Development of a Machine Learning Model to Predict Blood Glucose Levels

Desarrollo de un Modelo Basado en Aprendizaje Automático para Predecir los Niveles de Glucosa en Sangre

DOI: https://dx.doi.org/10.17981/ingecuc.21.2.2025.15

Artículo de Investigación Científica.

Fecha de Recepción: 03/10/2025, Fecha de Aceptación: 09/12/2025.

Omar Andres Aguirre-Aguirre
Universidad Nacional de Colombia, Sede de La Paz. La Paz, Colombia
oaguirre@unal.edu.co

Eduardo Nicolas Guevara-Povea
Universidad Nacional de Colombia, Sede de La Paz. La Paz, Colombia
edguevarap@unal.edu.co

Juan Pablo Hoyos-Sanchez
Universidad Nacional de Colombia, Sede de La Paz. La Paz, Colombia
jhoyoss@unal.edu.co

To cite this paper
O. Aguirre-Aguirre, E. Guevara-Povea & J. Hoyos-Sanchez "Development of a Machine Learning Model to Predict Blood Glucose Levels," INGE CUC, vol. 21, no. 2, 2025. DOI: https://dx.doi.org/10.17981/ingecuc.21.2.2025.15

Abstract

Accurate monitoring of blood glucose levels is essential for the effective management of type 1 diabetes, especially given the complexity and individual variability of this disease. Therefore, to prevent accidents and/or alterations, it is necessary to develop machine learning models capable of predicting glucose levels in the short term (1 hour). This paper proposes a solution following a six-step research methodology: data collection and processing, model training and selection, evaluation metrics, validation, prediction visualization, and application of grid search. The algorithms used were linear regression, decision trees, Random Forest, XGBoost, and TabNet. The dataset included continuous measurements of glucose, insulin administration, carbohydrate intake, and physical activity levels recorded using portable devices.
The models were evaluated using standard metrics (MAE, MAPE, RMSE, R^2, EVR) on the training and test sets. The results showed that the best predictive model, Random Forest, had an RMSE of only 0.75 and an R^2 coefficient of determination of 0.94, evidencing high precision and generalization capacity. It was also found that the manual adjustment of hyperparameters exceeded the grid search.

Keywords: Machine learning; type 1 diabetes; models; prediction; glucose.

Resumen

Un monitoreo preciso de los niveles de glucosa en sangre es fundamental para el manejo efectivo de la diabetes tipo 1, especialmente debido a la complejidad y variabilidad individual de esta enfermedad. Por tanto, para prevenir accidentes y/o alternaciones es necesario desarrollar modelos de aprendizaje automático capaces de predecir a corto plazo (1hora) los niveles de glucosa. En este artículo se propone una solución siguiendo una metodología de investigación que consta de 6 pasos: recolección y procesamiento de datos, entrenamiento y selección del modelo, métricas de evaluación, validación, visualización de las predicciones y aplicación de grid search. Los algoritmos usados fueron regresión lineal, árboles de decisión, Random Forest, XGBoost y TabNet. El dataset incluyó mediciones continuas de glucosa, administración de insulina, ingesta de carbohidratos y niveles de actividad física registrados mediante dispositivos portátiles.
Los modelos fueron evaluados mediante métricas estándar (MAE, MAPE, RMSE, R^2, EVR) sobre los conjuntos de entrenamiento y prueba. Los resultados mostraron que el mejor modelo predictor, Random Forest, tuvo un RMSE de solo 0.75 y un coeficiente de determinación R^2 de 0.94, evidenciando alta precisión y capacidad de generalización. También se encontró que el ajuste manual de hiperparámetros superó al grid search.

Palabras clave: Aprendizaje automático, diabetes tipo 1, modelos, predicción, glucosa.

INTRODUCTION

Reports presented by the World Health Organization (WHO) indicate that non-communicable chronic diseases (NCDs) constitute a major global challenge, as they account for approximately 71% of annual global deaths, corresponding to 41 million fatalities [1]. When ranked by mortality, cardiovascular diseases lead with about 18 million deaths per year, followed by cancer with 9.3 million, and then respiratory diseases and diabetes with 4.1 million and 2 million deaths, respectively.

