IJBPAS, October, 2026, 15(10): 4893-4909

ISSN: 2277–4998

International Journal of Biology, Pharmacy and Allied Sciences (IJBPAS)

'A Bridge Between Laboratory and Reader'

www.ijbpas.com

EFFICIENT WAY OF PREDICTING EARLY DIABETICS USING RECURRENT NEURAL DIABETICS PREDICTION (RNDP)THROUGH RECURSIVE FEATURE ELIMINATION WITH FISHER SCORE (RFEFS)

RENUGADEVI G[1], OSWALT MANOJ S[2], SASI KALA RANI K[3], KOUSIKA N[4]

[1,2,3,4]Department of Computer Science and Engineering, Sri Krishna College of Engineering and Technology, Anna University, Coimbatore, Tamil Nadu, India

[1,4]Sri Krishna College of Engineering and Technology, Coimbatore, Tamil Nadu, India

[2,3]Alliance School of Advanced Computing, Bengaluru, Karnataka, India

***Corresponding Author: Dr. Renugadevi G: E Mail: renu.mcett@gmail.com**

Received 25th Dec. 2025; Revised 26th Jan. 2026; Accepted 15th May 2026; Available online 1st Oct. 2026

https://doi.org/10.31032/IJBPAS/2026/15.10.10530

ABSTRACT

Early diabetes prediction is an important component of preventive healthcare because it allows for targeted medicines and lifestyle changes to reduce the risk of diabetes and its associated health concerns, such as neuropathy, kidney disease, and cardiovascular disease. It is a proactive method to controlling and enhancing public health outcomes associated with diabetes. Large datasets can be analysed using powerful computing techniques such as machine learning and deep learning to predict diabetes risk. These models might include a variety of variables, such as clinical data, lifestyle information, and genetic factors. In the proposed research, hybrid Recursive feature elimination with Fisher score (H-RFFS) method is a parameter collection technique is used to eliminate unwanted data form dataset and processed with prediction method. A deep learning model called Recurrent Neural Diabetics Prediction is utilised to predict early diabetics and their types. The input has been taken as sequential data and time-series prediction tasks. Using Recurrent Neural Diabetics Prediction to forecast early diabetes entails analysing sequential health data to estimate a person's risk of developing diabetes in the future. Early diabetic prediction is performed by using data set Mendeley Data Repository. Dataset that includes fasting blood glucose levels, BMI, HbA1c values, weight,

diabetes family history, food, exercise habits, and other health-related information. This model should provide improved performance matric such as Accuracy 90%, Precision 90%, Specificity 98 %and Sensitivity 88%.

Keywords: Recurrent Neural Diabetics Prediction, Neural Network, hybrid Recursive feature elimination, Fisher score method, dataset

1. INTRODUCTION

According to 2017 research published by the IDF, there were four hundred and twenty-five million diabetics worldwide. Additionally, research projected that by 2045, there will be 625 million. All of the endocrine illnesses that make up diabetes mellitus result from a relative or absolute lack of the substance known as "insulin" These disorders affect the body's ability to absorb glucose. All forms of metabolism are violated by the sickness, which has a chronic course. Diabetes will be separated into four essential categories: T1 diabetes, T2 diabetes, Ges diabetes mellitus, and a few kinds of diabetes originating from additional causes.

Diabetes is an extensive amount of glucose producing in our body also we can call it sugar, as your cells' key cause of liveliness. T1 diabetes and T2 diabetes are the 2 main forms of the illness. Individuals with diabetes need to take action to maintain their plasma sugar stages within a goal range as well as often check their levels. In order to avoid problems, which can impact the kidneys, eyes, nervous system, and cardiovascular system, it is imperative that diabetes be well managed.

Predicting the future progression and outcomes of diabetes in an individual can be challenging, as it depends on various issues, encompassing the diabetes kind (Type 1 or Type 2), genetic predisposition, lifestyle choices, and access to medical care. However, there are some general trends and risk factors that can help in making predictions regarding diabetes: Family History, Type of Diabetes, Body Weight and Lifestyle, Blood Sugar Levels, Complications, Medication and Treatment, Age.

Earlier notification of diabetes particularly Diabetes type 2, is crucial for preventive and proactive management. Early prediction is essential because it provides an opportunity for individuals should implement lifestyle modifications that can help postpone or prevent the onset of diabetes, such as eating a healthier diet and getting more exercise. For those at higher risk, healthcare providers may recommend more frequent monitoring and, in some cases, medication to manage blood sugar levels.

Early diabetes prediction using deep learning techniques involves leveraging machine learning models, particularly deep neural networks, to analyse and identify patterns in data that can be indicative of a person's risk of developing diabetes. Identify the most relevant features for diabetes estimate. While a large variety of features can be addressed by deep learning models, selecting the most useful ones can improve the performance of models and lower determining demands.

