ISSN: 2688-8645

Prediction and Analysis of Blood Glucose Levels based on Tabnet

Huazhong Yang

School of computer, Yangtze University, Jingzhou, Hubei, China

Abstract

Background: Blood glucose level prediction plays a significant role in the management of diabetes. Accurate prediction of blood glucose levels helps patients and doctors to make informed decisions regarding diet, exercise, and medication. The use of machine learning algorithms for blood glucose prediction has gained attention in recent years. Tabnet is one such algorithm that has shown promising results in various prediction tasks. Aim: The aim of this study is to evaluate the performance of Tabnet for blood glucose level prediction and compare it with other commonly used algorithms, including LR, DT, SVM, RF, and EN. Methods: A dataset of blood glucose levels of diabetic patients was used for this study. The dataset was preprocessed, and features were selected using correlation-based feature selection. Tabnet and other algorithms were trained on the dataset using 5-fold cross-validation. The performance of each algorithm was evaluated using root mean squared error (RMSE) and mean squared error (MSE). Results: The experimental results showed that Tabnet performed the best in terms of RMSE and MSE, with values of 0.5097 and 0.2523, respectively. The LR algorithm had an RMSE of 0.5126 and an MSE of 0.2629, while the DT algorithm had an RMSE of 0.7543 and an MSE of 0.5689. The SVM algorithm had an RMSE of 0.5165 and an MSE of 0.2663, while the RF algorithm had an RMSE of 0.5188 and an MSE of 0.2691. The EN algorithm had an RMSE of 0.5547 and an MSE of 0.3077. Conclusion: In this study, Tabnet was found to be the best algorithm for blood glucose level prediction compared to other commonly used algorithms. The results demonstrate the potential of Tabnet for predicting blood glucose levels in diabetic patients, which can assist in effective diabetes management.

Keywords

Blood Glucose Level; Machine Learning; Tabnet; Attentional Mechanisms.

1. Introduction

Blood glucose level prediction and analysis have become increasingly crucial in modern medicine. Diabetes, which is the most common metabolic disorder, affects approximately 463 million adults globally. The high prevalence of diabetes has made it necessary to develop techniques that can predict and analyze blood glucose levels accurately [1]. The ability to predict blood glucose levels can aid in early detection and management of diabetes, thereby reducing the risk of complications such as cardiovascular disease, neuropathy, and kidney failure. In this paper, we will investigate the use of Tabnet for predicting and analyzing blood glucose levels. The tabnet [2],[3] algorithm will be trained on a dataset of blood glucose measurements, and the resulting model will be used to predict the blood glucose levels of new patients. The performance of the model will be evaluated using various metrics, including mean absolute error, mean squared error, and correlation coefficient. Tabnet is a deep learning algorithm that has been successfully applied in many fields, including natural language processing, image recognition, and medical diagnosis. Tabnet is a unique algorithm that combines the strengths of decision trees and neural networks to achieve highly accurate predictions while also being highly interpretable.

ISSN: 2688-8645

The paper's significance lies in the potential for Tabnet to be used in clinical settings to predict and monitor blood glucose levels accurately. This technology can aid physicians in making informed decisions regarding diabetes management, leading to improved patient outcomes. Furthermore, the interpretability of Tabnet can help physicians understand the factors that influence blood glucose levels, leading to more targeted treatment plans. Overall, the goal of this paper is to explore the potential of Tabnet for predicting and analyzing blood glucose levels, which can aid in improving diabetes management and reducing the risk of complications associated with the condition. The key contributions of this study are as follows:

1. This paper proposes a deep learning algorithm called Tabnet with attention mechanism for predicting blood glucose levels.
2. Optimizing the hyperparameters of Tabnet for predicting blood glucose levels.
3. Data preprocessing techniques were used to process the blood glucose dataset, resulting in improved prediction accuracy.

2. Related Works

Blood glucose level prediction and analysis have been the subject of extensive research in recent years [4], with numerous approaches proposed for predicting and monitoring blood glucose levels. In this section, we will discuss some of the related works that have been proposed for blood glucose level prediction and analysis.

