of hidden layers is identified by various testing. Small number of layers gives poor result. Too many number of layers leads to over fitting. Optimum number of layers are determined by conduction various experiments. Training a neural network requires a certain number of RBM epochs to converge into an optimal value. When it fails converge to optimal value poor results were obtained. Testing with various numbers of iterations leads to better results. Table I shows the Pretraining parameters and its values used to train deep belief networks.

TABLE I. PRETRAINING PARAMETERS OF DEEP BELIEF NETWORKS

S.No	Parameters	Value	Explanation
1	Hidden activation functions	[ReLu, ReLU, ReLU]	(Rectifier Linear Unit)
2	Number of Hidden Layers	3	Hidden Layers values
3	Number of Hidden Units	[500 500 1000]	Total Number of Hidden units in the neural network
4	Input Activation function	Sigmoid	Sigmoid activation function
5	RBM epochs	10	Number of RBM epochs

Weight initialization takes important role in getting speedy results. Too small or large value of weight also leads to poor results. Table II shows fine-tuning parameters and its values of Neural Network.

TABLE II. FINE-TUNING PARAMETERS OF NEURAL NETWORK

S.No	Parameters	Value	Explanation
1	No. of epochs	5	Total Number of iterations
2	Global Learning Rate	0.01	Weight updates during the training.
3	No. of Layers	4	Number of Layers
4	Network topology	[500 500 1000 2]	Number of units in Layers
5	Activation function	[ReLu, ReLU, ReLU, softmax]	(Rectifier Linear Unit) and Softmax
6	Training method	SGD	(Stochastic Gradient Descent) with momentum
7	Momentum	0.5	Speed up the learning process
8	Max. Norm Constraints	4	
9	Weight Initialization parameter type	8	Gaussian Distribution with 0 mean
10	Batch size	100	Mini-batch, Full Batch

The Fig. 3 shows the comparison between training and validation error in various number of epochs.

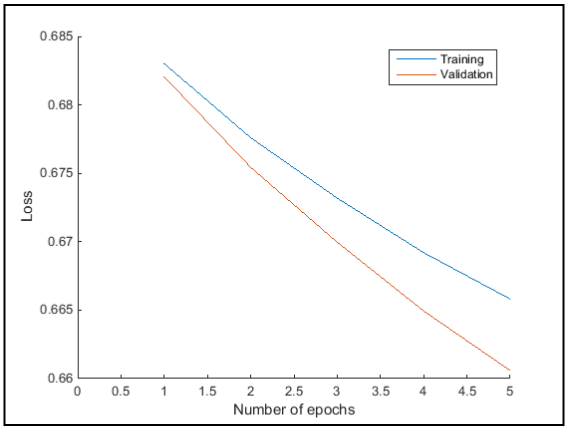

Figure 3. Comparison between training and validation error.

Table III represents comparison of different classification algorithms namely Naïve Bayes, Radial Basis Feed Forward – Neural Network (RBF-NN), Decision Tree (DT), Logistic Regression(LR)[1],Random Forest(RF)[1] and SVM with proposed DBM method using various measures. From Table III it is observed that DBN shows better results.

TABLE III. COMPARATIVE PERFORMANCE OF CLASSIFICATION ALGORITHMS ON VARIOUS MEASURES.

Algorithm	Recall	Precision	F1 Measure
Deep Belief Networks(DBN)	1.0	0.6791	0.808
Naïve Bayes (NB)	0.759	0.763	0.760
RBF-NN	0.761	0.756	0.757
Decision Tree (DT)	0.738	0.735	0.736
Logistic Regression(LR)	0.73	0.73	0.73
Random Forest (RF)	0.71	0.72	0.72
Support Vector Machine (SVM)	0.424	0.651	0.513

Figure 4. shows the comparison of F1 measure between proposed and conventional methods.

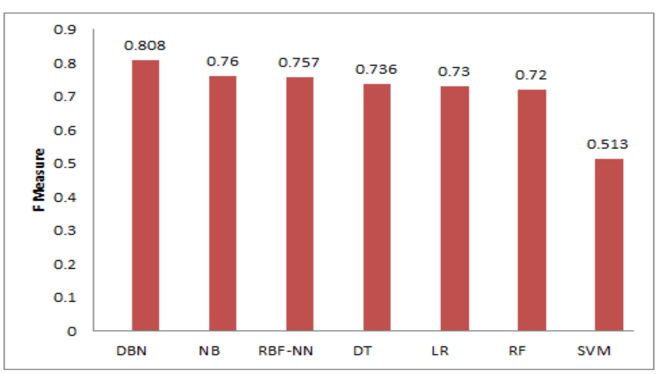

Figure 4. Comparison of F1 measure between proposed and conventional methods.