The multilayer perceptron (MLP), which is sometimes referred to as the Deep NN is a development of artificial observation. The i/p layer, unseen layers, and o/p layer are the three groups that DNN's layers fall under based on where they are located. The i/p stage is typically top of network, the o/p stage is typically the b/m layer, and the intermediate levels remain typically the concealed layers.

Consider using more advanced architectures like recurrent neural networks (RNNs) to capture sequential information in time-series data, which is relevant for diabetes prediction. The following are the primary findings of this work:

- The proposed model called RNDP (Recurrent Neural Diabetics Prediction), can predict not only if someone would develop diabetes in the future but also whether they will develop Type1D or Type2D of the condition.
- The model receives normalisation layers that enable it to memorise the assumptions drawn from the training set and address fresh information.
- During the training period, drops out randomly sets a fraction "G" of the given i/p values became ‘0’ with each update. Dropout regularization's main area remains to stop overfitting.
- The model makes adjustments to the hyperparameters.
- The window for compiling previous gradients has a fixed size. Recursively, the weakening regular of completely previous four-sided ramps is defined as the total of the gradients.
- At each stepin time t, the running average is then exclusively dependent upon the gradient that is now in effect
- The experimental findings demonstrate the efficacy and suitability of the suggested RNDP model; additionally, the diabetes's precision during training data sets for Pima Indians and Mendeley Diabetes are both optimal.

- **Figure 1** shows the selected features are given to Recursive neural model.

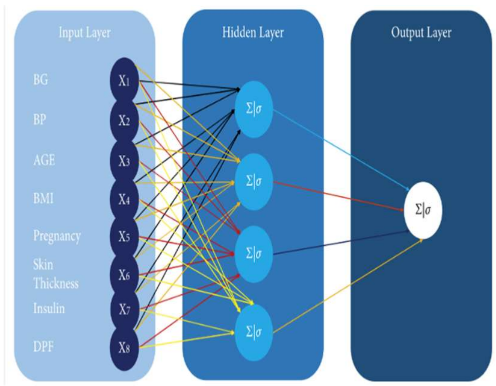

Figure 1: Recursive Neural Model for Diabetics Prediction

2. RELATED WORK

FNU Jyotsna *et al*., **[1]** Prognostic investigation is a procedure that practices historical and current data to uncover information and forecast upcoming events. It combines a number of ML algorithms, DM strategies, and statistical techniques. Furthermore, the article may discuss different strategies and interventions for managing cardiovascular risk in diabetic individuals. These approaches might encompass lifestyle modifications, pharmacological treatments, glycaemic control, blood pressure management, lipid-lowering therapies, and cardiovascular risk assessment tools.

Muhammad Exell *et al*. **[2]** the major goal of this work is to create prediction models utilising supervised machine learning approaches that can accurately forecast the occurrence or likelihood of diabetes in individuals. The importance of various characteristics or risk factors in predicting diabetes. This could include using feature selection approaches to determine the most important diabetes predictors. Contribute to the field of healthcare and medicine by establishing effective diabetes predictive models that can help healthcare practitioners with early diagnosis, prevention, and management.

Chun-Yang Chou. *et al* **[4]** In this paper goal is to apply machine learning techniques to healthcare difficulties, notably forecasting the onset of diabetes and enabling early intervention to enhance patient outcomes.

In this reference paper to find the possibilities and difficulties for DL based approaches in natural science and treatment have also been examined, in an effort to solve the difficulties associated with implementing DL techniques in healthcare today **[5]**. Li *et al*. **[6]** demonstrated Classical Center for DL lifecycle management,

Kim *et al*. **[7]** proposed a DN structure using Support Vector Machine to provide enough organizational deepness and full-bodied arrangement accuracy. To simulate various datasets has been taken from many resources.

Yousefi *et al*. **[8]** designed a novel prediction model that forecasts disease complications by employing a progressive concealing variables technique. Hammoudeh *et al*. **[9]** presented DL as an efficient method for forecasting clinic preadmissions in patient role with diabetes in their study.

Pham *et al*. **[10]** presented next DL-based approach to find health care progressions using the med-rec of patients. Bae *et al*. S. Bae *et al*,in this work may intend to create a predictive model that uses genetic polymorphisms to determine a person's likelihood of getting type 2 diabetes. This predictive model may help with early detection and care for people who are more likely to develop the condition **[11]**.

Zhu *et al*. **[13]** proposed a CNN algorithm to predict type 1 diabetic patients' future blood glucose levels. The model was an altered variant of WaveNet, which proved to be highly beneficial for processing acoustic signals.

Ryu, K.S.*et al*, The paper describes a deep learning model for estimating people who may have undiagnosed diabetes, an essential but often overlooked health condition with serious public health implications. The study's goal is to offer a reliable and efficient tool for detecting people at risk of undiagnosed diabetes, allowing for timely interventions and healthcare management **[14]**.

