Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 9s (2025)
https://www.jneonatalsurg.com

Multi-Layer Neural Network (Mlnn)-Based Classification Of Diabetes

C. B. Pavithra[1]

[1]Assistant Professor, Department of Computer Applications, KG College of Arts and Science, Coimbatore

Email ID: pavithra.b@kgcas.com

Email ID: cbpavithra@gmail.com

Cite this paper as: C. B. Pavithra, (2025) Multi-Layer Neural Network (Mlnn)-Based Classification Of Diabetes. *Journal of Neonatal Surgery*, 14 (9s), 424-433

ABSTRACT

Diabetes Mellitus (DM) is found to be common among people nowadays. Factors like age, obesity, absence of exercise, genetic factors, life style, improper diet, increased blood pressure, etc. cause diabetes. People with diabetes are highly prone to diseases in the heart and kidney, stroke, problem in the eye, damage in the nerve etc., as glucose levels are not accurately controlled. Necessary information for diagnosing diabetes is to be gathered by conducting diverse tests and suitable treatment must be offered depending on diagnosis. The existing schemes do not offer improved classification as well as prediction accuracy. In this paper, a Deep Learning (DL)-based diabetes prediction model is proposed for offering improved classification based on external factors which cause diabetes in addition to other factors including Glucose, Age, BMI, Insulin, etc. The proposed Long Short-Term Memory (LSTM)-based no-prop Multi-Layer Neural Network (LSTM-MLNN) offers improved classification accuracy in contrast to existing methods of classification of DM. Experimental analysis is performed based on sensitivity, specificity and accuracy.

Keywords: *Diabetes, Machine Learning, Deep Learning, Support Vector Machine (SVM), Long Short-Term Memory (LSTM), Proposed Multi-Layer Neural Network (MLNN), Random Forest (RF)*

1. INTRODUCTION

In the recent past, chronic illnesses related to upsurge in blood sugar have radically increased. Diabetes, a metabolic disease identified with increased levels of blood sugar along with organ damage like kidney failure seems to be common [1,2]. Insufficiency in insulin production as well as utilization that causes diabetes can be identified by intelligent models designed for finding diseases related to clinical data [3, 4].

This disease called Diabetes Mellitus (DM) is found to be common in adults. It is based on genetics as well as diagnostic criteria in addition to region's epidemics. Beta cells differ from infants to adults [5 - 7].

DM is categorized into Type 1, 2 and 3 [8, 9].

- **Type 1:** This type of diabetes is developed when sufficient insulin is not secreted by beta cells. It occurs in people under age of 30 who are insulin-dependent.
- **Type 2:** This kind of diabetes occurs when insulin is created by beta cells and body is incapable of using it. It is non-insulin-based and is seen in adults. It is due to secretion loss of insulin. Obesity, genetics and unhealthy lifestyle lead to the increase. It happens in middle age. It occurs before Type 3 in females as well as ethnic populations. It influences adolescents besides children.
- **Type 3:** This is called gestational diabetes which happens during trimesters of pregnancy. This goes off inevitably subsequent to child birth or may develop to Type 2 diabetes. It is seen based on maternal features during second as well as third pregnancy trimesters besides biomarkers during gestation.

To offer better treatment and support healthcare, an automatic information system to identify DM is essential [10, 11].

A novel method of classification is proposed using no-prop MLNN algorithm. The proposed algorithm involves 2 main stages namely, training and testing stage.

In every stage, appropriate features are determined using attribute-selection, and NN is trained as a multi-layer, commencing with normal (Type 1 & 2), and healthy or gestational diabetes. Classification is performed using MLNN framework.

An intelligent model that plays a dominant role in medical analysis depending on ML schemes with improved efficacy as well as accuracy is essential. The proposed framework is shown in Figure 1.

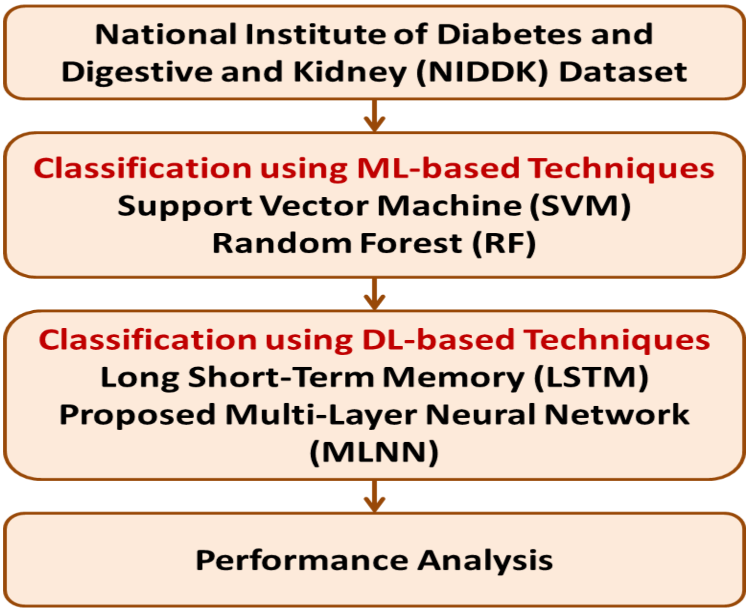

Figure 1: Propounded Framework

The paper is organised as follows. Section 2 details about work done associated with classification of diabetes. Section 3 deals with Machine Learning (ML)-based schemes for predicting diabetes. Section 4 focuses on the Deep Leaning (DL)-based techniques for classifying diabetes. Finally, the outcomes are discussed in Section 5. Conclusion is given in Section 6.