In recent decades, diabetes has shown a marked increase, currently affecting 537 million adults worldwide [2]. In particular, type 1 diabetes, also known as juvenile or insulin-dependent diabetes, is estimated to account for 5–10% of all diagnoses and primarily affects children, adolescents, and young adults. Approximately 1.2 million individuals under the age of 20 are estimated to live with this condition, with an annual increase in new cases of nearly 3%.

This sustained rise in incidence has placed a considerable burden on healthcare systems worldwide. According to The Lancet Diabetes & Endocrinology, the direct and indirect costs associated with this disease amount to approximately USD 27,000 per patient per year in developed countries [3]. Moreover, significant geographical differences exist: Finland, for example, reports 62.3 cases per 100,000 inhabitants, whereas in China and Japan the incidence is below 3 cases per 100,000 [4]. This pronounced disparity suggests that the development of the disease involves a complex interaction of social, genetic, and environmental factors.

As reported in [2], exposure to certain environmental pollutants, changes in Western dietary patterns, and alterations in the gut microbiome may be driving the increase in cases. This trend is confirmed by a meta-analysis showing that the incidence rate of type 1 diabetes in most countries has increased annually by 3–4%, with particularly high prevalence in children under the age of five [5].

In response to this scenario, the Global Diabetes Compact initiative was launched by the WHO [6], focusing on improving early diagnosis, access to treatment, and the prevention of complications. The initiative places special emphasis on low- and middle-income countries, where major barriers to insulin access and the availability of continuous glucose monitoring systems remain a substantial challenge.

Strategies for the early detection and risk assessment of type 1 diabetes have undergone significant changes in recent years. Traditional assessment methods have relied on various clinical approaches and specific biomarkers, including:

A. Immunological Markers: The detection of specific antibodies targeting pancreatic beta cells—such as islet cell antibodies (ICA), insulinoma-associated antigen-2 (IA-2), and glutamic acid decarboxylase (GAD)—has enabled the prediction of disease development with an accuracy of 85–90% [7].

B. Genetic Studies: Analysis of HLA (Human Leukocyte Antigen) alleles has proven crucial for risk identification, leading to the discovery of more than 50 associated loci [8].

C. Metabolic Tests: Monitoring beta-cell function through glucose tolerance tests, both intravenous (IVGTT) and oral (OGTT), has been established as an essential evaluation method. Studies from the TEDDY program have shown that these metabolic alterations can be detected years before clinical diagnosis [9].

However, these traditional approaches present significant limitations, including difficulty in integrating complex interactions among multiple variables, limited capacity to process longitudinal data, and predictive accuracy that rarely exceeds 90% [5][6][7].

Within this context, machine-learning-based predictive models offer an alternative approach capable of handling large data volumes, non-linear relationships, and inter-individual variability. This research is framed within this purpose, focusing on the design and evaluation of various machine learning (ML) models aimed at predicting blood glucose levels with a one-hour forecasting horizon in patients with type 1 diabetes.

RELATED WORK

Estimating glucose levels in individuals with type 1 diabetes represents a significant challenge, as multiple factors are involved, including both physiological and behavioral aspects. Early approaches relied on traditional statistical and mathematical methods, such as least squares and autoregressive models [13]. Although simple and computationally inexpensive, these approaches have shown limitations when faced with individual variability and the influence of external factors, such as physical activity levels or stress responses [10], [11], [12].

With the recent advancement of intelligent systems—particularly machine learning (ML) predictive models—there has been a surge in their use due to increased computational capabilities. Neural architectures, in particular, have gained prominence due to their adaptability to various contexts. For instance, recurrent neural networks (RNNs), and especially long short-term memory (LSTM) networks, stand out for their ability to detect complex temporal patterns [14], [15]. This capability is crucial in glucose monitoring, where variations follow nonlinear dynamics.

Other proposals, such as those presented in [13], have suggested the development of hybrid approaches that combine physiological representations with machine learning algorithms. These models integrate information on glucose metabolism with historical data to improve the accuracy and reliability of predictions [13]. In parallel, deep neural networks (DNNs) and convolutional neural networks (CNNs) have enabled the identification of complex patterns thanks to their capacity to handle large volumes of data that are not readily captured by traditional models [16], [17].