Isfafuzzaman Tasin *et al*, In this research, the authors are likely to investigate various machine learning algorithms and explainable AI methodologies in order to create prediction models capable of properly identifying persons at risk of getting diabetes. Machine learning algorithms enable the study of enormous datasets comprising a wide range of patient information, such as demographics, medical history, lifestyle factors, and biomarkers **[16]**.

Adnan Qayyum, Junaid Qadir, Muhammad Bilal, and Ala Al-Fuqaha **[18]** provides an overview of healthcare applications that use security and privacy approaches, along with their accompanying issues. We describe various ways to enable secure and privacy-preserving machine learning for healthcare applications. Finally, we discuss current research problems and potential directions for future study.

Hasan, M.K., *et al*, The research investigates the use of ensemble learning approaches for the prediction of diabetes, a common and serious chronic health condition. The authors study how combining predictions from various machine learning classifiers improves the accuracy and reliability of diabetes prediction models **[19]**.

Nadeem, M.W.; Goh, H.G.*et al* Machine learning algorithms for diabetes prediction. The design combines data from different sources to create a coherent dataset that better aligns with machine learning techniques. The methods use SVM and

ANN machine learning classifiers to detect and anticipate crucial events for people with disabilities. The proposed architecture achieves an accuracy of 84.67%, which is approximately 1.8% higher than the best-performing model in the literature **[20]**.

Olisah, C.C., *et al*, The research looks into the use of modern data preprocessing techniques and machine learning algorithms to predict and diagnose diabetes mellitus, a common and chronic metabolic condition with serious public health consequences. The study's goal is to improve the accuracy and efficiency of diabetes mellitus prediction and diagnosis by combining large datasets and cutting-edge machine learning approaches **[22]**.

Nesreen Samer El *et al*, The research looks into the use of artificial neural networks (ANNs) as a predictive tool for diabetes mellitus, a common chronic metabolic condition with serious public health consequences. The work intends to construct reliable and efficient diabetes prediction models using ANNs and comprehensive datasets, allowing for prompt treatments and personalised healthcare plans **[24]**.

However, the estimate procedures could expect individual the occurrence or absenteeism of the likelihood of taking the illness in the impending. In this proposed work, novel RNDP technique is to find diabetics types and what percentage of develop diabetics.

Gather relevant healthcare data, including medical records, patient history, and clinical measurements such as BG levels, BP, BMI, cholesterol stages, lifestyle factors. Ensure data privacy and compliance with relevant regulations (e.g., HIPAA in the United States). Preprocessing the given datasets to manage incomplete data, regularize the parameter and do change categorial var.

3. PROPOSED METHODOLOGY

Deep learning is used to forecast diabetes by analysing relevant data and making predictions about the likelihood of an individual developing diabetes in the future. Predicting early diabetes using deep learning and the Mendeley dataset entails using a dataset accessible on the Mendeley platform that has significant features for early diabetes detection. Search for diabetes-related datasets on the Mendeley platform or other sources like Kaggle, UCI Machine Learning Repository, or academic repositories.

Explore the dataset to understand its structure, features, and target variable(s). Ensure the dataset contains features that are indicative of early signs or risk factors associated with diabetes.

Collect a dataset with relevant features for early diabetes prediction. Make sure the dataset includes characteristics like glucose levels, BMI, age, family history, blood pressure, and other clinical markers. Preprocess the dataset by removing missing values, normalising numerical characteristics, and encoding categorical variables.

Use the Hybrid Recursive Feature Elimination with Fisher Score (H-RFFS) technique to determine the most useful features for diabetes prediction. H-RFFS uses the benefits of recursive feature elimination and Fisher score to rank and pick features based on their relevance to the target variable. This phase reduces the dimensionality of the feature space, which improves the prediction model's efficiency in **Figure 2**.

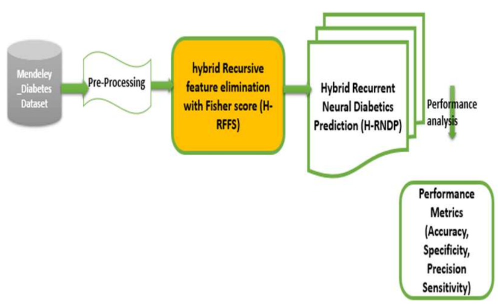

Figure 2: Proposed Architecture Model for Early Diabetics

Because RNNs are best suited for sequential data, divide the dataset into sequences or time-series data if necessary. For example, you may organise data by time periods or sequential observations for each individual. Divide the pre-processed data into training, validation, and testing sets. Train the RNN model with the training data and test its performance with the validation set. To monitor model performance and avoid overfitting, keep track of parameters like accuracy, precision, recall, F1-score, and ROC-AUC during training.

Use the test set to examine the trained model's ability to reliably forecast early diabetes.

