ESKİŞEHİR TEKNİK ÜNİVERSİTESİ BİLİM VE TEKNOLOJİ DERGİSİ
C- YAŞAM BİLİMLERİ VE BİYOTEKNOLOJİ

Eskişehir Technical University Journal of Science and Technology
C -Life Sciences and Biotechnology

Estuscience – Life, 2026, 15 (1) pp. 46-65, DOI: 10.18036/estubtdc.1674345

RESEARCH ARTICLE

ENHANCING DIABETES PREDICTION WITH INTERPRETABLE MACHINE LEARNING: A COMPARATIVE ANALYSIS OF ADDITIVE–MULTIPLICATIVE NEURAL NETWORKS AND KOLMOGOROV–ARNOLD NETWORKS

Şeyda DEMİREL TATLI [1], Kürşad AYTEKİN [2], Melih AGRAZ [3, 4, 5*]

[1] Department of Finance-Banking and Insurance, Giresun University, Giresun, Türkiye
seyda.demirel@giresun.edu.tr, - 0000-0002-8736-5162

[2] Department of Anatomy, Department of Orthopedics and Traumatology, Faculty of Medicine, Giresun University, Türkiye
kursadaytekin@gmail.com, - 0000-0002-6969-1183

[3] Department of Data Science and Analytics, Giresun University, Giresun, Türkiye
[4] Department of Medicine, Brown University, Providence, RI, 02903, USA
[5] Division of Cardiology, Brown University Health, Providence, RI, 02903, USA
melih.agraz@giresun.edu.tr, - 0000-0002-6597-7627

Abstract

This study investigates the effectiveness of machine learning (ML) models in diagnosing diabetes and identifying the most influential predictors using the PIMA Indians Diabetes dataset. A particular emphasis is placed on novel neural network architectures, especially the Additive and Multiplicative Neurons Network (AMNN), introduced as a key innovation in this work.

The dataset underwent comprehensive preprocessing, including handling missing values, feature scaling, and addressing class imbalance via the SMOTE algorithm. To interpret the importance of predictors, five feature selection techniques (Correlation, Boruta, MRMR, RFE, Random Forest) and two explainable AI (XAI) tools (SHAP and LIME) were applied.

A total of eight machine learning algorithms were tested and evaluated based on accuracy, recall, F1-score, and AUC-ROC. Among all models, AMNN achieved the best performance, with an accuracy of 0.7576, recall of 0.7576, F1-score of 0.7618, and AUC-ROC of 0.8206. MLP-2 and XGBoost also showed competitive results. Kolmogorov-Arnold Networks (KAN), while not outperforming other models, demonstrated moderate success and offered interpretability advantages due to its flexible activation structure.

Consistently, glucose, BMI, age, and pregnancy count were found to be the most significant predictors across feature selection and XAI evaluations. These results align with existing clinical insights into diabetes risk.

In conclusion, this study highlights the potential of the AMNN model as a powerful and interpretable tool for early diabetes detection. These findings suggest that AMNN offers a compelling balance between performance and interpretability, making it suitable for real-world medical applications. The integration of feature selection and XAI techniques supports model transparency, paving the way for its application in clinical decision-making. Future work should focus on enhancing generalizability through larger datasets and hybrid modeling strategies.

Keywords

Diabetes classification,
Kolmogorov–Arnold networks models,
Additive and Multiplicative neurons network models,
Machine learning

Time Scale of Article

Received :11 April, 2025
Accepted : 22 November 2025
Online date : 27 January 2026

*Corresponding Author: melih.agraz@giresun.edu.tr

1. INTRODUCTION

Diabetes is one of the most widespread chronic diseases worldwide, presenting a growing public-health challenge due to its serious complications such as cardiovascular disorders, kidney failure, neuropathy, and blindness [1]. According to the World Health Organization, the global diabetes burden continues to increase each year, leading to substantial healthcare costs and reduced quality of life [49]. Early and accurate diagnosis is therefore critical to prevent long-term damage and improve patient health. However, conventional diagnostic approaches often rely on limited biomarkers and may fail to detect complex, nonlinear interactions within clinical data. With rapid technological advancements shaping the modern world, artificial intelligence and machine learning have become pivotal tools in revolutionizing medical studies, particularly in fields like genetics [43], drug discovery [44], and the diagnosis of diseases [21]. Despite numerous studies applying conventional ML models such as Random Forests, Support Vector Machines, and standard neural networks to diabetes prediction, challenges remain in balancing model performance, interpretability, and generalizability. This gap motivates the exploration of novel neural architectures capable of both high accuracy and transparent reasoning. Machine learning methods accelerate the diagnostic process, reducing the time patients spend in healthcare facilities, and automate healthcare systems, thereby decreasing the need for human labor. Machine learning algorithms can process large and diverse datasets efficiently. This enables fast and accurate predictions at a reduced computational cost. Therefore, many researchers prefer machine learning techniques for disease diagnosis [2]. In this context, the present study introduces the Additive and Multiplicative Neurons Network (AMNN) as a novel and interpretable neural architecture designed to enhance diabetes diagnosis performance. Additionally, the study investigates the Kolmogorov–Arnold Networks (KAN), another emerging model known for its flexible activation structure and interpretability advantages. As a recently developed model, KAN has begun to find applications particularly in the field of biology [50]. By integrating explainable AI (XAI) tools (SHAP and LIME) and advanced feature selection methods, this work not only seeks to improve predictive accuracy but also to uncover the most influential clinical predictors driving diabetes risk.