The integration of IoT devices, such as wearable sensors and continuous glucose monitors, has transformed the data collection process. This connectivity enables continuous and real-time monitoring, thereby improving the quality of the data repositories used for developing predictive models [12], [18]. Likewise, the personalization of models based on individual variables—such as genetics, dietary habits, and insulin response—has demonstrated improvements in both predictive accuracy and clinical usefulness [19], [20].

Despite these advances, significant technical challenges remain. High inter-patient variability, the scarcity of labeled data, and limitations in model generalization hinder large-scale clinical deployment. Emerging research directions focus on transfer learning, the use of next-generation sensors, and the integration of real-time artificial intelligence to enhance model adaptability [21], [22], [23].

Finally, Martinsson et al. [25] emphasize that factors such as the rate of carbohydrate absorption and interactions with different types of insulin directly affect glycemic stability. This underscores the need for models that are not only accurate but also capable of integrating multiple variables within personalized contexts.

METHODOLOGY

The methodology followed in this paper consists of the following key stages.

A. Data Collection and Processing

A dataset collected from the BrisT1D Blood Glucose Prediction Competition [25] was used. The dataset is based on a study that gathered information from young adults in the United Kingdom with type 1 diabetes. Continuous glucose monitoring (CGM) devices, as well as insulin pumps and smartwatches, were employed to record blood glucose measurements, insulin administration, carbohydrate intake, and physical activity levels.

Temporal feature engineering was performed to prepare the data for glucose prediction tasks. Initially, data cleaning was carried out by filling missing values using strategies based on column-wise or row-wise averages, as appropriate. Subsequently, categorical activity variables were encoded using a one-hot encoding process, and numerical variables were transformed using standardization or normalization techniques. In addition, temporal features such as hour, minute, and second were extracted from the original timestamps. The data were structured into five-minute intervals, generating samples that represent specific

moments in time. Each sample incorporates accumulated information from the previous six time intervals (i.e., the previous six hours) [25].

B. Model Training and Selection

Different machine learning algorithms were tested to evaluate which one was best suited to data prediction:

1. Linear Regression: A mathematical model that assumes a linear relationship between the predictor variables and the target variable.
2. Decision Tree: Based on a hierarchical structure constructed by iteratively splitting the data using decision rules derived from the features.
3. Random Forest: An ensemble method based on a collection of decision trees (bagging). Its novelty lies in the random selection of features to reduce correlation and increase diversity among the trees. In regression problems, the final decision is obtained by averaging the predictions of all trees.
4. Gradient Boosting (GBM/GBoost): An ensemble technique in which weak learners (trees) are built sequentially, with each new tree aiming to correct the errors of the previous one, ultimately leading to error reduction.
5. TabNet: A neural-network-based model that employs attention mechanisms to dynamically select which features to use at each step. It is specifically designed for tabular data.

The final model was selected based on its performance as measured by the chosen evaluation metric.

C. Evaluation Metrics

The primary metric used was the Root Mean Squared Error (RMSE), as it reflects the prediction error in the same units as the target variable. It is defined as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{y_i}\right)^2} \quad (1)$$

Where y_i represents the real values. $\hat{y}_i$ the predicted values, and n the total number of samples. In addition, other metrics were used to assess and visualize the performance of each model, such as the Mean Squared Error (MSE):

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{y_i}\right)^2 \quad (2)$$

The Mean Absolute Error (MAE) was also considered, defined as:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \widehat{y_i}\right| \quad (3)$$

Where $|.|$ denotes the absolute value, thus preserving the error in the same units as the data. The Mean Absolute Percentage Error (MAPE), defined as:

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \widehat{y_i}}{y_i}\right| \times 100 \quad (4)$$

measures the error in relative terms and is expressed as a percentage. Furthermore, the Coefficient of Determination (R^2 or R-squared) was included as an indicator of the proportion of variance explained by the model, defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \widehat{y_i}\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2} \quad (5)$$

Finally, the Explained Variance Ratio (EVR) was incorporated, which measures the proportion of the total variance explained by the model and is given by:

$$\mathrm{EVR} = 1 - \frac{\text{error variance}}{\text{total variance}} \quad (6)$$

This metric is similar to but is adapted for more general cases.