4. IMPLEMENTATION

4.1. Data Set:

In proposed work there are two data sets has been taken to predict early diabetics. The dataset have many parameters. It have many Independent Var(IV) and one Target Var. Independent variables include the number of pregnancies the patient has had, their BMI_1, insulin level -1, age_1, and so on.

Another data set is taken from the Specialised Centre for Endocrinology and Diabetes-Al-Kindy Teaching Hospital and the Medical City Hospital laboratory provided the data, which were gathered from the Iraqi society. To create the diabetes dataset, patient files were obtained, data was extracted, and the information was arrived into the folder. The data are laboratory analyses and medical information. Data attributes consist of: Number of Patients, BG_1, Age_1 Gender_1, Cr_1, BMI_1, U, Chol_1, FP_1, including total, LDL_1, VLDL_1, TG_1 and HDL_1Cholest, HBA1C_1, Class_1 of diabetics(D,ND,PD)

4.2 Hybrid Recursive Feature Elimination with Fisher Score (RFFS) Method

The Hybrid Recursive Feature Elimination with Fisher Score (RFFS) method combines the Recursive Feature Elimination (RFE) technique and the Fisher Score for feature ranking. The RFE approach recursively removes the least significant features using a selected estimator, whereas the Fisher Score assesses each feature's discriminatory power.

Proposed novel feature selection method Recursive Feature Elimination with Fisher Score (RFFS), how works step-by-step overview:

Step1: **Calculate Fisher Score:** Fisher Score is a calculation of the discriminative control of a feature in distinctive between different classes in a classification problem. It is calculated using the following formula in equation 1:

Fisher Score (FS) = (mean of between - class variance_1) / (mean of within-class variance_1) ----------*(1)*

Step2: **Rank Features:** Calculate FS score for all parameters in the dataset and rank them in down order based on their scores. Higher FS Score indicate more discriminative power.

Step3: Select Features: Initially, select all the features.

Step4: Build a Classifier: Train a classifier using the selected features.

Step5: Evaluate Performance: Evaluate the performance of the classifier using a performance metric.

Step6: **Feature Elimination:** Eliminate the feature with the bottom FS Score from the selected parameter.

Step7: Repeat: Repeat steps 4 to 6 until you reach a predefined n of parameter or until the representation's presentation starts deteriorating. This process is recursive, as you iteratively eliminate the least important features.

Step 8: Final Model: After eliminating features, you are left with a subdivision of topographies that subsidize greatest to the representation's performance. You can then build your final model using this reduced feature set.

By iteratively eliminating the least important features, RFE with Fisher Score helps in simplifying the model, reducing dimensionality, and potentially improving model generalization. It is an effective way to discovery a balance among the n of features and the representation's performance, which can be crucial for avoiding overfitting and improving model interpretability.

4.2.1. Algorithm_RFFS:

```
# Step 1: Load the dataset containing features relevant to early diabetes prediction
X, y = load_diabetes_dataset ()
```

Step 2: Preprocess the dataset (e.g., handle missing values, scale features, encode categorical variables)

Step 3: Apply the Hybrid RFFS Algorithm

Use appropriate statistical measures (e.g., chi-squared) for calculating Fisher scores

Step 4: Train and evaluate machine learning models using the selected features

X_selected_features = hybrid_rffs(X, y, estimator, k)

Step 5: Split the dataset into training and testing sets

X_train, X_test, y_train, y_test = train_test_split(X_selected_features, y, test_size=0.2, random_state=42)

Step 6: Train machine learning models (e.g., logistic regression, decision trees, random forests) on the training set

model = train_model (X_train, y_train)

Step 7: Evaluate the trained model on the testing set

accuracy, precision, recall, f1_score, roc_auc = evaluate_model (model, X_test, y_test)

Step 8: Analyze the results and iterate if necessary

sns.heatmap(diab_data.corr(),cmap="YlGnBu")

plt. show ()

diab_data = diab_data. replace ({'CLASS': {'Y': 1, 'N': 0, 'P':1}})

The association between each support are pictured using heatmap. From the output, the flatboat flags indicate more association.

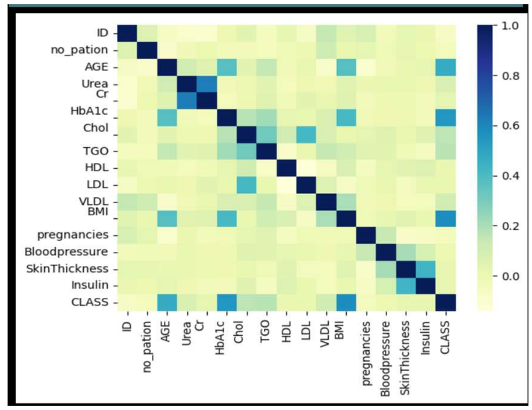

Figure 3: Heat Map for diabetic's prediction

We notice the correlation between pairs of features, like age_1 and pregnancies_1, or BMI_1 and skin thickness_1 etc. **Figure 3**.