In this study, the PIMA dataset was used for machine learning algorithms. The dataset originally was obtained from the National Institute of Diabetes and Digestive and Kidney Diseases. It is designed to predict whether an individual has diabetes based on a range of diagnostic indicators, while also highlighting the factors that contribute to diabetes. Notably, all participants in the dataset are Pima Indian women aged 21 or above [3]. There is extensive research in the literature related to the PIMA Diabetes dataset. Kayaer and Yildirim [4] examined the General Regression Neural Network (GRNN) structure on the Pima Indians Diabetes dataset and found that it achieved similar classification accuracy to more complex neural network models with a simpler structure. They compared GRNN with commonly used neural network models such as Multi-Layer Perceptron (MLP) and RBF, obtaining an accuracy rate of 80.21% for GRNN and 81% for ARTMAP-IC.4 Karatsiolis and Schizas [5] aimed to improve these accuracy rates by using a modified Support Vector Machine (SVM) strategy, achieving an accuracy of 82.2% for diabetes classification on the Pima Indians Diabetes dataset. Yangın [6] compared decision trees, Random Forest, Gradient Boosting, and XGBoost (eXtreme Gradient Boosting) machine learning algorithms on two different datasets, including the Pima Diabetes dataset, and identified the highest classification accuracy with 82.35%. Ganesh and Sripriya [7] examined different classification methods on the Pima Indians Diabetes dataset, evaluating the advantages and disadvantages of techniques used for diabetes prediction and demonstrating that preprocessing improves classification accuracy. Lakhwani et al. [8] developed an artificial neural network (ANN)-based automated diabetes diagnosis system using the Pima Indians Diabetes dataset and used cumulative gain charts to measure model quality. Patra and Khuntia [9] improved the k-nearest neighbors (KNN) classifier using a new standard deviation-based distance calculation method on the Pima Indians Diabetes dataset, achieving an accuracy of 83.2%. Mousa et al. [10] compared Long Short-Term Memory (LSTM), Random Forest (RF), and Convolutional Neural Network (CNN) models for detecting diabetes using the same dataset, showing that LSTM achieved the highest accuracy (85%) and was

particularly effective in capturing temporal patterns. Chang et al. [11] developed an e-diagnosis system for diabetes using the Pima Indians Diabetes dataset and three interpretable machine learning models: Naïve Bayes, Random Forest (RF), and J48 decision trees. The performance of the models was calculated using accuracy, precision, recall, and specificity. Among these, the Random Forest (RF) model achieved the highest accuracy (79.57%). However, further comparisons indicated that Naïve Bayes worked effectively with a smaller number of features, whereas J48 excelled in maintaining high sensitivity. In addition, Farsana and Poulose [12] analyzed the PIMA dataset by comparing hybrid CNN models with traditional machine learning approaches to aid in the early diagnosis and prevention of diabetes. Their hybrid CNN model attained an accuracy of 73%.

The objective of this study is to enhance diabetes diagnosis using various machine learning algorithms and to determine the most influential features in classification performance. To achieve this, different models were trained and evaluated, and the best-performing algorithm was identified based on key metrics. Furthermore, feature importance was analyzed using both feature selection techniques and explainable artificial intelligence (XAI) methods. The primary innovation of this study is the implementation of the Additive and Multiplicative Neurons Network (AMNN), which outperformed all other models. Additionally, the study introduces the Kolmogorov-Arnold Networks (KAN) model as an alternative approach to exploring flexible neural architectures.
The remaining sections of this study are organized as follows: The second section introduces the Pima Indians Diabetes dataset and describes the preprocessing steps applied to the dataset. The third section discusses model selection criteria and presents the selected models. The fourth section compares the methods used in the study and reports the findings. Finally, the last section presents the conclusions and recommendations based on the obtained results.

2. METHODS

2.1. Dataset

The PIMA dataset consists of 768 observations and 9 variables. Among the 768 individuals in the dataset, 268 have diabetes, while 500 do not have diabetes. The independent variables for diabetic individuals are described in Table 1.

Table 1. Variables in the PIMA Indians Diabetes dataset.

Variable	Unit	Description
Pregnancies	Count	Number of pregnancies
Glucose	mg/dL	2-hour plasma glucose concentration
Blood Pressure	mmHg	Diastolic blood pressure
Skin Thickness	Mm	Triceps skin fold thickness
Insulin	μU/ml	2-hour serum insulin level
BMI	kg/(m^2)	Body mass index (BMI)
Diabetes Pedigree Function	Numeric	Diabetes pedigree function
Age	Years	Age of the individual
Outcome	Categorical	Class label (0: non-diabetic, 1: diabetic)

To assess the suitability of the dataset, we draw a t-SNE plot, as shown in Figure 1.

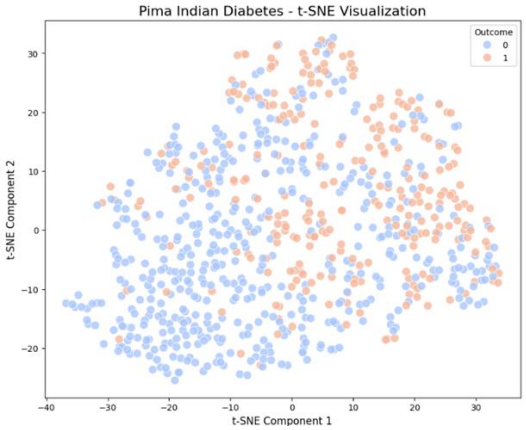

Figure 1. Two-dimensional t-SNE visualization of the PIMA Indians Diabetes dataset showing the distribution and separability of diabetic (orange) and non-diabetic (blue) cases.