Traditional statistical models: Traditional statistical models such as regression and time series analysis have been widely used for blood glucose level prediction. These models use historical blood glucose data to make predictions about future blood glucose levels. However, these models are limited by their inability to capture the complex relationships between the variables that influence blood glucose levels. In 2020, Jinli He et al [5]. proposed a method for blood glucose concentration prediction called Kernel Canonical Correlation Analysis with Particle Swarm Optimization and Error Compensation. Their approach utilized only historical blood glucose data as input, instead of complex multi-dimensional inputs. Previous studies have shown that Canonical Correlation Analysis (CCA) can effectively predict blood glucose, but only considering the linear relationship between historical blood glucose values and predicted values is regrettable [6]. To address this limitation, the authors introduced a kernel function to identify the non-linear relationship between blood glucose. However, the introduced kernel function requires adjusting some parameters, and manual parameter adjustment can result in deviation. Therefore, this study explores the role of Particle Swarm Optimization (PSO) to reduce the deviation caused by manual parameter adjustment. Additionally, the authors proposed an error compensation for CCA to improve prediction accuracy. Finally, based on the prediction results of PSO-KCCA, a personalized hypoglycemic warning threshold was proposed, which considers the influence of model accuracy on prediction results. The experimental results confirmed the effectiveness of PSO-KCCA in blood glucose prediction, and the proposed EC-CCA successfully reduced delay in time series prediction. The personalized hypoglycemic warning threshold takes into account the influence of model accuracy on prediction results.

Deep learning models: Deep learning models such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) [7] have been proposed for blood glucose level prediction. These models have shown excellent performance in predicting blood glucose levels but are often limited by their lack of interpretability. In 2022, Mehrad Jaloli et al [8]. proposed a method for predicting blood glucose (BG) levels using stacks of convolutional neural networks and long short-term memory units. The method predicts BG levels for 30-, 60-, and 90-minute prediction horizons (PH) based on historical glucose measurements, meal information, and insulin intakes. The experimental results demonstrate that for the 90-minute PH, their model achieved a mean absolute error of 17.30 ± 2.07 and 18.23 ± 2.97 mg/dL, root mean square error

ISSN: 2688-8645

of 23.45 ± 3.18 and 25.12 ± 4.65 mg/dL, coefficient of determination of 84.13 ± 4.22% and 82.34 ± 4.54%, and in terms of continuous glucose-error grid analysis, 94.71 ± 3.89% and 91.71 ± 4.32% accurate predictions, 1.81 ± 1.06% and 2.51 ± 0.86% benign errors, and 3.47 ± 1.12% and 5.78 ± 1.72% erroneous predictions. The proposed method achieved better glucose prediction compared to existing methods in the literature and demonstrated potential for real-life applications due to its universality. To achieve the best prediction results, we propose using the Tabnet model with an attention mechanism to predict blood glucose levels. This method can accurately predict blood glucose levels with high precision and accuracy, providing a good reference for future blood glucose level predictions.

3. Materials and Methods

3.1. Dataset

The dataset comes from the Tianchi Precision Medicine Competition - AI-assisted Diabetes Genetic Risk Prediction. The dataset contains two files, the training file "d_train.csv" and the testing file "d_test.csv". The first line of each file is the field name, and each subsequent line represents an individual. The files contain a total of 42 fields, including numerical, character, and date types, and some fields may have missing data for some individuals. The first column represents the individual ID number. The last column of the training file is the label column, which is the target blood glucose value that needs to be predicted.

3.2. Data Preprocessing

Data preprocessing is an essential step in the process of building predictive models. It involves cleaning, transforming, and organizing raw data into a suitable format for analysis [9]. Data preprocessing is necessary because raw data often contains errors, missing values [10], and inconsistencies that can negatively affect the accuracy of predictive models. In the context of the blood glucose level prediction problem, data preprocessing is particularly important because the dataset contains multiple data types and missing values. Data cleaning involves identifying and correcting errors and inconsistencies in the dataset. Handling missing data involves dealing with missing values in the dataset. This may involve imputing missing values, dropping rows or columns with too many missing values, or using algorithms that can handle missing values directly.