4.2.2 Feature Selection:

Feature selection in deep learning model is done by inbuilt and selecting the importance features is major task to predicting early diabetics. In figure4 shows the importance from two data sets. Glucose and BMI have top important features. Based on glucose and BMI value early prediction model developed.

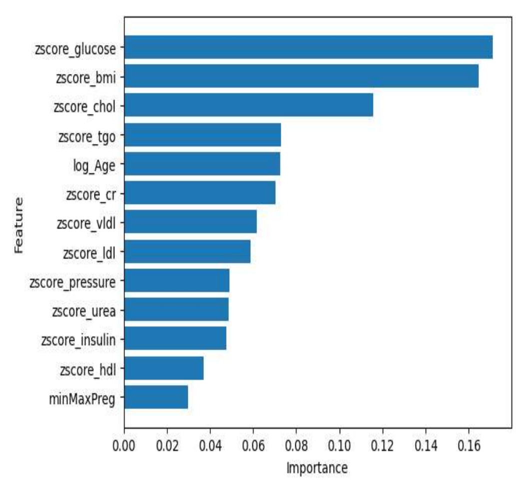

Figure 4: Importance Features Selection

The connection between each column is visualized using heatmap. From the output, the lighter colors specify additional association in **Figure 4**.

4.3. Building and Training the RNDP Model:

The RNDP model is made up of many decision neuron pools, each of which can be considered a tiny neural network. Each decision neuron pool makes independent decisions, and the final forecast is usually based on a majority vote or an average of predictions from all pools. The RNDP model combines elements of random forests and neural networks to create a hybrid model for classification tasks. The steps are:

- Create several decision neuron pools, each with its own architecture (e.g., number of layers, neurons per layer).
- Train each decision neuron pool independently using the training data.
- Training the RNDP Model: During training, each decision neuron pool learns independently from the data, sometimes utilising distinct subsets of features or samples.
- The forecasts of all decision neuron pools are combined using a voting or averaging process. Evaluation.
- Evaluate the RNDP model's performance using relevant metrics such as accuracy, precision, recall, F1-score, and ROC AUC.
- Use techniques like cross-validation to ensure a thorough examination of the model's performance.
- Experiment with hyperparameters, including decision neuron pool quantity, architecture, regularisation, and learning rates.
- To identify ideal hyperparameters, use approaches such as grid search and random search.
- Interpretability and Analysis: Understand how the RNDP model predicts early diabetes by analysing its decisions.
- Analyse the model's performance on various subsets of data to identify areas for improvement.

Assume for now that you have a data stream with T time steps. The features at a agreed time step r can be represented by the i/p vector X[r] for each period step r.

Concealed State: The RNDP keeps track of a concealed state vector, X[r], at each time step r. This vector is updated at each time step and contains information from prior time steps. Zeros or another value can be used to initialise the hidden state X[0].

Recurrent Computation: A recurrent formula that combines the i/p data at the present time stage and the previous concealed state is used to calculate the concealed state at each present stage. The method can be stated as in equation (2):

$X[r] = q(Z * Y[r] - M * X[r-1] + a)$-----(2)

Y is the weight matrix for the input data at time step r.

M is the Heavier Matrix for the previous concealed state X[r-1].

a is the bias term.

q is an Activation Function.

Final Computation: At each time step can be calculated based on the concealed state at that time step in equation (2):

$Y[r] = g(V * X[r] + c)$----------(3)

V is the Weigh Matrix for the Concealed state X[r].

c is the bias term.

g is an activation function that depends on the specific task, such as the SF function for binary classification or linear activation for regression.

RNDP model for early diabetes prediction

pseudo-code in Fig 5 demonstrates how to develop and train the RNDP model for early diabetes prediction. Here, need to write functions to load the dataset, combine predictions, calculate evaluation metrics, and maybe tune hyperparameters based on your individual needs and dataset peculiarities. Consider fine-tuning the architecture and hyperparameters to boost the model's performance.

```
from sklearn.ensemble import RandomForestClassifier
from sklearn.neural_network import MLPClassifier
from sklearn.ensemble import VotingClassifier
from sklearn.model_selection import train_test_split

# Step 1: Load and preprocess the dataset
X, y = load_diabetes_dataset()
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state

# Step 2: Build the RNDP model
decision_neuron_pools = []
for _ in range(num_pools):
    mlp = MLPClassifier(hidden_layer_sizes=(50, 50), activation='relu', solver='adam'
    rf = RandomForestClassifier(n_estimators=100, random_state=42)
    pool = VotingClassifier(estimators=[('mlp', mlp), ('rf', rf)], voting='soft')
    decision_neuron_pools.append(pool)

# Step 3: Train the RNDP model
for pool in decision_neuron_pools:
    pool.fit(X_train, y_train)

# Step 4: Combine predictions from all decision neuron pools
predictions = []
for pool in decision_neuron_pools:
    predictions.append(pool.predict(X_test))

# Step 5: Final prediction based on majority vote or averaging
final_prediction = majority_vote(predictions)  # Implement majority vote function

# Step 6: Evaluate the RNDP model
accuracy = calculate_accuracy(final_prediction, y_test)
precision, recall, f1_score = calculate_precision_recall_f1(final_prediction, y_test)
roc_auc = calculate_roc_auc(final_prediction, y_test)
```