As seen in Figure 1, the dataset exhibits certain clustering patterns; some data points form distinct clusters, while others appear more scattered. The colors in the visualization indicate that data points labeled as Outcome = 0/no diabetes (blue) and Outcome = 1/diabetes (orange) are somewhat distinguishable but not completely separable. Thus, it can be concluded that the dataset is suitable for a classification problem.

2.2. Data Preprocessing

The dataset was preprocessed using Python (version 3.10.12). Descriptive statistics of the dataset are provided in Table 2.

Table 2. Descriptive statistics of the PIMA Indians Diabetes dataset.

Variable	Mean	Std	Min	25%	50%	75%	Max
Pregnancies	3,85	3,37	0,00	1,00	3,00	6,00	17,00
Glucose	120,89	31,97	0,00	99,00	117,00	140,25	199,00
Blood Pressure	69,11	19,36	0,00	62,00	72,00	80,00	122,00
Skin Thickness	20,54	15,95	0,00	0,00	23,00	32,00	99,00
Insulin	79,80	115,24	0,00	0,00	30,50	127,25	846,00
BMI	31,99	7,88	0,00	27,30	32,00	36,60	67,10
Diabetes Pedigree Function	0,47	0,33	0,08	0,24	0,37	0,63	2,42
Age	33,24	11,76	21,00	24,00	29,00	41,00	81,00

Additionally, the following data preprocessing methods were applied throughout the entire analysis
Missing Data: Although there are no explicitly missing values in the dataset, certain variables contain an excessive number of zero values, which are considered missing. These zero values were treated as missing observations and imputed using the median.

Data Scaling: An analysis of the dataset revealed that variables have different magnitudes. For instance, the Age variable ranges from 21 to 81, whereas the Diabetes Pedigree Function variable ranges from 0.07 to 2.42. To eliminate these differences, min-max scaling and Z-score normalization were applied.

Imbalanced Data: Imbalanced data refers to datasets in which the distribution of classes is markedly unequal, a common challenge in classification problems [13]. In the Pima dataset, there is a pronounced

imbalance in class representation. In the Pima dataset, there is a noticeable difference between diabetic (34.9%) and non-diabetic (65.1%) individuals. This class imbalance can hinder the accurate prediction of the minority class (diabetic individuals). To address this issue, the Oversampling (SMOTE) technique was applied during the training process [14].

2.3. Feature Engineering

2.3.1. Feature Selection Methods

In this study, after the data were splitted into train and test sets, the feature selection methods described below were applied to the training data.

Correlation: When selecting features, it is essential to examine the relationships between variables. Independent variables should be highly correlated with the dependent variable; however, high correlations among independent variables can cause multicollinearity, which is undesirable. One of the feature selection techniques, CFS (Correlation Feature Selection), determines a subset of features using a correlation matrix. When analyzing the correlation matrix, it is observed that the variables with a correlation value of $0.20^{+}$with the outcome variable are ['Glucose', 'BMI', 'Pregnancies', 'Age']. We plot the correlation matrix, as seen in Figure 2 below.

Figure 2. Pearson correlation matrix heatmap for all variables in the PIMA dataset.

Boruta: Another feature selection method used is Boruta [15]. This method analyzes the relationships between dependent and independent variables to determine which variables are important for the predictive model. Based on the Random Forest (RF) algorithm, Boruta assesses the importance of each variable by comparing it with randomly generated "shadow" variables. Through this comparison, it identifies the most significant real features [16]. Selected important features: ['Pregnancies', 'Glucose', 'BMI', 'DiabetesPedigreeFunction', 'Age']. Rejected features: ['BloodPressure', 'SkinThickness', 'Insulin'].

Maximum Relevance - Minimum Redundancy (MRMR): MRMR is a powerful feature selection technique that that enhances the relevance of selected features to the output while reducing redundancy among them [17]. This method selects the most relevant features to the output while ensuring minimal similarity between selected features. The formula for this method is given in Equation 1 [18].

$$MRMR = max\ \left(\frac{1}{|S|}\sum_{f\in S} I(f,c) - \frac{1}{|S|^2}\sum_{f_i,f_j\in S} I\left(f_i,f_j\right)\right) \quad (1)$$

Selected important features: ['Glucose', 'BMI', 'Pregnancies', 'DiabetesPedigreeFunction', 'Age'].

Recursive Feature Elimination (RFE): This method iteratively eliminates the least important features to select the best-performing set of features. It begins by training step with all available features. Then, in each iteration, it identifies and eliminates the feature with the least significance, retraining the model with the other features. This process continues until the predetermined number of significant features is reached [19]. Feature ranking using RFE: [1, 1, 2, 3, 4, 1, 1, 1]. Selected important features: ['Pregnancies', 'Glucose', 'BMI', 'DiabetesPedigreeFunction', 'Age'].

Random Forest (RF): Random Forest (RF): RF is a widely used for feature selection and classification problems. In this method, the importance scores of variables are calculated, and high-scoring features are retained in the model. RF is an ensemble model consisting of multiple decision trees. It assesses each feature's significance by examining its influence on the splits within decision trees, using measures like Gini index or entropy gain [20,25,47]. The feature importance scores from the RF method are provided in the appendix, but the five most important features are ['Glucose', 'BMI', 'Age', 'DiabetesPedigreeFunction', and 'Pregnancies'].

The features selected by the five different feature selection methods are summarized in Table 3.