Data normalization is a common preprocessing technique used to transform data into a standardized format [11]. It involves rescaling the values of numerical data to have a standard range, usually between 0 and 1 or -1 and 1, so that they can be more easily compared and analyzed. Standardizing data can help to improve the performance of predictive models by reducing the impact of variables with large numerical ranges. It also ensures that the variables are on a similar scale, which can help to avoid numerical instabilities in algorithms that are sensitive to the scale of the data. The formula for standardizing data is:

$$x' = \frac{\left(x - mean(x)\right)}{std(x)} \tag{1}$$

where x is the original value, mean(x) is the mean of the data, std(x) is the standard deviation of the data, and x' is the standardized value.

This formula calculates the z-score of each value, which represents the number of standard deviations that the value is away from the mean. Values with a z-score of 0 are at the mean, while values with a z-score of 1 or -1 are one standard deviation away from the mean in the positive or negative direction, respectively. Overall, data standardization is a valuable technique for preprocessing data in the blood glucose level prediction problem, as it can help to improve the accuracy and stability of predictive models.

ISSN: 2688-8645

3.3. Evaluation Metrics

In this article, we conducted comparative experiments using various algorithms, including multiple linear regression, decision tree, support vector machine, random forest, ElasticNet, and LightGBM. We evaluated the performance of these algorithms using the RMSE and MSE metrics. The formulas presented below represent the mathematical equations used to assess the performance of the algorithms. We will compare the performance of the algorithms based on these metrics.

RMSE (Root Mean Square Error) is a commonly used metric in regression analysis to measure the difference between predicted values and actual values [12]. It is calculated as the square root of the average of the squared differences between the predicted and actual values.

For a given predictive model with n data samples, where the predicted values are denoted as $\hat{y}_i$ and the actual values are denoted as y_i, the RMSE can be expressed as:

$$RSME = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \quad (2)$$

Here,$\sum y$ represents the square root,$\sum_{i=1}^{n} y$ denotes the sum of the squared errors for all n data samples, and $\frac{1}{n}$ represents the average value.

Mean Squared Error (MSE) is a commonly used metric in statistical analysis and machine learning for evaluating the quality of a model's predictions [13]. MSE measures the average squared difference between the predicted and actual values of a variable.

The formula for MSE is as follows:

$$MSE = \frac{1}{n} * \sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \quad (3)$$

where n is the number of data points, yi is the actual value of the variable, and ŷi is the predicted value. The calculation of MSE involves squaring the difference between the predicted and actual values, and then averaging these squared differences over all data points. The resulting value represents the average amount by which the predicted values deviate from the actual values, in terms of squared units. The significance of MSE is that it provides a quantitative measure of the accuracy of a model's predictions. A lower MSE indicates that the model is better at predicting the target variable, while a higher MSE suggests that the model's predictions are less accurate. In research and practical applications, MSE is often used as a benchmark for comparing the performance of different models or techniques. MSE is commonly used in statistical analysis and machine learning, and serves as a benchmark for comparing the performance of different models or techniques."

3.4. Multiple Linear Regression

Multiple linear regression is a statistical technique used to model the relationship between multiple independent variables and a dependent variable [14]. In other words, it is used to estimate how a dependent variable changes as the values of two or more independent variables change, while holding all other variables constant.

The primary purpose of multiple linear regression is to understand the relationships between variables and to make predictions based on these relationships. It can also be used to identify which independent variables are most strongly related to the dependent variable.

The formula for multiple linear regression is as follows:

$$y = \beta 0 + \beta 1x1 + \beta 2x2 + \ldots + \beta_n x_n + \varepsilon$$

ISSN: 2688-8645

where y is the dependent variable, β0 is the intercept, β1, β2, ..., βn are the coefficients of the independent variables x1, x2, ..., x_n , and ε is the error term. The coefficients represent the amount by which the dependent variable changes when the corresponding independent variable changes by one unit, while holding all other independent variables constant. The process of multiple linear regression involves estimating the values of the coefficients that best fit the data, using methods such as least squares regression. The resulting model can be used to predict the value of the dependent variable for new values of the independent variables.