Figure 5: Code for to develop and train the

6. RESULTS AND DISCUSSION

Diabetics prediction is classified by using DL algorithm and performance metrics are calculated using formulas. Predicting the occurrence of diabetes in its early stages is critical for effective intervention and control. Several performance criteria can be used to assess the efficacy of predictive models for early diabetes detection. Here are some common metrics:

Confusion Matrix:

The confusion matrix is an effective tool for assessing the effectiveness of classification models, especially those used to predict early diabetes. The confusion matrix is a table that summarises the model's actual and anticipated classifications. Here's how the confusion matrix is normally organised, Confusion matrix for this work is given in **Table 1**.

Table 1: Confusion Matix

	Predicted Negative (Non-Diabetic)	Predicted Positive (Early Diabetic)
Actual Negative (Non-Diabetic)	*TN*	*FP*
Actual Positive (Early Diabetic)	*FN*	*TP*

True Positive (TP): Instances that are actually positive (early diabetics) and are correctly predicted as positive.

False Positive (FP): Instances that are actually negative (non-diabetics) but are incorrectly predicted as positive.

False Negative (FN): Instances that are actually positive (early diabetics) but are incorrectly predicted as negative.

True Negative (TN): Instances that are actually negative (non-diabetics) and are correctly predicted as negative.

Accuracy: The accuracy is calculated by using confusion matrix. Addition of true positive and false positive. The formula to calculate accuracy is in equation (4):

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad \text{-----(4)}$$

Sensitivity: The sensitivity of a test is to find diabetics patients details. The amount of true positive inpatient cases. The formula to calculate sensitivity is in equation (5):

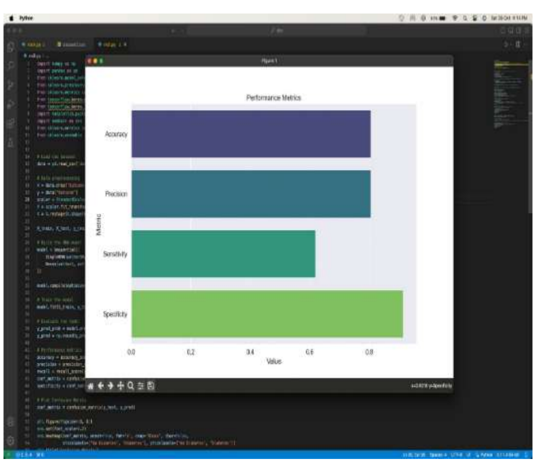

Figure 6

$$Sensitivity = \frac{TP}{TP + FN} \quad \text{--(5)}$$

Specificity: Specificity is used to calculate number non diabetics details. Formula to calculate specificity is in equation (6):

$$Specificity = \frac{TN}{TN + FP} \quad \text{--------(6)}$$

Precision Score Precision is the fraction of predicted positives/negatives events that are actually positive/negatives is in Equation(7).

$$Precision\ (positive\ class) = \frac{TP}{TP+FP} = \frac{True\ Positive}{Number\ of\ cases\ predicted\ as\ positive}$$

$$Precision\ (negative\ class) = \frac{TN}{TN+FN} = \frac{True\ Negative}{Number\ of\ cases\ predicted\ as\ negative}$$

---------(7)

Area around the receiver characteristic characteristic curve (ROC AUC): ROC AUC quantifies the model's ability to distinguish between early diabetics and non-diabetics at various threshold levels. It shows the genuine positive rate versus the false positive rate.

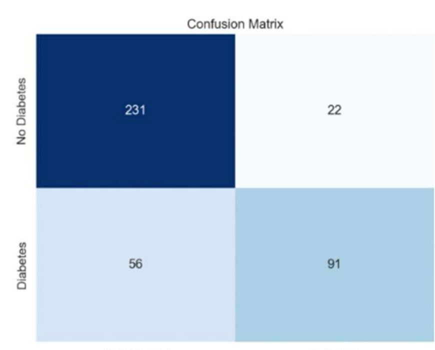

Figure 7: Confusion matix for Mendeley Daibetics data sets

A confusion matrix for the Mendeley Diabetes dataset in **Figure 7**. A confusion matrix is a certain table structure that enables for the visualisation of an algorithm's performance, which is often supervised learning. Each row of the matrix represents occurrences in a predicted class, and each column represents instances in an actual class (or vice versa).

Figure 8: Performance Matrix for early diabetic prediction

"Performance Matrix for early diabetic prediction," appears to relate to a matrix that assesses the performance of a model built for diabetes early detection. This matrix could include a variety of measures that are often used to evaluate the effectiveness of predictive models in **Figure 8**.