2. RELATED WORK

The works done by diverse authors based on ML and DL are discussed in detail in this section.

Dutta et al (2018) [12] have proposed an automated model to find key antecedents of Diabetic Retinopathy (DR). The propounded model is trained using Back Propagation NN (BPNN), Convolutional Neural Network (CNN) and Deep Neural Network (DNN). DL-based models outperform NN. DL models are efficient in quantifying features including blood vessels, exudates, fluid drip, micro aneurysms and haemorrhages into diverse classes. It computes weights which show the severity degree of patient's eye. The main challenge is to find the precise verdict of every class thresholds. Weighted Fuzzy C-Means (WFCM) algorithm is used for detecting target class thresholds.

Ayon & Islam (2019) [13] have proposed an approach for diagnosing diabetes using DNN by training attributes using 5 and 10 fold Cross-Validation (CV). The proposed scheme is applied to Pima Indian Diabetes (PID) dataset to predict diabetes based on numerous medical factors.

In the work proposed by Gadekallu et al (2020) [14], the dataset is normalized using scalar scheme to remove irrelevant attributes, inclusion of which increases the burden on ML model. Predominant features are extracted by applying Principal Component Analysis (PCA). Fire-Fly Algorithm (FFA) is applied for reducing dimensions. The reduced dataset is given to DNN model for classification.

Gadekallu et al (2020) [15] have used PCA-based DNN model using Grey Wolf Optimization (GWO) for classifying extracted features. GWO aids in choosing ideal parameters for training DNN model. The input is normalised followed by reduction of dimensionality using PCA. GWO is used for choosing optimal hyper parameters and training is performed using DNN.

Gupta et al (2022) [16] have proposed prediction models based on DL and Quantum Machine Learning (QML) schemes. Rejection of outliers, filling of missing values along with normalization is performed to improve performance of models. The models are trained by using PIMA dataset. Following pre-processing for cleaning of dataset, efficiency of techniques is analysed based on statistical parameters. QML is employed for predicting diabetes. Optimum amount of layers for both models are analysed using extended experimentation.

Mushtaq et al (2022) [17] have proposed a model for predicting and classifying diabetes based on glucose, Body Mass Index (BMI), age and insulin. Schemes like Tomek and SMOTE are used for data balancing. Further, outliers are removed. It is seen that Random Forest (RF) outperforms other classifiers. Voting classifier is also assessed.

3. CLASSIFICATION USING ML-BASED METHODS

In this section, ML-based methods of classification are discussed.

3.1 Support Vector Machine (SVM)

SVM [18], a supervised classifier is employed for regression and classification. It is applied for solving classification-based problems. It classifies Data Points (DPs) by using appropriate hyperplane in multi-dimensional space. Hyperplane is a decision boundary which classifies DPs. The hyperplane categorizes DPs with maximum margin amid classes and hyperplane.

The main aim of SVM algorithm is to determine a hyperplane in n-dimensional space which particularly categorizes DPs. The hyperplane's dimensions vary with quantity of features. In 2 input features are involved, then hyperplane will be a line. In case it is 3, then hyperplane is a 2-D plane. With more than 3 features, the complexity increases. A sensible choice is the best hyperplane which shows huge separation or margin amid 2 classes.

It is advisable to choose a hyperplane whose distance to adjoining DP on every side is more. Such hyperplanes are known as maximum-margin hyperplane or hard margin.

The kernel takes an input space with low-dimensions and converts it into a high-dimensional one. The non-separable problems are split into separable problems. It finds its use in separation problems which are non-linear in nature. For the kernel, complex data transformations are performed and data based on labels or outputs are defined.

Merits of SVM

- Efficient in handling high-dimensional cases.
- It makes efficient use of memory as it involves a subset of training Points (TPs) in Decision Function (DF) called support vectors.
- Diverse kernel functions may be specified for DFs and it is likely to mention custom kernels.

SVM takes labelled training data for every category and classifies text. In contrast to NNs, they have their own advantages: increased speed with improved performance with restricted amount of samples.

$$Minimize\ a_0 \ldots a_m : \sum_{j=1}^{n} \quad Max\left\{0,1-\left(\sum_{i=1}^{m} \quad a_i.x_{ij}+a_0\right)y_j\right\}+\ \lambda \sum_{i=1}^{m} \quad (a_i)^2 \quad (1)$$

Where

n - Quantity of data points

m - Quantity of attributes

x_{ij} - i^{th} attribute of j^{th} DP

$y_j = 1\ or - 1$ based on DP

3.2 Random Forest (RF)

Random Forest (RF) [19] classifier makes numerous Decision Trees (DTs) from arbitrarily chosen subgroup of training dataset. It aggregates votes from diverse DTs to choose the class of test objects.