Table 3. Features selected by different feature selection methods

Methods	**Features Selected**				
Correlation	Glucose	BMI	Age	Pregnancies	
Boruta	Glucose	BMI	Pregnancies	PedigreeFunction	Age
MRMR	Glucose	BMI	Pregnancies	PedigreeFunction	Age
RFE	Glucose	BMI	Pregnancies	PedigreeFunction	Age
Random Forest	Glucose	BMI	Age	PedigreeFunction	Pregnancies

Using the majority voting method [22] it is observed that 'Glucose', 'BMI', 'Age', and 'Pregnancies' were consistently selected across all methods, highlighting their significance for the model. In methods where feature ranking is important (Correlation, MRMR, and RF), "Glucose" and "BMI" were ranked among the top two features, while the rankings of other features varied by method. However, since ranking is not crucial for Boruta and RFE, features were selected based on their presence rather than their order. Based on these observations, "Glucose" and "BMI" are identified as the most critical features for diabetes diagnosis, followed by "Pregnancies", "DiabetesPedigreeFunction", and "Age".

2.3.2. Explainable Machine Learning (XAI) Approaches

Explainable machine learning (XAI) methods aim to make model decisions understandable and interpretable. After training a model on the Pima dataset, widely used XAI methods such as SHAP (SHapley Additive ExPlanations) or LIME (Local Interpretable Model-Agnostic Explanations) [23]. The SHAP summary plot is provided in Figure 3.

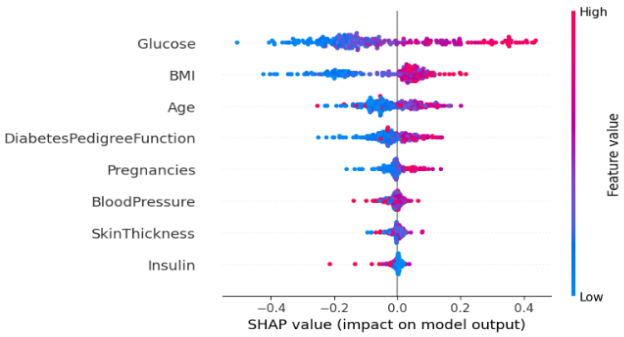

Figure 3. SHAP (SHapley Additive exPlanations) summary plot displaying feature importance and impact direction for diabetes prediction. Each point represents a patient, with red indicating high feature values and blue indicating low values. Glucose emerges as the most influential predictor, followed by BMI, Age, and Pregnancies. Horizontal spread indicates the magnitude of each feature's impact on model predictions.

When analyzing the SHAP plot, it is observed that "**Glucose**" is the most influential feature for diabetes diagnosis. Higher glucose levels (represented by red points) substantially raise the likelihood of a diabetes diagnosis, whereas lower glucose levels (blue points) reduce this probability. After glucose, the next most influential feature is "**BMI**", which also plays a significant role in predicting diabetes risk. As BMI increases, the model assigns a higher likelihood of diabetes diagnosis, highlighting the importance of weight management in diabetes prediction.

2.3.3. Toy LIME example on patient 163

LIME is a model-agnostic technique that provides interpretability by approximating complex model predictions locally with simpler, interpretable models, thus offering insights into the factors influencing individual predictions. LIME visualization is presented on the patient number 163 in Figure 4.

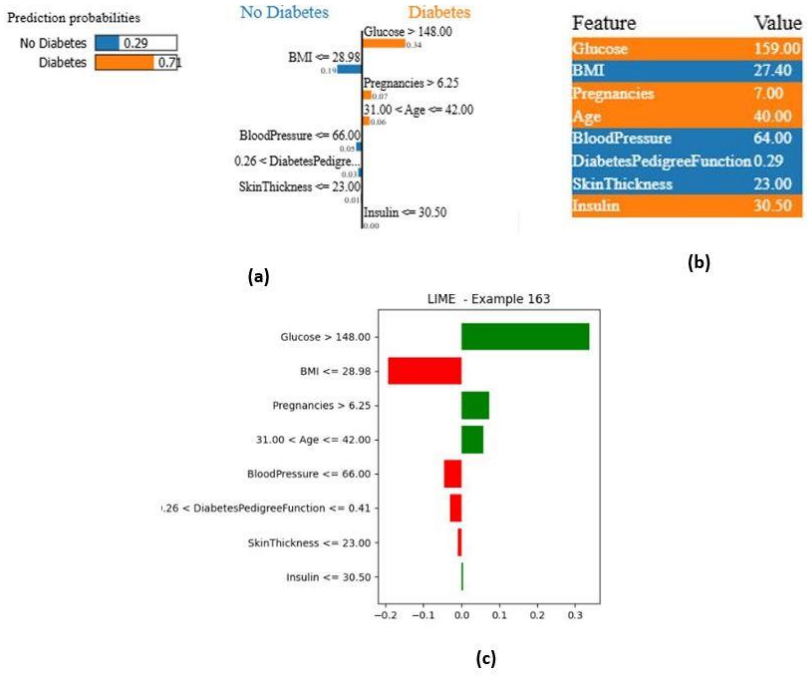

Figure 4. LIME analysis for Patient 163: **(a)** Prediction probabilities showing 71% diabetes likelihood, **(b)** Patient's clinical feature values, and **(c)** Feature contribution plot demonstrating that Glucose >148 increases risk by 34%, while BMI <28.98 reduces risk by 19%, with Age (31-42) and Pregnancies (>6.25) contributing 6% and 7% respectively.

According to the LIME results on patient number 163, BMI, Glucose level, Age, and Number of Pregnancies significantly influence the risk of diabetes. A BMI below 28.98 contributes to a 19% reduction in risk, while a BMI above this threshold is related with a higher risk of diabetes. Glucose level is the most influential factor, with a 34% effect; high glucose levels (above 148) substantially increase the risk of diabetes. The age range of 31-42 years contributes to a 6% increase in risk, and the number of pregnancies above 6.25 has a 7% effect on diabetes risk. DiabetesPedigreeFunction and other factors such as Skin Thickness and Insulin level have less significant contributions to the overall risk. These findings suggest that while glucose level and BMI are the strongest predictors, age and pregnancy history also play a role, though to a lesser extent.