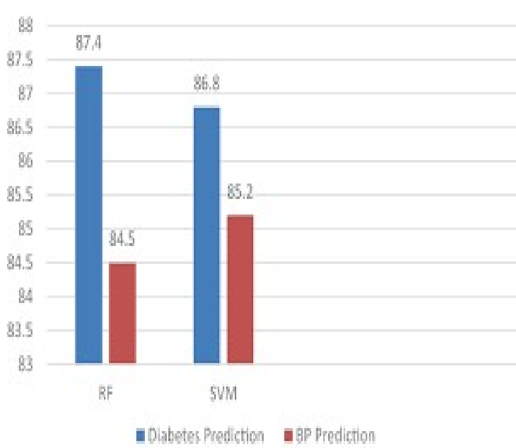

Figure 9: Result for early diabetes prediction using BP

Early detection of diabetes using blood pressure is an important field of study and application in healthcare and medical science. Blood pressure, together with other clinical factors, can predict the risk of developing diabetes or prediabetes in **Figure 9**.

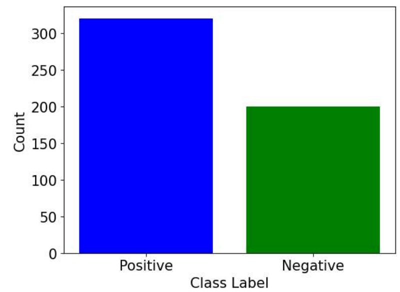

Figure 10: Class Label for Diabetics Positive or Negative

The confusion matrix contains a detailed breakdown of accurate and erroneous classifications, such as true positives, true negatives, false positives, and false negatives. These features, taken together, provide a thorough picture of the early diabetes prediction model's performance, allowing researchers and practitioners to evaluate its efficacy and make informed decisions regarding its value and future enhancements.

It is crucial to remember that, while blood pressure can be a valuable indicator, it is often used in conjunction with other clinical indicators and risk factors when predicting and managing diabetes. Integrated techniques that take into account various factors provide a more comprehensive assessment of an individual's risk profile, allowing for targeted interventions in diabetes prevention and management. In **Figure 10**. Class Label for Diabetics Positive or Negative.

V. CONCLUSION AND FUTURE WORK

Healthcare monitoring and prediction technology save many lives, particularly when patients are located far away. Early diagnosis of diabetes is crucial for selecting the best course of therapy. For the purpose of predicting diabetes type, a novel Recursive Neural Diabetics Prediction algorithm model (RNDP model) was presented in this work. Two datasets totalling more than a thousand records each were used to pretrained the deep neural network. Because there were few epochs used in the training phase, the approach can operate quickly on all mobile platforms. The efficacy is demonstrated by the experimental results. This model should provide improved performance matric such as Accuracy 90%, Precision 90%, Specificity 98 %and Sensitivity 88%. In the future, efforts can be long and upgraded for computerized diabetes investigation by improving performance metrics.

REFERENCES

[1] FNU Jyotsna, Areeba Ahmed, Kamal Kumar, Paramjeet Kaur, Mitul Hareshkumar Chaudhary, Sagar Kumar, Ejaz Khan, Bushra Khanam, Syeda Urooba Shah, Giustino Varrassi, Mahima Khatri, Satesh Kumar, Kishan Ashokbhai Kakadiya," Exploring the Complex Connection Between Diabetes and Cardiovascular Disease: Analyzing Approaches to Mitigate Cardiovascular Risk in Patients With Diabetes", Cureus 15(8): e43882 Article DOI: 10.7759/cureus.43882.(2023)

[2] Muhammad Exell, Fransiskus Xaverius, Gustian Paul Sendani a, Kristien Margi Suryanigrum a, Rezki Yunanda a,"Diabetes prediction using supervised machine learning ", Volume 216, 2023, Pages 21-30, Procedia Computer Science (2023), https://doi.org/10.1016/j.procs.2022.12.107.

[3] H. Liu, H. Kou, C. Yan, L. Qi, Keywords-driven and popularity-aware paper recommendation based on undirected paper citation graph. Complexity 2020, 2085638, 15 pages (2020).

[4] Chun-Yang Chou, Ding-Yang Hsu, Chun-Hung Chou," Predicting the Onset of Diabetes with Machine Learning Methods", J. Pers. Med. 2023, 13(3), 406; https://doi.org/10.3390/jpm13030406 (2023)

[5] T Gupta, A.; Katarya, R. Social media based surveillance systems for healthcare using machine learning: A systematic review. J. Biomed. Inform. 2020, 108, 103500.