This ensemble scheme is efficient in performing regression as well as classification using several DTs, and Bootstrap with Aggregation called bagging. Several DTs are combined for determining the output not depending on a single DT. In case of bootstrapping, row and feature samplings are performed, thus forming datasets for each model.

It aids in modelling predictions and performing behavior analysis by building on DTs. It includes several DTs indicating a separate occurrence of classification of input into RF. It considers each instance, the one with majority votes as prediction.

Every tree in classification considers input from samples in dataset. Features are arbitrarily chosen that are used for growing the tree in every node. Each tree must not be pruned till the end, wherein decisive prediction is reached. RF facilitates classifiers with weak correlations to build a strong classifier.

Merits of Random Forests

- RFs show estimates of variable significance, say neural nets.
- They offer a better method for handling missing data by substituting with the variable that appears several times in a specific node.
- RFs offer increased accuracy.
- It is capable of handling huge datasets owing to its ability to work with several variables.

- It can balance datasets with an infrequent class
- It is capable of handling variables quickly making it appropriate for complex tasks.

4. CLASSIFICATION OF DIABETES TYPES USING DL-BASED METHODS

In this section, DL-based methods of classification are discussed.

4.1 Long Short-Term Memory (LSTM)

LSTM is a kind of Recurrent Neural Network (RNN) [20]. Neurons are connected to previous states besides layer- wise inputs. RNNs are applied to sequential data or data that can evaluate contextual view. They are good in text categorisation [21]. They recall information with time. It faces several challenges when it operates with Gradient-based learning schemes. It is capable of handling increased duration involved in determining dependencies.

It includes a memory cell that stores information. It involves input, forget and output gates to handle memory. The units disseminate a substantial feature that appears at the beginning of input sequence over increased distance. It determines long-distance dependencies. Ensuing equations enable determination of time-step (t) for every hidden state:

Candidate layer:

$$I_t = \sigma(U_i.[H_{t-1}, Y_t] + B_i) \quad (2)$$

Forget Gate:

$$F_t = \sigma\left(U_f.[H_{t-1}, Y_t] + B_f\right) \quad (3)$$

Hidden state:

$$O_t = \sigma(U_o.[H_{t-1}, Y_t] + B_o) \quad (4)$$

Input Gate:

$$G_t = tanh(U_c.[H_{t-1}, Y_t] + B_c) \quad (5)$$

Output Gate:

$$C_t = F_t \odot C_{t-1} + I_t \odot G_t \quad (6)$$

Memory state:

$$H_t = O_t \odot tanh(C_t) \quad (7)$$

where,

H_t, H_{t-1} - Present and former hidden states

Y_t - Present LSTM input. Embedding of present word (U_t)

U and B - Weight vectors

$\odot$ - Element-wise multiplication

σ - Sigmoid function

4.2 Proposed Multi-Layer Neural Network (MLNN)

MLNN focuses on categorizing diverse kinds of diabetes by relating attributes offered by various datasets. The propounded scheme produces 4 sets of data by relating attributes.

The main aim is to reduce the number of attributes. The most suitable attribute is chosen depending on priority as well as frequency. After optimising features, classification is performed.

The proposed methods use Neural Network (NN) for classification which primarily includes 2 processes namely, training and testing phases. No-prop MLNN is used to categorize different kinds of diabetes. Once the weights of attributes are computed, attributes are involving in training 3 NNs.

In the training stage, initially the tested data fit into Normal (Type 1) NN. In case of normal data, it belongs to Normal (Type 2) NN. If it is normal again, data belongs to normal (gestational) diabetes. The framework of NN involves multiple layers providing better efficacy for classification.

Training Stage

The input layer of training stage includes 8 attributes from PID, while output layer includes 2 values for every NN. The following classes are assigned to different types of diabetes.

- **Class 1:** Normal without diabetes
- **Class 2:** Type 1
- **Class:** Type 2
- **Class 4:** Gestational

NN with 8 attributes is used for predicting Type 1 diabetes. Type 2 diabetes may be forecast using same network with pre-processed features. The main aim is to determine the general values of hidden layers. Layers are linked among themselves to establish a network. The transmitted signal activates the neuron and it includes an activation function which plays the role of a threshold. These functions aid in determining weights. The stability of the hidden layer is ensured when NN derives least error from the training dataset. Scaled Conjugate Gradient (SCG), a quick-supervised learning algorithm aids in reducing the error in the network. SCG aids the hidden layers to get stable values. Also, NNs are trained with associated dataset, and the values of the hidden layer are determined.

Testing Stage

In this stage, sequential testing is applied to classify data depending on diverse diabetes types. In this phase, MLNN is trained using known data of patients. The trained NN employs 3 dissimilar layers that are similar to neural training network with minimal variations. Input layer resembles training stage. Hidden layers pass through substantial variations as they get values from training stage.