2.4. Model Selection

Model selection in data mining is critical for identifying the best model for a given task. This process often involves evaluation criteria that measure accuracy, generalizability, and efficiency. Commonly used metrics include accuracy, the F1-score, cross-validation performance, and the area under the ROC curve (AUC-ROC). It is also important to consider model complexity and parameter tuning to avoid overfitting or underfitting. In addition, the efficiency and speed of the model play a significant role in the selection process.

A confusion matrix is a valuable tool for visually comparing the performance of different models. Table 4 provides a confusion matrix for a binary classification problem, with the corresponding results detailed below:

- True Positive (TP): Positive units correctly classified as positive
- False Positive (FP): Units classified as positive but actually negative.
- True Negative (TN): Negative units correctly classified as negative.
- False Negative (FN): Units classified as negative but actually positive.

Table 4. Confusion matrix for the binary classification problem

CONFUSION MATRIX		**ACTUAL VALUE**	
		Positive: Diabetes	Negative: No Diabetes
PREDICTED VALUE	Positive: Diabetes	True Positive (TP)	False Positive (FP)
	Negative: No Diabetes	False Negative (FN)	True Negative (TN)

Based on the matrix presented in Table 4, the corresponding performance metrics are calculated. The performance measures provided sequentially are accuracy, precision, recall (sensitivity), F1-score, and ROC curve.

Accuracy: Defined as the fraction of instances that are correctly classified relative to the total number of classified instances. It is commonly used as a measure of prediction performance in classification problems. However, using accuracy alone can be misleading, so employing multiple performance metrics provides a more comprehensive evaluation. $Accuracy = \frac{TP+TN}{TP+FP+TN+FN}$

Precision: Shows the ratio of truely classified positive ratios to the total number of predicted positive ratios. $Precision = \frac{TP}{TP+FP}$

Recall (Sensitivity): Calculates the model's ability to correctly identify real positive instances in the test set. It is calculated as the proportion of true positive rates to the total number of real positive rates. $Recall = \frac{TP}{TP+FN}$

F1-Score: Since working with two different metrics can sometimes be problematic, researchers may prefer a single number to measure performance. F1-score is calculated as a combination of recall and precision. $F1 - Score = \frac{2*Recall*Precision}{Recall+Precision}$

ROC Curve: The Receiver Operating Characteristic (ROC) curve is a probability curve for different classes. The x-axis shows the false positive rate, while the y-axis shows the true positive predictions. This curve illustrates how well the classifier performs in making predictions. The area under the curve (AUC) takes values in the range of [0,1] and serves as a measure of model performance. A higher AUC value reflects a stronger capability of the model to differentiate among the classes within the dataset. In binary classification problems, an ROC curve visualizes the balance between the model's performance to correctly identify positive units and its tendency to misclassify negative units as positive. In summary, ROC curves are highly useful visual tools for comparing classification methods [24].

2.5. Machine Learning Methods

Classification problems aim to determine the class of a new observation based on the analysis of a given dataset. Various methods can be used for solving classification problems. In this study, machine learning methods frequently used in the literature for diabetes diagnosis are considered.

Logistic Regression (LR): A statistical model used in binary classification problems where the dependent variable is discrete. It is widely applied in computer science, various applied sciences, and real-world problems. Logistic regression explains the relation between a binary dependent variable and a set of independent variables using a logistic function (logit function) [26]. The probability of an event occurring is given by Equation 2. The logit function is calculated as Equation 3 [27].

$$P\ (Y) = \frac{e^{\beta_0+\beta_1 X}}{1+ e^{\beta_0+\beta_1 X}} = \frac{1}{1+ e^{-(\beta_0+\beta_1 X)}} \tag{2}$$

$$ln\ \left(\frac{P(Y)}{1-P\ (Y)}\right) \tag{3}$$

Multi-Layer Perceptron (MLP): An MLP is a fundamental type of artificial neural network with a feedforward architecture, typically composed of at least one hidden layer. It includes an input layer, one or more hidden layers, and an output layer, all of which are fully connected. MLPs utilize nonlinear activation functions such as ReLU, Sigmoid, and Tanh to catch complex patterns and relationships. They are trained through the backpropagation algorithm, and their weights are updated using optimization techniques based on gradient descent. MLPs are extensively used for classification, regression, and time series prediction tasks, particularly when addressing nonlinear problems [28].

Extreme Gradient Boosting (XGBoost): XGBoost is a high-performance, optimized gradient boosting algorithm. It builds upon the gradient boosting framework originally proposed by Friedman [29] and and includes various enhancements that make it especially effective for large datasets and high-dimensional feature spaces. XGBoost has shown superior performance across a variety of applications [30].

Extra-Trees (ET): Extra-Trees is an ensemble learning method developed by Geurts [31] as an extension of the Random Forest algorithm proposed by Breiman [20]. It constructs multiple decision trees while increasing randomness to prevent overfitting. Unlike Random Forest, Extra-Trees selects split points randomly, leading to faster tree construction. This approach reduces computational cost while maintaining generalization performance. Extra-Trees is used in classification, regression, and feature importance evaluation, and it performs well on noisy datasets [31].

K-Nearest Neighbors (KNN): KNN is a non-parametric approach introduced by Fix and Hodges [32] and commonly employed for both classification and regression tasks. In classification, it determines a label by looking at the K nearest neighbors and choosing the class that appears most frequently among them. Since KNN searches for the closest neighbors for each data point, the computational complexity increases as the dataset grows. As KNN is distance-based, normalizing the training data can significantly improve its accuracy.