D. Validation

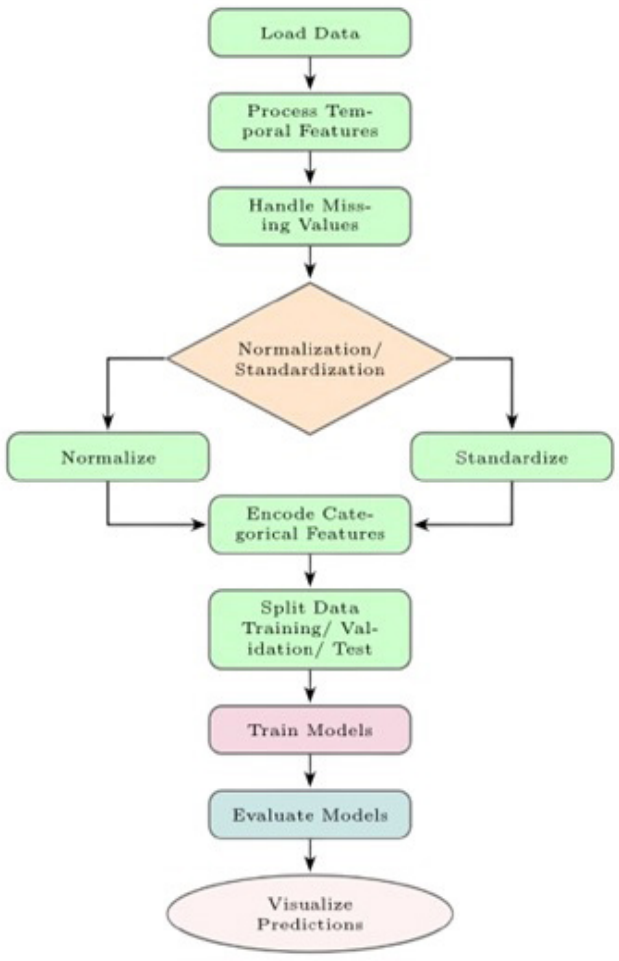
The original dataset was divided into two parts: 80% of the data was used for training the algorithms, while the remaining 20% was used as the test set to evaluate the models developed during training.

E. Results Visualization

Scatter plots were generated to assess whether the models correctly captured the relationships between the input features and blood glucose levels one hour ahead for each sample.

A summary of the methodology proposed in this article, as described in the previous subsections, is presented in Figure 1. Each stage is designed so that the data are processed and transformed before being used in the training phase and subsequently in the evaluation of the trained models. The final stage includes the visualization of the predictions generated by each model.

Figure 1. Proposed methodology flow diagram.



F. Grid Search Application

In this step, a grid search procedure was performed to select the optimal hyperparameters for each model, with the aim of comparing the results obtained using this method against those achieved previously without hyperparameter tuning. Table 1 lists the ranges of values used in the hyperparameter search.

TABLE 1. RANGE OF PARAMETER VALUES FOR THE GRID SEARCH.

Modelo	Parámetro	Valor
Linear regression	Fit intercept	[True, False]
Decision tree	max_depth	[None, 5, 10, 20, 30]
	min_samples_split	[2, 3, 4, 5,10]
	min_samples_leaf	[1, 2, 3,10]
	max_features	[None, ‘auto’, ‘sqrt’, ‘log2’]
Random forest	n_estimators	[10, 50, 100, 150, 200]
	max_depth	[None, 5, 10, 20, 30]
	min_samples_split	[2, 3, 4, 5,10]
	min_samples_leaf	[1, 2, 3, 5]
	max_features	[None, ‘sqrt’, ‘log2’]

XG Boost	n_estimators	[10, 50, 100, 150, 200]
	max_depth	[6, 10, 20, 30]
	Learning_rate	[0.3,0.5]
TabNet	n_d	[8,16]
	n_a	[8]
	n_steps	[3,5]
	gamma	[1.3,1.5]
	lambda_sparse	[0.001,0.01]

RESULTS

In order to quantify model performance in terms of generalization capability, the machine learning models were evaluated using the standard metrics previously defined: RMSE, MSE, MAE, MAPE, EVR, and the Coefficient of Determination (R^2).