In contrast, output layer is empty during testing stage. This stage focusses on classifying the unlabelled data into diverse groups depending on types of diabetes. This phase uses 3 NNs that are trained for diverse kinds of diabetes.

Firstly, unlabelled data is tested for type 1 diabetes by employing an initial NN. Data not listed under diabetes (normal) is chosen and given to second NN for determining likelihood of type 2 diabetes. Gestational diabetes is tested distinctly during pregnancy period as it is seen only in pregnant women.

5. RESULTS AND DISCUSSION

Diabetes involves several metabolic disorders wherein, the level of blood sugar remains high for a longer period. Symptoms include repeated urination, high thirst and hunger. It may lead to several problems when left untreated. Critical problems include diabetic ketoacidosis, Hyperosmolar Hyperglycemic State (HHS) or even death. Extended complications may include Cardio Vascular Disease (CVD), chronic kidney disease, stroke, foot ulcers and damage eyes.

To carry out investigations, dataset is taken from National Institute of Diabetes and Digestive and Kidney (NIDDK) diseases. The main goal is to forecast presence of diabetes depending on particular diagnostic measurements involved in the dataset. Numerous constraints are placed on selecting instances from huge database. All patients are females of age more than 21 years of Pima Indian heritage.

A ML model is designed to precisely envisage the presence of diabetes. The datasets includes numerous predictor and target variables.

The predictor variables include following:

- Pregnancies
- Blood Pressure
- Glucose
- Skin Thickness
- BMI
- Insulin
- DiabetesPedigreeFunction
- Age
- Outcome (0 or 1)

768 observation units are taken into consideration with 9 variables. Some features include ‘0’ which means missing value. It is replaced by NaN.

Categorical variables are converted into numerical values using Label and One Hot Encoding methods. Figure 2 shows the selection of parameters using Forward Elimination method.

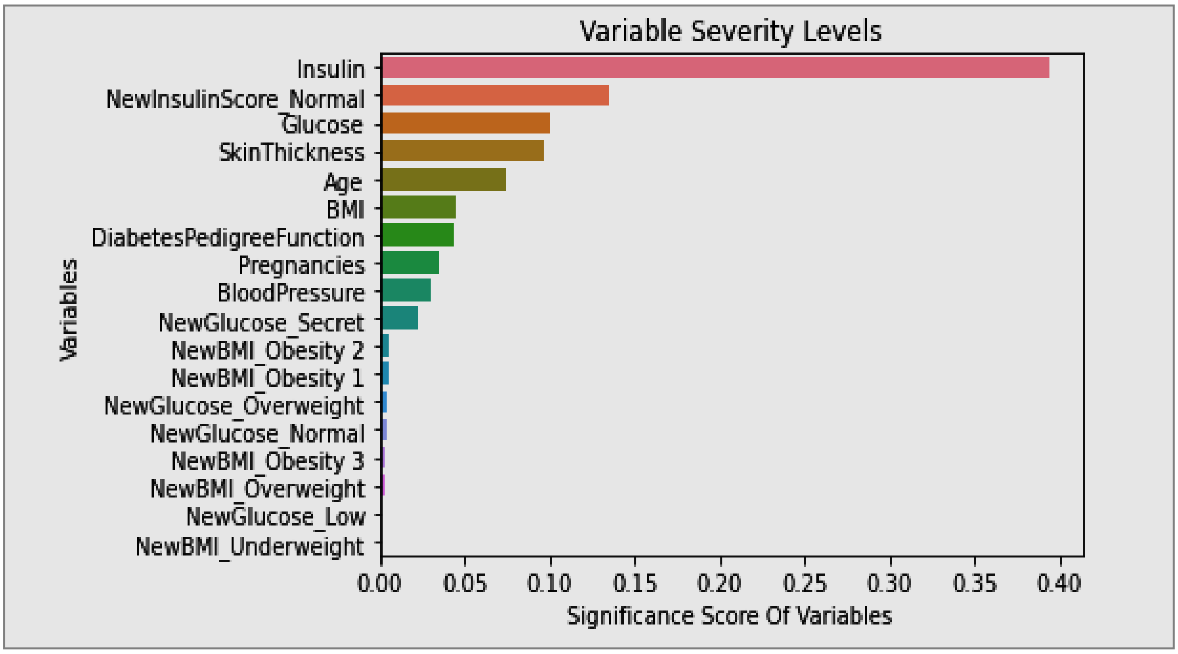

Figure 2: Selection of Parameters using Forward Elimination Method

Access to dataset correlation is provided. The relationship amid variables is also examined.

- **Positive correlation:** The value of correlation>0. The value of a variable increases with that of another
- **No correlation:** Correlation = 0
- **Negative correlation:** The value of correlation <0. The value of a variable increases while that of another decreases.