The significance of multiple linear regression is that it allows us to analyze the complex relationships between multiple variables and to make predictions based on these relationships. It is widely used in fields such as economics, finance, and social sciences, where it is used to analyze the relationships between variables and to make forecasts based on these relationships.

3.5. Decision Tress

Decision tree regression is a machine learning algorithm used for modeling the relationship between a dependent variable and multiple independent variables [15]. It involves constructing a decision tree to model the relationship between the variables. The mathematical formula for decision tree regression involves recursively partitioning the data into subsets based on the independent variables, and then calculating the mean or median value of the dependent variable for each subset. This process is repeated until a stopping criterion is met, such as a minimum number of observations in each subset or a maximum depth of the tree.

The decision tree is constructed by selecting the independent variable that provides the best split at each level of the tree, based on a metric such as information gain or mean squared error. The resulting tree can be represented as a series of if-then statements that specify the decision rules for predicting the dependent variable based on the input variables. Let X1, X2, ..., Xp be the independent variables, and Y be the dependent variable. The decision tree regression formula can be written as:

1. Divide the data into subsets based on the independent variable that provides the best split, using a metric such as information gain or mean squared error.
2. Calculate the mean or median value of Y for each subset.
3. Repeat steps 1 and 2 for each subset, until a stopping criterion is met.
4. Construct the decision tree by selecting the independent variable that provides the best split at each level of the tree, based on the chosen metric.
5. Use the decision tree to predict the value of Y for new values of the independent variables, by following the decision rules in the tree.

3.6. Support Vector Machine

Support vector machine (SVM) is a machine learning algorithm used for classification and regression analysis [16]. It works by finding the hyperplane that best separates the data into different classes with a maximal margin between the classes. SVM works by transforming the input data into a high-dimensional feature space where the classes are more easily separable. Then it finds the hyperplane that separates the classes with a maximal margin, meaning the distance between the hyperplane and the closest data points of each class is as large as possible. SVM can handle both linearly separable and non-linearly separable data. In the case of non-linearly separable data, SVM uses the kernel trick to map the input data into a higher-dimensional space where the classes are separable by a hyperplane.

3.7. Random Forest

Random Forest is a widely used ensemble learning method that combines multiple decision trees to form a strong and robust predictive model [17]. Its popularity is due to its ability to handle high-dimensional data with complex interactions, as well as its resistance to overfitting

ISSN: 2688-8645

and noise. The core idea of Random Forest is to generate a forest of decision trees by randomly selecting subsets of the data and features for each tree. This approach reduces the correlation between individual trees, thereby improving the overall accuracy and robustness of the model. The final prediction is obtained by averaging the predictions of all trees in the forest.

Suppose we have a training set consisting of n observations and p predictor variables, denoted by (x1, y1), (x2, y2), ..., (xn, yn), where xi = (xi1, xi2, ..., xip) is a vector of predictor variables, and yi is the corresponding response variable. The goal of Random Forest regression is to build a model that can predict the response variable y for a new observation x.

The Random Forest algorithm can be broken down into the following steps:

For each tree in the forest:

a. Draw a random sample of n observations with replacement from the training set (bootstrapping).

b. Draw a random sample of m features from the p available features, where m << p.

c. Grow a decision tree using the bootstrapped sample and selected features.

To make a prediction for a new observation x:

a. Apply each decision tree in the forest to x.

b. Average the predictions across all trees to obtain the final prediction.

To compute the prediction of a single decision tree, we first split the training data into two regions R1 and R2, based on a splitting rule that maximizes the reduction in variance of the response variable y. Specifically, we minimize the following sum of squared errors:

$$SSE=\sum_{i\in R1}(yi-\widehat{y_1})^2+\sum_{i\in R2}(yi-\widehat{y_2})^2$$

where $\widehat{y_1}$ and $\widehat{y_2}$ are the mean response values in regions R1 and R2, respectively.