Gaussian Naive Bayes (GNB): GNB is a type of Naive Bayes classifier that assumes a Gaussian (Normal) distribution for continuous data. The algorithm relies on Bayes' Theorem to compute conditional probabilities and make class predictions. GNB assumes that features are independent (the Naive assumption), making it simple, fast, and computationally efficient. Since GNB considers features to be normally distributed, it estimates class probabilities using the mean and variance of each feature. GNB is particularly effective for high-dimensional datasets and small sample sizes. It is widely applied in text classification, medical diagnosis, fraud detection [33].

Kolmogorov-Arnold Networks (KAN): Kolmogorov-Arnold Networks (KANs) represent a novel type of neural network that adopts a basically different learning paradigm compared to conventional MLPs. In contrast to MLPs, which rely on fixed activation functions at the nodes (or "neurons"), KANs place adaptable activation functions on the edges. Although this might seem like a minor adjustment, it substantially boosts both performance and interpretability. In KANs, each weight is replaced by a univariate function-often modeled as a spline-thus eliminating the use of linear weights altogether. The nodes in a KAN basically gather incoming signals without applying any nonlinear transformations. The foundation of KAN lies in the Kolmogorov–Arnold representation theorem, which asserts that any continuous multivariate function can be represented as a finite combination of univariate continuous functions and addition operations. Specifically, Kolmogorov-Arnold theorem [48] expresses a continuous function $f(x_1, x_2, \ldots, x_n)$ as in Equations 4.

$$f(x) = f(x_1, x_2, \ldots, x_n) = \sum_{q=1}^{2n+1} \Phi_q \left(\sum_{P=1}^{n} \emptyset_{q.p} \left(x_p \right) \right) \tag{4}$$

where Φ_q:$\mathbb{R} \rightarrow \mathbb{R}$ and $\emptyset_{q.p}$:$[0,1] \rightarrow \mathbb{R}$ are univariate continuous functions. Building on this principle, the KAN replaces conventional neuron activations with learnable spline-based univariate functions, enabling it to approximate nonlinear mappings with fewer parameters and improved interpretability. For a given input vector $X = [x_1, x_2, \ldots, x_N]$, the KAN output is defined as in Equations 5.

$$\hat{y} = \sum_{i=1}^{M} w_i \Phi_i \left(\sum_{j=1}^{N} \psi_{i,j} \left(x_j \right) \right) \tag{5}$$

Here Φ_i and $\psi_{i,j}$denote the nonlinear spline transformations learned during training, while w_i are trainable weights analogous to those in conventional neural networks. M and N represent the model width and input dimension, respectively. By leveraging spline-based activation functions and a reduced parameterization, the KAN achieved high generalization performance while maintaining interpretability. This balance between expressive power and transparency makes KAN particularly suitable for AI-assisted clinical decision support systems, where model explainability is critical [34].

2.5.1 Artificial neural network approach with additive and multiplicative neurons (AMNN)

Deep Learning-Based Alternative Model: Neural Network Approach with Additive and Multiplicative Neurons (AMNN): Traditional artificial neural networks (ANNs) are primarily based on additive processing, where each neuron computes a weighted sum of its inputs and applies an activation function.

However, this structure may fall short in modeling complex nonlinear relationships in certain applications. To address this limitation, more flexible models have been developed that integrate both additive and multiplicative neurons. AMNN offer an enhanced ability to learn richer representations by employing both summation (∑) and multiplication (∏) operations on the inputs. Specifically, multiplicative neurons are beneficial for modeling interactions between inputs. This approach helps overcome the limited linear combination capability of traditional neural networks, enabling the model to capture more intricate relationships. Models utilizing both additive and multiplicative neurons have been developed as an alternative to conventional ANNs, aiming to achieve more effective results in specific problems by allowing the network to learn more complex dependencies between input features [45, 46]. As shown in Figure 5, the model includes two parallel computational pathways. (1) a linear component based on McCulloch–Pitts neurons (Σ, f_1), and (2) a nonlinear component based on multiplicative neurons (Π, f_1). The activation functions for these components are defined as Equations (6-7).

$$f_1(x) = x, \qquad f_2(x) = \frac{1}{1+e^{-x}} \tag{6-7}$$

The hidden layer computes outputs for both components as Equations (8-11).

$$net_1 = \sum_{j=1}^{m} w_{1j} y_{t-j} + b_1, \qquad o_1 = f_1(net_1) = (net_1) \tag{8-9}$$

$$net_2 = \sum_{j=1}^{m} w_{2j} y_{t-j} + b_{2j}, \qquad o_2 = f_1(net_2) = \frac{1}{1+\exp(-net_2)} \tag{10-11}$$

Finally, the outputs of both paths are combined in the output layer as shown in Equation (12-13).

$$\hat{y} = w_{31} o_1 + w_{32} o_2 + b_3 \tag{12-13}$$

This formulation allows the AMNN to learn both additive and multiplicative dependencies within the same framework, where w_{31} and w_{32} are adaptively learned to balance the contribution of linear and nonlinear components during training. Such a structure provides a more expressive approximation capability than conventional additive-only neural networks. As illustrated in the model architecture diagram in Figure 5, the network comprises two parallel paths originating from the input layer. The first path includes a conventional neuron that performs a weighted summation (i.e., linear combination), while the second path consists of a computational unit that multiplies all input features together. These two pathways are subsequently merged through summation to generate a unified representation. A hidden layer with ReLU activation follows this combined output, and finally, a sigmoid-activated output neuron is employed for binary classification.