The Python code developed for this study is publicly available online at the following repository: https://github.com/usuario3pm/Predicci-on-de-Niveles-de-Glucosa-en-Sangre-para-la-Diabetes-Tipo-1

A. Performance with Manual Hyperparameter Tuning

Table 2 presents the results obtained after an iterative manual tuning of the hyperparameters on the training dataset. When examining the MAE metric, the Random Forest model stands out by achieving the lowest absolute error, with a value of only 0.49. This behavior is consistent across the other metrics, with a MAPE of 6.47% and an RMSE of 0.74, explaining approximately 94% of the data variance (R^2 and EVR). These results suggest that Random Forest more faithfully captures the nonlinear fluctuations in glucose levels, which is essential for preventing critical events such as hypoglycemia or hyperglycemia.

In contrast, TabNet exhibited the poorest performance, likely due to suboptimal hyperparameter selection or a reduced ability to model complex tabular relationships in this specific context. The Decision Tree and XGBoost models yielded intermediate results, while Linear Regression was outperformed by more advanced models, although it may still be considered due to its simplicity and low computational cost. Overall, Random Forest emerges as the most suitable model for accurate and reliable glucose monitoring in this study.

TABLE 2. RESULTS OF THE MODEL METRICS DURING THE TRAINING PHASE WITH MANUAL HYPERPARAMETER TUNING.

Models/Metrics	**MAE**	**MAPE (%)**	**MSE**	**RMSE**	**R^2**	**EVR**
Linear Regression	1.53	20.33	4.26	2.06	0.53	0.53
Decision Tree	1.42	19.03	3.66	1.91	0.59	0.59
Random forest	**0.49**	**6.47**	**0.55**	**0.74**	**0.94**	**0.94**
XGBoost	1.25	16.89	2.80	1.67	0.69	0.69
TabNet	1.54	21.03	4.27	2.07	0.52	0.53

In the analysis of the glucose level prediction results on the test set (see Table 3), the outcomes are consistent with those observed during the training phase (see Table 2), confirming that the models were properly trained without underfitting or overfitting.

The Random Forest model remains the best-performing approach, achieving the lowest RMSE (0.74), MAE (0.49), and MAPE (6.47%), while maintaining similar R^2 and EVR value (0.94), This indicates a strong generalization capability and reliable performance on previously unseen data. In contrast, TabNet and Linear Regression exhibited the poorest performance, with high errors across all metrics. Decision Tree and XGBoost provided intermediate results, offering lower computational complexity but at the expense of reduced predictive accuracy.

TABLE 3. RESULTS OF THE MODEL METRICS DURING THE TEST PHASE WITH MANUAL HYPERPARAMETER TUNING.

Models/Metrics	MAE	MAPE (%)	MSE	RMSE	R^2	EVR
Linear Regression	1.55	20.59	4.36	2.09	0.52	0.52
Decision Tree	1.44	19.28	3.74	1.93	0.59	0.59
Random forest	**0.50**	**6.59**	**0.57**	**0.75**	**0.94**	**0.94**
XGBoost	1.26	17.10	2.85	1.69	0.69	0.69
TabNet	1.56	21.31	4.36	2.09	0.52	0.52

B. Comparison with Grid Search

Subsequently, Grid Search was applied to optimize the hyperparameters of each model (see Table 4). However, the results were generally inferior to those obtained through manual tuning, even though the manually selected hyperparameter values were included in the search space, except for linear regression. In particular, the Random Forest model exhibited a notable degradation in performance (RMSE = 1.08 and R^2 = 0.87), indicating a tendency toward overfitting during the automated search process. This suggests that the grid search selected parameters that performed well on the training data but generalized poorly to unseen data, resulting in overfitting.