When compared with conventional classifiers such as Naïve Baye's, Decision Tree, Logistic Regression and Random Forest. DBN classifier model gives more prediction accuracy when measured in terms of precision, recall and f1 statistical measures using UCI machine repository diabetes datasets. Similarly synthetic datasets also can be experimentally testes by collecting diabetes mellitus dataset collected from various hospitals.

V. CONCLUSION AND FUTURE WORK

The deep belief prediction model was successfully designed, trained, validated and tested for performance benchmark using Pima Indian Diabetes Dataset. Various experiments with different datasets and parameters are experimented and results are obtained and tabulated. The comparison of results shows that the deep belief model is more effective in terms of recall, precision and F1 measure. This proposed deep belief network based model can be constructed using optimization techniques and tested for other kinds of diseases prediction as a future work.

ACKNOWLEDGMENT

This research work was carried out with the financially support of RUSA – Phase 2.0 grant sanctioned vide Letter No. F 24-51 / 2014-U, Policy (TNMulti-Gen) Dept. of Edn. Govt. of India, Dt.09.10.2018 at Alagappa University, Karaikudi, Tamilnadu. India.

REFERENCES

[1] Micheal Dutt, Classification of diabetes and cardiac arrhythmia using Deep Learning, Master Thesis, Department of ICT, Faculty of Science and Engineering, University of Agder,2018.

[2] UCI Machine Repository : https://archive.ics.uci.edu/ml/machine-learning-databases/pima-indians-diabetes/.

[3] Deepti Sisodia, Dilip Singh Sisodia, Prediction of Diabetes using Classification Algorithms, International Conference on Computational Intelligence and Data Science (ICCIDS 2018), Procedia Computer Science 132 (2018) 1578–1585.pp 1578-1585.

[4] J. Vijayashree, J. Jayashree. An Expert System for the Diagnosis of Diabetic Patients using Deep Neural Networks and Recursive Feature Elimination. International Journal of Civil Engineering and Technology, 8(12), 2017, pp. 633-641.

[5] Dilip Kumar Choubey, Sanchita Paul, GA_MLP NN: A Hybrid Intelligent System for Diabetes Disease Diagnosis, I.J. Intelligent Systems and Applications, 2016, 1, 49-59

[6] Xue-Hui Meng, Yi-Xiang Huang, Dong-Ping Rao, Qiu Zhang , Qing Liu ,Comparison of three data mining models for predicting diabetes or prediabetes by risk factors, Kaohsiung Journal of Medical Sciences (2013) 29, 93-99.

[7] Schulze MB, Heidemann C, Schienkiewitz A, Bergmann MM,Hoffmann K, Boeing H. Comparison of anthropometric characteristics in predicting the incidence of type 2 diabetes in the EPIC-Potsdam study. Diabetes Care 2006;29:1921-3.

[8] Nyamdorj R, Qiao Q, Soderberg S, Pitkaniemi JM, Zimmet PZ, Shaw JE, et al. BMI compared with central obesity indicators as a predictor of diabetes incidence in Mauritius. Obesity 2009;17:342-8.

[9] Jia Z, Zhou Y, Liu X, Wang Y, Zhao X, Wang Y, et al. Comparison of different anthropometric measures as predictors,of diabetes incidence in a Chinese population. Diabetes Res Clin Pract 2011;92:265-71.

[10] Lonn M, Mehlig K, Bengtsson C, Lissner L. Adipocyte size predicts incidence of type 2 diabetes in women. Faseb J 2010; 24:326-31.

[11] Sushant Ramesh, H. Balaji, N.Ch.S.N Iyengar and Ronnie D. Caytiles, Optimal Predictive analytics of Pima Diabetics using Deep Learning, International Journal of Database Theory and Application Vol.10, No.9 (2017), pp.47-62.

[12] Ramalingaswamy Cheruku, Damodar Reddy Edla, Venkatanareshbabu Kuppili, Diabetes Classification using Radial Basis Function Network by Combining Cluster Validity Index and BAT Optimization with Novel Fitness Function, International Journal of Computational Intelligence Systems, Vol. 10 (2017) 247–265. doi.org/10.2991/ijcis.2017.10.1.17

[13] Ciresan, D., Meier, U. Masci, J. & Schmidhuber, J. Multi-column deep neural network for traffic sign classification. Neural Networks 32, 333–338 (2012).

[14] P.Prabhu, N Anbazhagan, A Neural Network Based Collaborative Filtering Model, International Journal of Science and Technology and Management,(4)(1),pp.164-175(),2015.

[15] P Prabhu, N Anbazhagan, A new hybrid algorithm for business intelligence recommender system, International Journal of Network Security & Its Applications,)(6)(2),2014,pp.43