3.8. ElasticNet

ElasticNet is a linear model used for regression analysis that aims to find the best model coefficients by simultaneously considering both L1 and L2 regularization [18]. It can be viewed as a compromise between Lasso regression and Ridge regression, allowing for a balance between the effects of L1 and L2 regularization by controlling the weights between the regularization terms. In high-dimensional datasets, ElasticNet can reduce unnecessary variables in the model and improve the predictive power of the model. In ElasticNet, the loss function is defined as the sum of squared error between the predicted value and the actual value, plus the L1 and L2 penalties of the coefficient values. The optimization problem is then formulated as finding the coefficient values that minimize the loss function, subject to the constraints imposed by the L1 and L2 penalties. The optimization problem can be solved using various numerical optimization algorithms, such as coordinate descent or gradient descent.

3.9. Tabnet

Tabnet is an interpretable neural network based on self-attention mechanism, which is used for classification and regression tasks of structured data [19]. Compared to traditional neural networks, Tabnet provides interpretability and visualization capabilities while maintaining high predictive performance through special sparsity constraints and feature selection mechanisms. The main feature of Tabnet is the adoption of self-attention mechanism, which can learn the correlation between different features and also has some interpretability. It also reduces unnecessary features, improves model generalization ability and robustness through sparsity constraints and feature selection mechanism. Tabnet also introduces a mechanism similar to boosting, by training multiple models at each stage and integrating them to further improve the model performance. During the training process, Tabnet learns the correlation between different features through alternating feature selection and feature reconstruction stages, and uses global attention mechanism to combine information from different features in

ISSN: 2688-8645

the final prediction stage [20]. In addition, Tabnet can provide the importance ranking of each feature in the model, providing support for the interpretability of the model. The architecture diagram of Tabnet can show the above features, Figure 1 shows the encoder architecture of Tabnet.

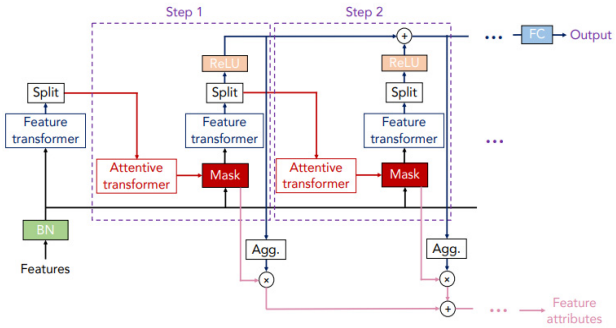

Figure 1. Tabnet encoder architecture

4. Experimental Results

In this study, the best prediction results were obtained by preprocessing the data, and the correlation heatmap showed the correlation between features. The larger the value, the more correlated the feature is with the target feature, and the more it affects the algorithm's performance. Figure 3 shows the correlation heatmap.

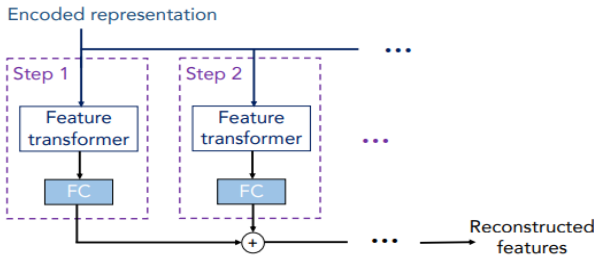

Figure 2. Shows the Tabnet decoder

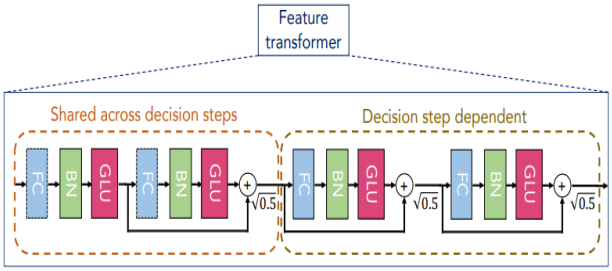

Figure 3. Shows a feature transformer block example – 4-layer network is shown, where 2 are shared across all decision steps and 2 are decision step-dependent. Each layer is composed of a fully-connected (FC) layer, BN and GLU nonlinearity

In this study, we propose an algorithm (Tabnet) that achieves the best prediction results with an RMSE of 0.5197 and an MSE of 0.2601, compared to other algorithms. In Figure 4, we compared the performance of five machine learning algorithms for predicting blood glucose levels with our proposed algorithm. The experimental results showed that our proposed algorithm had the best predictive performance, with an RMSE of 0.5213 and MSE of 0.2601. The Tabnet algorithm with attention mechanism was able to better fit the predicted results.