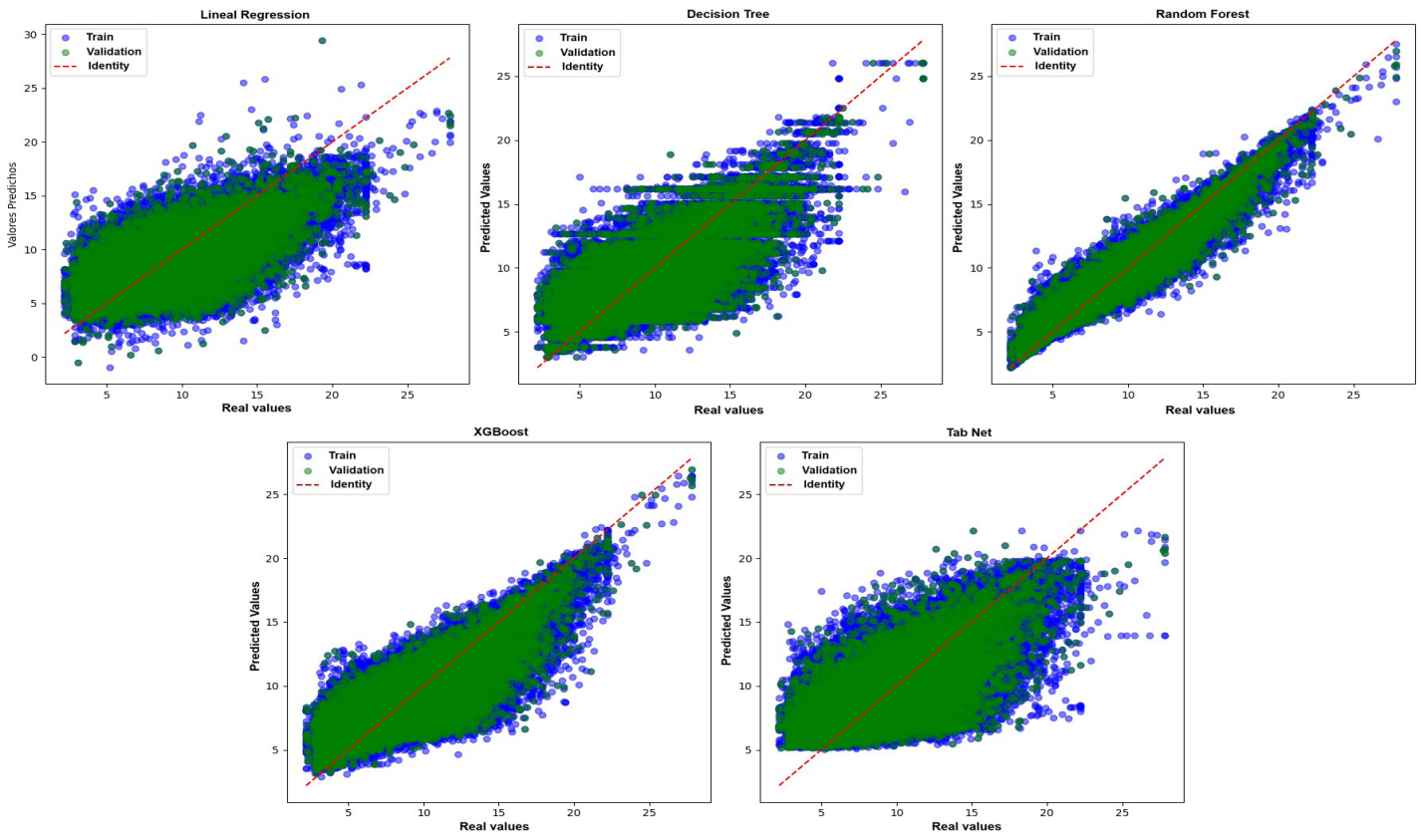
These findings indicate that, for this type of dataset, manual hyperparameter tuning guided by empirical criteria may be more effective than automated grid search strategies.

TABLE 4. RESULTS OF THE MODEL METRICS DURING THE TEST PHASE WITH GRID SEARCH TUNING.

Models/ Metrics	MAE	MAPE (%)	MSE	RMSE	R^2	EVR
Linear Regression	1.55	20.59	4.36	2.09	0.52	0.52
Decision Tree	1.58	21.20	4.47	2.11	0.51	0.51
Random forest	0.80	10.94	1.16	1.08	0.87	0.87
XGBoost	1.45	19.56	3.80	1.95	0.58	0.58
TabNet	1.50	20.02	4.11	2.03	0.55	0.55

Finally, to complete the analysis of the best models obtained through manual tuning, the scatter plots shown in Figure 2 were generated. The purpose of these plots is to evaluate the accuracy of each model by comparing their predictions with the actual values, as well as to identify systematic biases and the dispersion of the prediction error.

Figure 2. Scatter plots for blood glucose level prediction using the five models on the training and test sets. Blue points represent the training set, while green points correspond to the test set. Points located closer to the dashed lines indicate better model predictions.