Figure 3 shows the Input and Figure 4 shows Confusion Matrix. Figure 5 shows the classification of dataset.

```
In [2]: data = pd.read_csv("diabetes.csv")
        print("Successfully Imported Data!")
        data.head()

        Successfully Imported Data!

Out[2]:
```

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1

Figure 3: Input Data

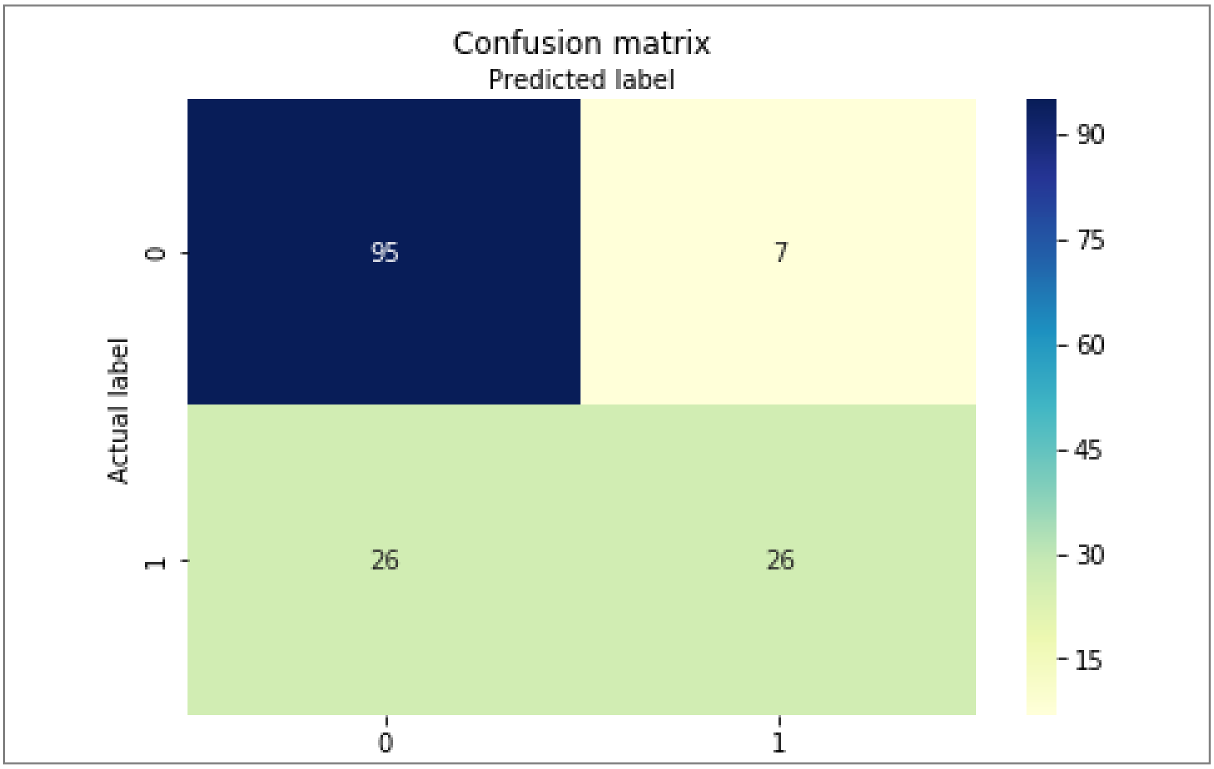

Figure 4: Confusion Matrix

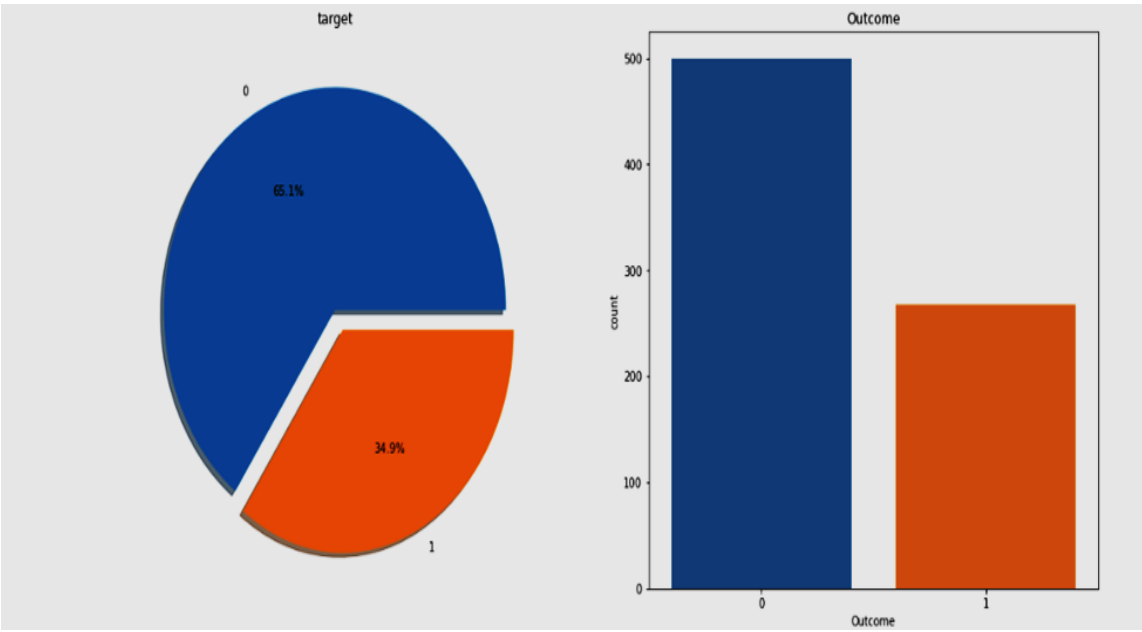

Figure 5: Classification of Dataset

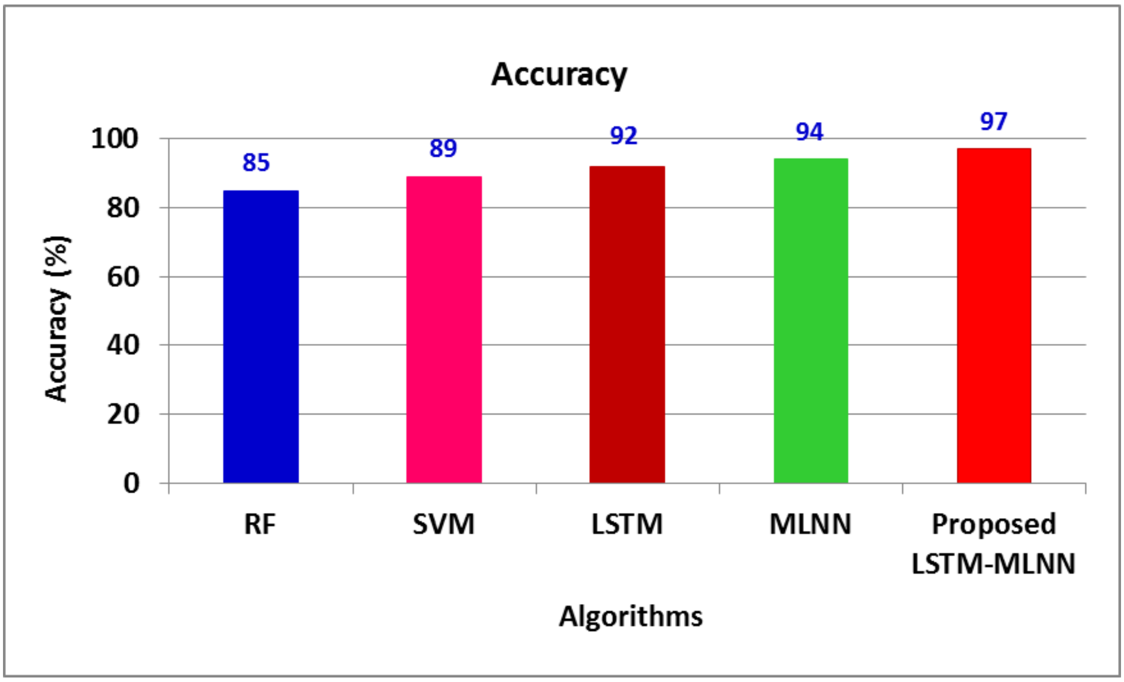

Figure 6: Accuracy

The Proposed LSTM-MLNN offers 12%, 8%, 5% and 3% better Accuracy in contrast to RF, SVM, LSTM and MLNN respectively (Figure 6).