ISSN: 2688-8645

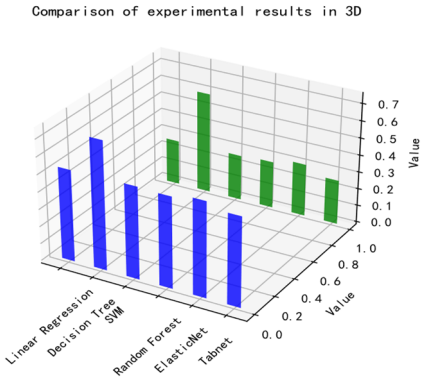

Figure 4. Tabnet and other 5 algorithms graph

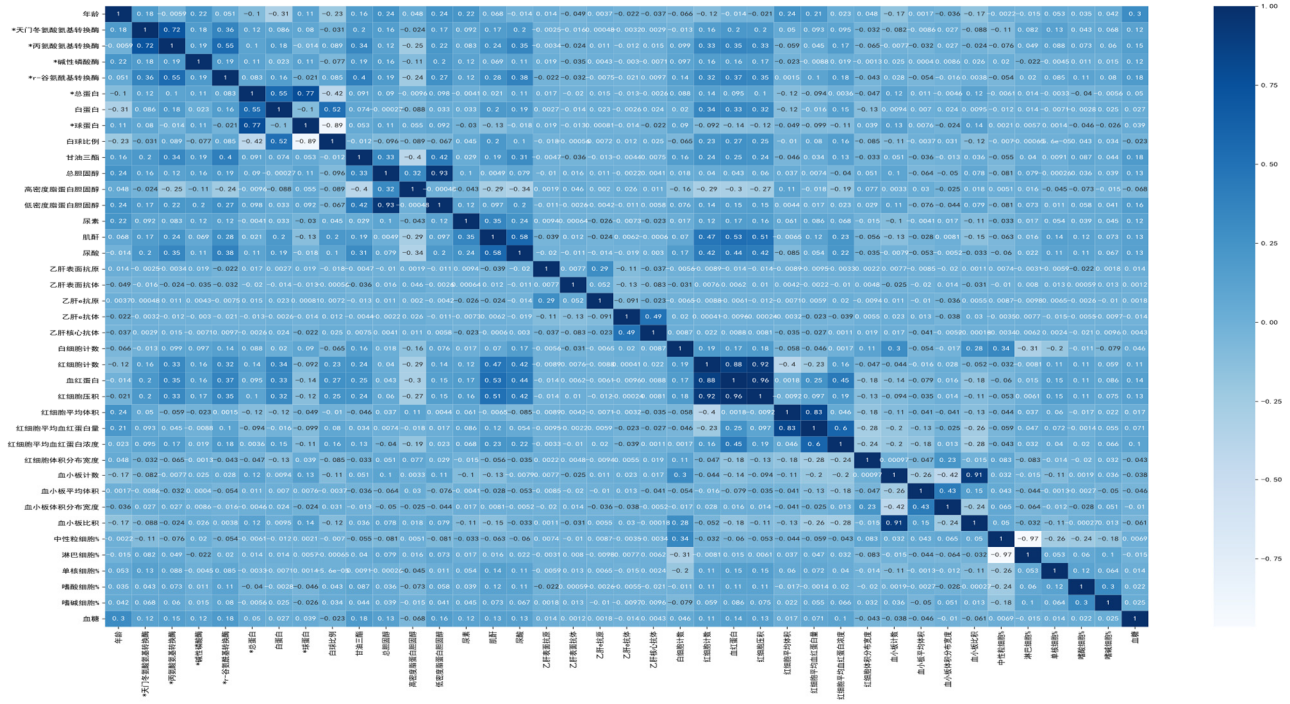

Figure 5. Heat matrix plot of the correlation of the eigenvalues

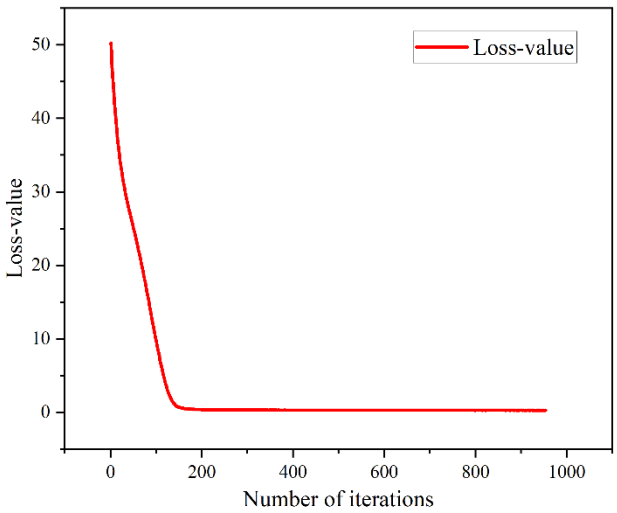

Figure 6. Loss function graph

ISSN: 2688-8645

Through the use of Tabnet for predicting blood glucose levels, a smaller value of the loss function indicates more accurate predictions. In our study, we set the number of iterations to 1000, and through experimentation, we determined that the optimal number of training iterations was 453, with early stopping at 953 iterations, at which point overfitting had occurred. Therefore, we selected the best training iteration of 453. For the validation set, the best RMSE was 0.5097, as shown in Figure 5. The loss function curve also indicates the good performance of the prediction results with a relatively small number of training iterations, resulting in precise predictions of blood glucose levels and cost savings in training. These findings demonstrate the practical value of this method in real-world applications.

Table 1 displays the outcomes of predicting blood glucose levels using Tabnet and five other machine learning algorithms. The table illustrates that Tabnet attains the most optimal prediction results, with linear regression as the second-best. Among the five selected machine learning algorithms, the decision tree algorithm presents the poorest performance in predicting blood glucose levels. Therefore, we have opted to employ the Tabnet algorithm for predicting blood glucose levels due to its remarkable prediction accuracy and robustness, offering a promising approach for medical blood glucose level prediction.

5. Conclusion