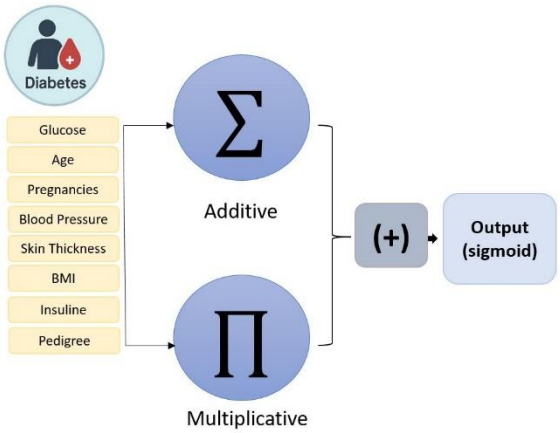

Figure 5. Architectural schematic of the proposed Additive and Multiplicative Neurons Network (AMNN). The architecture integrates two parallel computational pathways: (i) an additive neuron path, which performs a weighted summation of the input signals, and (ii) a multiplicative neuron path, which computes the product of all input features. The outputs of these two streams are subsequently combined and passed through a hidden layer with ReLU activation, followed by a sigmoid-activated output layer for binary classification.

3. RESULTS

In this study, the performance of various machine learning algorithms used in diabetes classification was compared. Within the scope of the analysis, 70% of the dataset was allocated for training and 30% for testing, followed by the application of 5-fold cross-validation. Different algorithms such as LR, KNN, MLP, XGB, ET, GNB, KAN and AMNN were evaluated in terms of accuracy, precision, recall, F1-score, and the AUC-ROC. MLP models were tested with different configurations. For instance, MLP-2 used a two hidden layer (100 and 50 neurons), the ReLU activation function, and the "Adam" optimization algorithm. In the KAN model, spline-based activation functions and 10 hidden layer units were employed. In the AMNN model, a hybrid deep learning architecture was constructed that combines additive and multiplicative neurons to better capture both linear and nonlinear relationships. The model consisted of two parallel processing paths: a summation path using ReLU-activated dense neurons (16 units) and a multiplicative path using a custom Lambda layer. These two paths were merged and followed by a hidden layer with 8 neurons (ReLU) and a sigmoid output neuron for binary classification. The AMNN model was compiled using the binary cross-entropy loss function and Adam optimizer, and trained over 100 epochs with a batch size of 32. Performance evaluation was conducted on the test dataset, with metrics such as accuracy, precision, recall, F1-score, and AUC-ROC being calculated. A summary of the performance comparison of all models is presented in Table 5.

Table 5. Performance metrics of selected machine learning algorithms (LR: Logistic Regression, MLP: Multi-Layer Perceptron, XGboost: Extreme Gradient Boosting, ET: Extra Trees, KNN: K-Nearest Neighbors, GNB: Gaussian Naive Bayes, KAN: Kolmogorov–Arnold Networks, AMNN: Additive and Multiplicative Neurons.)

Classification Report	Accuracy	Precision	Recall	F1-score	AUC-ROC
LR	0.7316	0.7386	0.7316	0.7342	**0.8347**
MLP-2	0.7359	0.7482	0.7359	0.7398	0.7888
XGboost	0.7489	0.7520	0.7489	0.7503	0.8332
ET	0.7489	0.7463	0.7489	0.7474	0.8286
KNN	0.7489	**0.7755**	0.7489	0.7542	0.8028
GNB	0.7273	0.7354	0.7273	0.7302	0.8015
KAN	0.7272	0.7185	0.7272	0.7186	0.8225
AMNN	**0.7576**	0.7733	**0.7576**	**0.7618**	0.8206

Examining Table 5 reveals that the AMNN model exhibited the best overall performance among the listed models, achieving the highest accuracy (0.7576), recall (0.7576), and F1-score (0.7618). It also delivered a strong precision score (0.7733) and a competitive AUC-ROC value of 0.8206. These results highlight AMNN's effectiveness in classification tasks, particularly its ability to balance sensitivity and specificity. The model's high precision suggests strong control over false positives, which is valuable in medical diagnostics. This success is likely due to AMNN's unique hybrid architecture, combining additive and multiplicative neurons for capturing both linear and nonlinear data relationships.

Following AMNN, the XGBoost, ET, and KNN models formed a mid-performance group, all attaining an accuracy of 0.7489. Among them, KNN stood out with the highest precision (0.7755), indicating reliable positive class identification. XGBoost and ET had competitive F1-scores (0.7503 and 0.7474, respectively), and solid AUC-ROC values above 0.82, reflecting good classification capacity.

The MLP-2 model, with an accuracy of 0.7359 and AUC-ROC of 0.7888, demonstrated moderate performance. LR (Logistic Regression) showed similar trends, achieving an AUC-ROC of 0.8347, the highest among all models, despite its lower accuracy (0.7316), highlighting its strength in distinguishing classes across thresholds.

Models such as GNB, KAN, and LR ranked at the bottom in terms of accuracy and F1-score, though their AUC-ROC values remained above 0.80. KAN, for instance, had a relatively low accuracy (0.7272), but its AUC-ROC (0.8225) suggests stable probability predictions, making it potentially useful in threshold-sensitive applications.

To further illustrate the discriminative performance of the models, Figure 6 presents the ROC curves, generated using the [51], for four representative classifiers: LR, XGBoost, KAN, and the proposed AMNN. The figure clearly demonstrates the superior area under the curve of the AMNN model compared with the conventional and ensemble-based approaches.