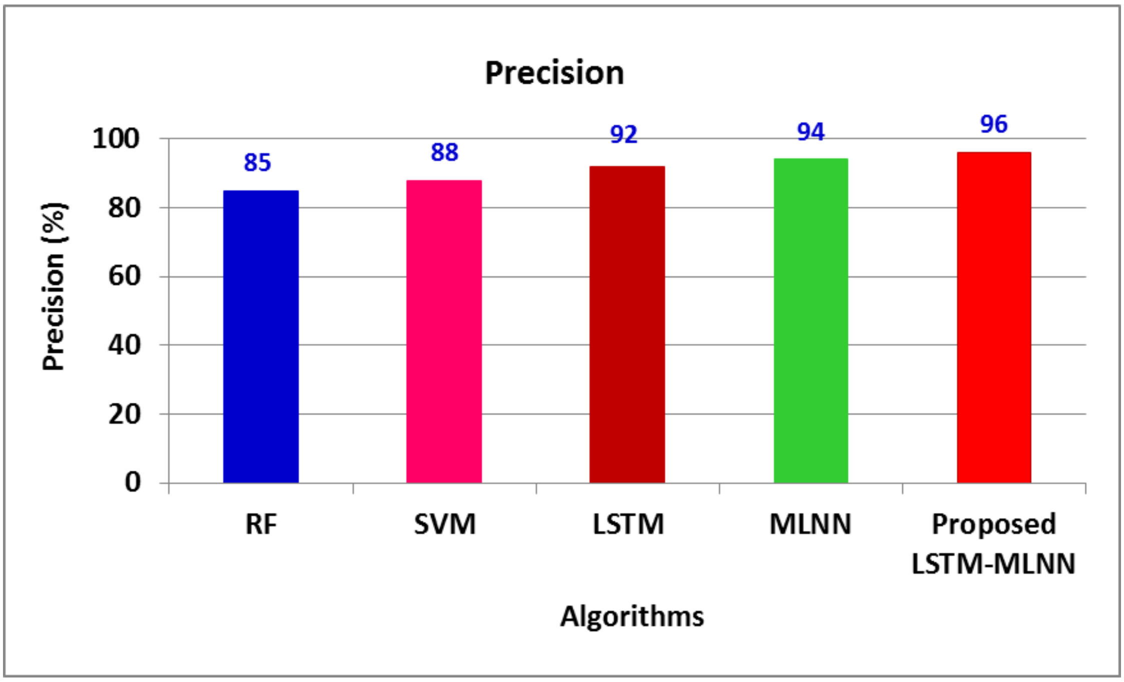

Figure 7: Precision

The Proposed LSTM-MLNN offers 11%, 8%, 4% and 2% better Precision when compared to RF, SVM, LSTM and MLNN respectively (Figure 7).

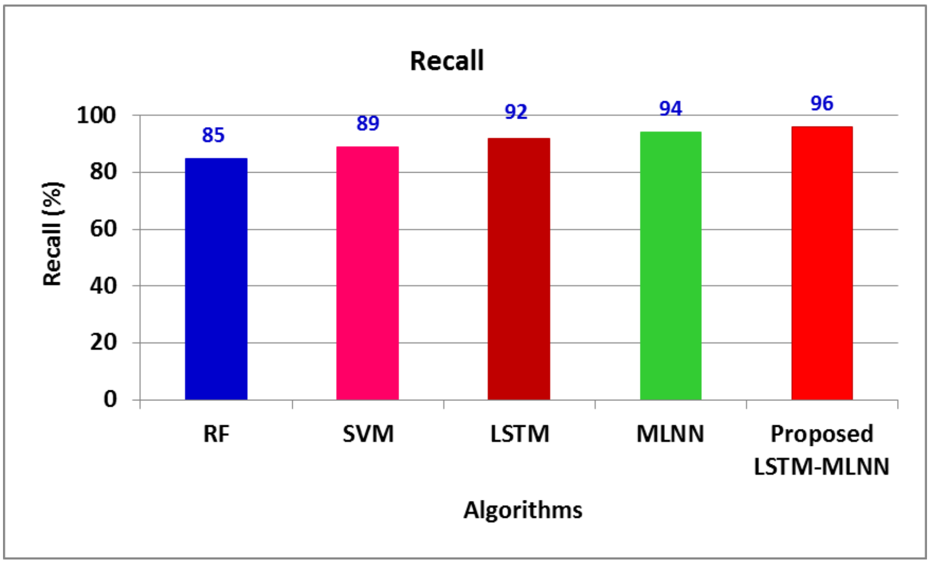

Figure 8: Recall

The Proposed LSTM-MLNN offers 11%, 7%, 4% and 2% better Recall in contrast to RF, SVM, LSTM and MLNN respectively (Figure 8).

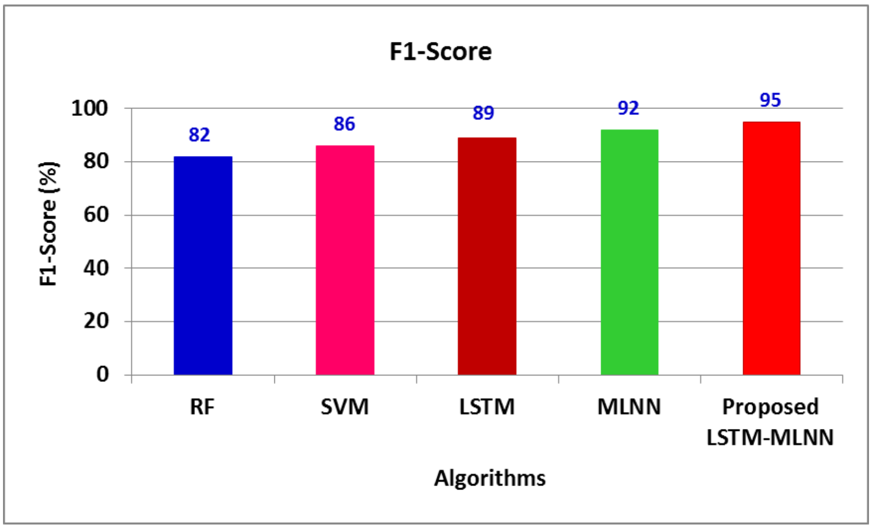

Figure 9: F1-Score

The Proposed LSTM-MLNN offers 13%, 9%, 6% and 3% better F1-Score when compared to RF, SVM, LSTM and MLNN respectively (Figure 9).