Effective management of blood glucose levels is crucial for individuals with diabetes, as it can prevent serious health complications. Machine learning algorithms have shown great potential in accurately predicting and analyzing blood glucose levels, thus helping clinicians make informed decisions and providing patients with more personalized care. In this study, we compared the performance of several machine learning algorithms, including Tabnet, LR, DT, SVM, RF, and EN, in predicting and analyzing blood glucose levels.

Table 1. Tabnet and the experimental results of five other algorithms

Model	LR	DT	SVM	RF	EN	Tabnet
RMSE	0.5126	0.7543	0.5165	0.5188	0.5547	0.5097
MSE	0.2629	0.5689	0.2663	0.2691	0.3077	0.2523

In conclusion, the Tabnet algorithm has demonstrated superior performance in the prediction and analysis of blood glucose levels compared to other commonly used machine learning algorithms, including LR, DT, SVM, RF, and EN. The Tabnet algorithm achieved an RMSE of 0.5097 and an MSE of 0.2523, which are significantly lower than the RMSE and MSE values obtained by the other algorithms. This study highlights the potential of Tabnet as an effective tool for predicting and analyzing blood glucose levels, which could have important clinical implications for patients with diabetes or other conditions that require careful monitoring of blood glucose levels. Further research is needed to explore the generalizability of these findings across different populations and to optimize the performance of Tabnet for specific applications. Overall, this study underscores the value of machine learning algorithms in healthcare and the importance of continued efforts to develop and refine these tools.

Acknowledgments

Natural Science Foundation.