[6] H. Miao, A. Li, S. Davis, A. Deshpande, in Proceedings- International Conference on Data Engineering. Model hub: deep learning lifecycle management (2017), pp. 1393–1394

[7] S. Kim, Z. Yu, R. Kil, M. Lee, Deep learning of support vector machines with class probability output networks. Neural Netw. 64,19–28 (2015)

[8] L. Yousefi, A. Tucker, M. Al-Luhaybi, L. Saachi, R. Bellazzi, L. Chiovato, in Proceedings- IEEE Symposium on Computer-Based Medical Systems. Predicting disease complications using a stepwise hidden variable approach for learning dynamic Bayesian networks (2018), pp. 106–111.

[9] A. Hammoudeh, G. Al-Naymat, I. Ghannam, N. Obied, Predicting hospital readmission among diabetics using deep learning. Procedia Comput. Sci. 141, 484–489 (2018)

[10] T. Pham, T. Tran, D. Phung, S. Venkatesh, Predicting healthcare trajectories from medical records: a deep learning approach. J. Biomed. Inform. 69, 218–229 (2017)

[11] S. Bae, T. Park, Risk prediction of type 2 diabetes using common and rare variants. Int. J. Data Min. Bioinform. 20(1), 77 - 90 (2018)

[12] K. Kannadasan, D.R. Edla, V. Kuppili, Type 2 diabetes data classification using stacked autoencoders in deep neural networks. Clin. Epidemiol. Glob. Health. 7(4), 530–535 (2018)

[13] Nadeem, M.W.; Goh, H.G.; Ponnusamy, V.; Andonovic, I.; Khan, M.A.; Hussain, M. A Fusion-Based Machine Learning Approach for the Prediction of the Onset of Diabetes. Healthcare 2021, 9, 1393.

[14] Ryu, K.S.; Lee, S.W.; Batbaatar, E.; Lee, J.W.; Choi, K.S.; Cha, H.S. A Deep Learning Model for Estimation of Patients with Undiagnosed Diabetes. Appl. Sci. 2020, 10, 421.

[15] S.P. Soniya, L. Singh, Application and need based architecture design of deep neural networks. Int. J. Pattern Recogn. Artif. Intell. 34(13), 2052014 (24 pages) (2020)

[16] Isfafuzzaman Tasin, Tansin Ullah Nabil, Sanjida Islam, Riasat Khan

Diabetes prediction using machine learning and explainable AI techniques , First published: 14 December 2022 https://doi.org/10.1049/htl2.12039 .

[17] Isfafuzzaman Tasin *et al*, In this research, the authors are likely to investigate various machine learning algorithms and explainable AI methodologies in order to create prediction models capable of properly identifying persons at risk of getting diabetes. Machine learning algorithms enable the study of enormous datasets comprising a wide range of patient information, such as demographics, medical history, lifestyle factors, and biomarkers.

[18] Mendeley Diabetics data set repository https://data.mendeley.com/datasets/wj9rwkp9c2/1 (Last updated in 2020)

[19] Akhtar, S., *et al*.: Prevalence of diabetes and pre-diabetes in Bangladesh: A systematic review and meta-analysis. BMJ Open 10, e036086 (2020)

[20] N Saranya, D Karthika Renuka and J Nikkesh kanthan Brain Tumor Classification Using Convolution Neural Network Journal of Physics: Conference Series Vol 2621 ISSN: 1742-6596

[21] Hasan, M.K., Alam, M.A., Das, D., Hossain, E., Hasan, M.: Diabetes prediction using ensembling of different machine learning classifiers. IEEE Access 8, 76516–76531, (2020)

[22] Muhammad Waqas Nadeem, Hock Guan Goh, Vasaki Ponnusamy, Ivan Andonovic, Muhammad Adnan Khan and Muzammil Hussain, "A Fusion-Based Machine Learning Approach for the Prediction of the Onset of Diabetes", Healthcare 2021, 9, 1393. https:// doi.org/10.3390/ healthcare 9101393.

[23] Cervantes, J., García-Lamont, F., Rodríguez, L., Lopez-Chau, A.: A comprehensive survey on support vector machine classification: Applications, challenges and trends. Neurocomputing 408, 189–215 (2020)

[24] Olisah, C.C., Smith, L., Smith, M.: Diabetes mellitus prediction and diagnosis from a data preprocessing and machine learning perspective. Comput. Methods Programs Biomed. 220, 1–12 (2022)

[25] Yu, L., Liu, L., Peace, K.E.: Regression multiple imputation for missing data analysis. Stat. Methods Med. Res. 29, 2647–2464 (2020)

[26] Nesreen Samer El Jerjawi, and Samy S. Abu-Naser," Diabetes Prediction Using Artificial Neural Network", International Journal of Advanced Science and Technology, Vol.121 (2018), pp. 55-64, https://dx.doi.org/10.14257/ijast.2018.121.05

[27] Jobeda Jamal Khanam, Simon Y. Foo "A comparison of machine learning algorithms for diabetes prediction", ICT Express, Volume 7, Issue 4, December 2021, Pages 432-439