The red line (identity line, $y = x$) serves as a reference for what a perfect prediction should be, that is, an indicator of good model generalization. It also allows the previously discussed results to be visualized graphically. In these plots, the Random Forest model exhibits a high concentration of points along the reference line in both the training and test sets, reinforcing its predictive capability. In contrast, the lower-performing models show greater dispersion and poorer alignment with the ideal line.

CONCLUSIONS

This comparative study proposed a methodology consisting of preprocessing, model development, evaluation, and visualization. The preprocessing phase included temporal feature engineering, where data cleaning tasks were performed, missing values were imputed using averaging strategies, categorical variables were encoded through one-hot encoding, and numerical variables were standardized or normalized. Additionally, the samples were structured into five-minute intervals, considering accumulated information from the previous six hours to predict blood glucose levels one hour ahead.

The results of the comparison among the five models showed that Random Forest was the best-performing prediction model, achieving an RMSE of only 0.75 and an R^2 of 0.94, consistently outperforming alternatives such as XGBoost, TabNet, and linear regression. This superiority suggests its suitability as a robust tool for clinical predictive monitoring applications, particularly in contexts where glycemic stability is critical.

It was found that manual hyperparameter tuning, based on metric-guided iterations, was more effective than the use of Grid Search, likely due to its ability to avoid overfitting.

Overall, the findings support the use of ensemble methods such as Random Forest in digital health environments aimed at personalized treatment. Nevertheless, it is important to consider the associated computational cost and to explore optimization strategies that enable efficient deployment on low-power devices or in resource-constrained environments.

Finally, this work lays the groundwork for future studies focused on predictive models integrated into clinical decision support systems, as well as their validation on real-world cohorts using real-time data. The incorporation of new physiological data sources, transfer learning techniques, and continuous update mechanisms could further enhance the adaptive capacity of these models in dynamic and personalized scenarios.

CRediT AUTHORSHIP CONTRIBUTION STATEMENT

O. Aguirre-Aguirre: Data curation, Conceptualization, Investigation, Methodology, Writing. E. Guevara-Povea: Data curation, Conceptualization, Investigation, Methodology, Writing. J. Hoyos-Sanchez: Conceptualization, Supervision, Methodology, Writing-review and editing.

FUNDING

This work did not receive external funding; support in terms of infrastructure and personnel was provided by the university.

ACKNOWLEDGMENTS

Juan Pablo Hoyos-Sanchez agradece a la Universidad Nacional de Colombia.

REFERENCIAS

[1] World Health Organization, Noncommunicable Diseases, Rehabilitation and Disability: Surveillance, Monitoring and Reporting (SMR). Geneva, Switzerland: WHO, 2022. https://www.who.int/publications/i/item/9789240047761

[2] International Diabetes Federation, IDF Diabetes Atlas, 10th ed. Brussels, Belgium: IDF, 2021. https://diabetesatlas.org/

[3] M. Romanello et al., "The 2022 report of the Lancet countdown on health and climate change," *The Lancet Diabetes & Endocrinology*, vol 400, pp.1619–1654, 2022. https://doi.org/10.1016/S0140-6736(22)01540-9

[4] Putula et al., "All-cause mortality and factors associated with it in Finnish patients with type 1 diabetes," *Journal of Diabetes and its Complications*, vol 38, no. 12, 2024. https://doi.org/10.1016/j.jdiacomp.2024.108881

[5] Mameli et al., "Lessons and gaps in the prediction and prevention of type 1 diabetes", *Pharmacological Research*, vol 193, 2023. https://doi.org/10.1016/j.phrs.2023.106821

[6] Gregory et al., "Global incidence, prevalence, and mortality of type 1 diabetes in 2021 with projection to 2040: a modelling study." *The Lancet Diabetes Endocrinology*, vol 10, no. 10, pp. 741–760, 2022. https://doi.org/10.1016/S2213-8587(22)00218-2

[7] N. Foster et al., "State of type 1 diabetes management and outcomes," *Diabetes Technology & Therapeutics*, vol 21, no. 2, pp. 66–72, 2019. https://doi.org/10.1089/dia.2018.0384

[8] S. Rich, "Type 1 Diabetes Genetics Consortium (T1DGC)," Version 7 [Dataset], NIDDK Central Repository, 2014. https://doi.org/10.58020/qrdt-eh40