6. CONCLUSION

In this paper, Multi-Layer Neural Network (MLNN)-based scheme is proposed for predicting diabetes. The performance of ML and DL-based schemes are compared. It is seen that the proposed scheme offers improved Accuracy, Precision, Recall and F-Measure. in the future, a hybrid model is to be developed for accurately predicting diabetes at early stages.

REFERENCES

[1] Yahyaoui, A., Jamil, A., Rasheed, J., & Yesiltepe, M. (2019, November). A decision support system for diabetes prediction using machine learning and deep learning techniques. In 2019 1st International informatics and software engineering conference (UBMYK) (pp. 1-4). IEEE.

[2] Nadesh, R. K., & Arivuselvan, K. (2020). Type 2: diabetes mellitus prediction using deep neural networks classifier. International Journal of Cognitive Computing in Engineering, 1, 55-61.

[3] Varghese, N. R., & Gopan, N. R. (2020). Performance analysis of automated detection of diabetic retinopathy using machine learning and deep learning techniques. In Innovative Data Communication Technologies and Application: ICIDCA 2019 (pp. 156-164). Springer International Publishing.

[4] Chang, V., Bailey, J., Xu, Q. A., & Sun, Z. (2022). Pima Indians diabetes mellitus classification based on machine learning (ML) algorithms. Neural Computing and Applications, 1-17.

[5] Refat, M. A. R., Al Amin, M., Kaushal, C., Yeasmin, M. N., & Islam, M. K. (2021, October). A comparative analysis of early stage diabetes prediction using machine learning and deep learning approach. In 2021 6th International Conference on Signal Processing, Computing and Control (ISPCC) (pp. 654-659). IEEE.

[6] Das, D., Biswas, S. K., & Bandyopadhyay, S. (2022). A critical review on diagnosis of diabetic retinopathy using machine learning and deep learning. Multimedia Tools and Applications, 81(18), 25613-25655.

[7] Fregoso-Aparicio, L., Noguez, J., Montesinos, L., & García-García, J. A. (2021). Machine learning and deep learning predictive models for type 2 diabetes: a systematic review. Diabetology & Metabolic Syndrome, 13(1), 1-22.

[8] Mercaldo, F., Nardone, V., & Santone, A. (2017). Diabetes mellitus affected patients classification and diagnosis through machine learning techniques. Procedia computer science, 112, 2519-2528.

[9] Olisah, C. C., Smith, L., & Smith, M. (2022). Diabetes mellitus prediction and diagnosis from a data preprocessing and machine learning perspective. Computer Methods and Programs in Biomedicine, 220, 106773.

[10] Daanouni, O., Cherradi, B., & Tmiri, A. (2020). Type 2 diabetes mellitus prediction model based on machine learning approach. In Innovations in Smart Cities Applications Edition 3: The Proceedings of the 4th International Conference on Smart City Applications 4 (pp. 454-469). Springer International Publishing.

[11] Alex, S. A., Nayahi, J. J. V., Shine, H., & Gopirekha, V. (2022). Deep convolutional neural network for diabetes mellitus prediction. Neural Computing and Applications, 34(2), 1319-1327.

[12] Dutta, S., Manideep, B. C., Basha, S. M., Caytiles, R. D., & Iyengar, N. C. S. N. (2018). Classification of diabetic retinopathy images by using deep learning models. International Journal of Grid and Distributed Computing, 11(1), 89-106.

[13] Ayon, S. I., & Islam, M. M. (2019). Diabetes prediction: a deep learning approach. International Journal of Information Engineering and Electronic Business, 12(2), 21.

[14] Gadekallu, T. R., Khare, N., Bhattacharya, S., Singh, S., Maddikunta, P. K. R., Ra, I. H., & Alazab, M. (2020). Early detection of diabetic retinopathy using PCA-firefly based deep learning model. Electronics, 9(2), 274.

[15] Gadekallu, T. R., Khare, N., Bhattacharya, S., Singh, S., Maddikunta, P. K. R., & Srivastava, G. (2020). Deep neural networks to predict diabetic retinopathy. Journal of Ambient Intelligence and Humanized Computing, 1-14.

[16] Gupta, H., Varshney, H., Sharma, T. K., Pachauri, N., & Verma, O. P. (2022). Comparative performance analysis of quantum machine learning with deep learning for diabetes prediction. Complex & Intelligent Systems, 8(4), 3073-3087.

[17] Mushtaq, Z., Ramzan, M. F., Ali, S., Baseer, S., Samad, A., & Husnain, M. (2022). Voting classification-based diabetes mellitus prediction using hypertuned machine-learning techniques. Mobile Information Systems, 2022, 1-16.

[18] Noble, W. S. (2006). What is a support vector machine?. Nature biotechnology, 24(12), 1565-1567.

[19] Breiman, L. (2001). Random forests. Machine learning, 45, 5-32.

[20] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

[21] Yenter, A., & Verma, A. (2017, October). Deep CNN-LSTM with combined kernels from multiple branches for IMDb review sentiment analysis. In 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON) (pp. 540-546).