References

[1] Ali, Jaouher Ben, et al. "Continuous blood glucose level prediction of type 1 diabetes based on artificial neural network." Biocybernetics and Biomedical Engineering 38.4 (2018): 828-840.

[2] Arik, Sercan Ö., and Tomas Pfister. "Tabnet: Attentive interpretable tabular learning." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 35. No. 8. 2021.

[3] Shah, Chiranjibi, Qian Du, and Yan Xu. "Enhanced Tabnet: Attentive interpretable tabular learning for hyperspectral image classification." Remote Sensing 14.3 (2022): 716.

[4] DECODE Study Group, and European Diabetes Epidemiology Group. "Glucose tolerance and cardiovascular mortality: comparison of fasting and 2-hour diagnostic criteria." Archives of internal medicine 161.3 (2001): 397-405.

[5] He, Jinli, and Youqing Wang. "Blood glucose concentration prediction based on kernel canonical correlation analysis with particle swarm optimization and error compensation." Computer Methods and Programs in Biomedicine 196 (2020): 105574.

[6] He, Jinli, and Youqing Wang. "Blood glucose concentration prediction based on kernel canonical correlation analysis with particle swarm optimization and error compensation." Computer Methods and Programs in Biomedicine 196 (2020): 105574.

[7] Banerjee, Imon, et al. "Comparative effectiveness of convolutional neural network (CNN) and recurrent neural network (RNN) architectures for radiology text report classification." Artificial intelligence in medicine 97 (2019): 79-88.

[8] Jaloli, Mehrad, and Marzia Cescon. "Long-term prediction of blood glucose levels in type 1 diabetes using a cnn-lstm-based deep neural network." Journal of Diabetes Science and Technology (2022): 19322968221092785.

[9] Abdallah, Zahraa Said, Lan Du, and Geoffrey I. Webb. "Data Preparation." (2017): 318-327.

[10] Liu, Yuehua, et al. "Missing value imputation for industrial IoT sensor data with large gaps." IEEE Internet of Things Journal 7.8 (2020): 6855-6867.

[11] Liu, Xueyan, et al. "Normalization methods for the analysis of unbalanced transcriptome data: a review." Frontiers in bioengineering and biotechnology 7 (2019): 358.

[12] Chicco, Davide, Matthijs J. Warrens, and Giuseppe Jurman. "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation." PeerJ Computer Science 7 (2021): e623.

[13] Hodson, Timothy O. "Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not." Geoscientific Model Development 15.14 (2022): 5481-5487.

[14] Rath, Smita, Alakananda Tripathy, and Alok Ranjan Tripathy. "Prediction of new active cases of coronavirus disease (COVID-19) pandemic using multiple linear regression model." Diabetes & Metabolic Syndrome: Clinical Research & Reviews 14.5 (2020): 1467-1474.

[15] Acharya, Mohan S., Asfia Armaan, and Aneeta S. Antony. "A comparison of regression models for prediction of graduate admissions." 2019 international conference on computational intelligence in data science (ICCIDS). IEEE, 2019.

[16] Ghosh, Sourish, Anasuya Dasgupta, and Aleena Swetapadma. "A study on support vector machine based linear and non-linear pattern classification." 2019 International Conference on Intelligent Sustainable Systems (ICISS). IEEE, 2019.

[17] Sun, Yuantian, et al. "Development of ensemble learning models to evaluate the strength of coal-grout materials." International Journal of Mining Science and Technology 31.2 (2021): 153-162.

[18] Hayslep, Matthew, et al. "Understanding district metered area level leakage using explainable machine learning." IOP Conference Series: Earth and Environmental Science. Vol. 1136. No. 1. IOP Publishing, 2023.

[19] Gavito, Andrea Treviño, Diego Klabjan, and Jean Utke. "Multi-Layer Attention-Based Explainability via Transformers for Tabular Data." arXiv preprint arXiv:2302.14278 (2023).

[20] Wang, Qinglun, et al. "GTFD-XTNet: A tabular learning-based ensemble approach for short-term prediction of photovoltaic power." IET Renewable Power Generation 16.12 (2022): 2682-2693.