[9] TEDDY Study Group, "The Environmental Determinants of Diabetes in the Young (TEDDY) Study," *Annals of the New York Academy of Sciences*, vol. 1150, pp. 1-13, 2008. https://doi.org/10.1196/annals.1447.062

[10] A. Facchinetti, "Continuous Glucose Monitoring Sensors: Past, Present and Future Algorithmic Challenges", *Sensors*, vol. 16, no. 12, 2016. https://doi.org/10.3390/s16122093

[11] Borle, Neil C et al. "The challenge of predicting blood glucose concentration changes in patients with type I diabetes." *Health informatics journal* vol. 27, no. 1, 2021. https://doi.org/10.1177/1460458220977584

[12] M. S. Farooq et. Al., "Role of Internet of Things in diabetes healthcare: Network infrastructure, taxonomy, challenges, and security model," *Digital Health*, vol. 9, 2023. https://doi.org/10.1177/20552076231179056

[13] Sparacino et al., "Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time-series," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 5, pp. 931-937, May 2007. https://doi.org/10.1109/TBME.2006.889774

[14] Jaloli, M., & Cescon, M. (2023). 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*, vol 17, no. 6. https://doi.org/10.1177/19322968221092785

[15] Contreras I, Vehi J., "Artificial Intelligence for Diabetes Management and Decision Support: Literature Review," *Journal of Medical Internet Research*, vol 22, no.2, 2020. https://doi.org/10.2196/10775

[16] Wonju S., et al. "A personalized blood glucose level prediction model with a fine-tuning strategy: A proof-of-concept study." *Computer methods and programs in biomedicine* vol. 211, 2021. https://doi.org/10.1016/j.cmpb.2021.106424

[17] Ghimire, Sarala et al. "Deep learning for blood glucose level prediction: How well do models generalize across different data sets?." *PloS one* vol. 19, no. 9, pp. e0310801. 25 Sep. 2024. https://doi.org/10.1371/journal.pone.0310801

[18] A. K. M and S. Baskar, "Internet of Things based Real-Time Hyperglycemia and Hypoglycemia Monitoring using Wearable Biosensors," *2022 International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 2022*, pp. 487-493. https://doi.org/10.1109/ICEARS53579.2022.9751878

[19] T. Zhu, K. Li, P. Herrero and P. Georgiou, "Personalized Blood Glucose Prediction for Type 1 Diabetes Using Evidential Deep Learning and Meta-Learning," *IEEE Transactions on Biomedical Engineering*, vol. 70, no. 1, pp. 193-204, Jan. 2023. https://doi.org/10.1109/TBME.2022.3187703

[20] D. Klonoff, et al. "Continuous glucose monitoring: A review of the technology and clinical use." *Diabetes research and clinical practice* , vol. 133, 2017. https://doi.org/10.1016/j.diabres.2017.08.005

[21] Khokhar, P. Bakhsh et al. “Advances in artificial intelligence for diabetes prediction: insights from a systematic literature review.” *Artificial intelligence in medicineK*, vol. 164, 2025. https://doi.org/10.1016/j.artmed.2025.103132

[22] M. T. I Rimon et al. “Advancements in Insulin Pumps: A Comprehensive Exploration of Insulin Pump Systems, Technologies, and Future Directions.” *Pharmaceutics*, vol. 16, no. 7, Jul. 2024. https://doi.org/10.3390/pharmaceutics16070944

[23] Dunstan, J., Villena, F., Hoyos, J. et al. “Predicting no-show appointments in a pediatric hospital in Chile using machine learning”. *Health Care Manag Sci*, vol. 26, pp. 313–329, 2023. https://doi.org/10.1007/s10729-022-09626-z

[24] F. Ladislav and D. Vašata. 2024. “Predicting Blood Glucose Levels with LMU Recurrent Neural Networks: A Novel Computational Model”, In Artificial Intelligence in Medicine: 22nd International Conference, AIME 2024, Salt Lake City, UT, USA, July 9–12, 2024, Proceedings, Part I. Springer-Verlag, Berlin, Heidelberg, 117–127. https://doi.org/10.1007/978-3-031-66538-7_13

[25] S. G. James, et. al., BrisT1D Blood Glucose Prediction Competition, 2024. Kaggle. https://kaggle.com/competitions/brist1d