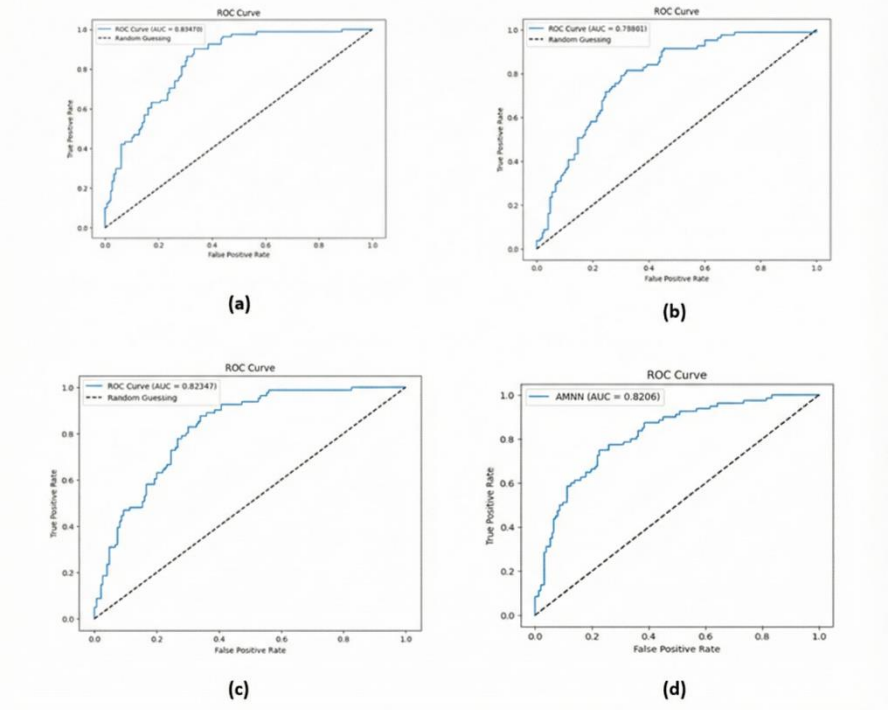

Figure 6. Receiver Operating Characteristic (ROC) curves of the selected classification models: (a) Logistic Regression (LR), (b) XGBoost, (c) Kolmogorov-Arnold Network (KAN), and (d) the proposed Additive and Multiplicative Neurons Network (AMNN). The figure highlights the superior discriminative performance of the proposed AMNN model compared to conventional and ensemble-based approaches.

Overall, these findings suggest that the AMNN model is a competitive alternative to traditional classifiers, excelling in both accuracy and precision. It is particularly well-suited for domains where minimizing false positives is critical. Ensemble-based models and tree-based learners (like XGBoost and ET) also proved effective, while simpler models like LR and GNB, despite lower classification metrics, still offer value in terms of interpretability and AUC-based ranking.

4. DISCUSSION

This study sought to evaluate the diagnostic performance of various machine learning algorithms for diabetes, with the aim of determining which model proved most effective. Using the PIMA Indians Diabetes dataset, the analysis revealed that the AMNN models achieved the highest performance. These models outperformed others in terms of accuracy and AUC-ROC scores, suggesting that deep learning-based approaches and ensemble models are highly effective for diabetes classification.

The application of XAI methods such as SHAP and LIME provided valuable insights into model decision-making processes. These methods highlighted that glucose level, BMI, and age are the most critical features in predicting diabetes. Among these, glucose level was found to be the strongest predictor, significantly influencing the model's ability to classify individuals as diabetic or non-diabetic. This finding aligns with clinical evidence, as elevated blood glucose levels are a primary diagnostic criterion for diabetes and play a central role in disease progression [35]. BMI, another key factor, is strongly associated with insulin resistance and metabolic dysfunction, making it a crucial indicator for early intervention [36]. Age-related changes in glucose metabolism and pancreatic function further emphasize the importance of age as a predictive variable in diabetes risk assessment [37]. Other important features included Diabetes Pedigree Function and the number of pregnancies. A strong family history of diabetes, as indicated by the Diabetes Pedigree Function, suggests a genetic predisposition, which is critical for identifying at-risk individuals before clinical symptoms appear [38]. Additionally, the number of pregnancies may reflect the impact of gestational diabetes, a known risk factor for type 2 diabetes in later life [39]. Recognizing these risk factors through machine learning models allows for earlier lifestyle modifications, targeted screening programs, and personalized treatment plans [40,41]. However, features like blood pressure, skin thickness, and insulin levels had a lower impact on the model’s predictions. While these factors may not be the strongest predictors, they still contribute to the overall assessment of metabolic health and should not be overlooked in clinical practice [42].

These findings are largely consistent with previous studies in the literature. For instance, Kayaer and Yıldırım [4] achieved 80.21% accuracy using a GRNN on the same dataset, while Patra and Khuntia [9] improved the KNN classifier to achieve 83.2% accuracy. Similarly, Mousa et al. [10] found that LSTM achieved the highest accuracy (85%) among various models applied to the PIMA dataset. In this study, although the LSTM model was not implemented, a number of classical and hybrid machine learning models were tested, including ensemble and deep learning-based approaches.

According to the experimental results, the best performance was achieved by the AMNN model, with an accuracy of 75.76%, a recall of 75.76%, and an F1-score of 76.18%. These metrics demonstrate a competitive performance compared to several previously reported models. Furthermore, ensemble models such as XGBoost and ET (Extra Trees) also showed promising results, with accuracy values of 74.89%, aligning with the findings of Yangın [6] and Chang et al. [11], who emphasized the strength of ensemble-based methods like XGBoost and RF. Additionally, KNN achieved a solid precision of 77.55%, which is consistent with the enhanced version presented by Patra and Khuntia [9].

On the other hand, logistic regression (LR) yielded a relatively lower accuracy of 73.16% but reached the highest AUC-ROC value of 0.8347, indicating a good balance between sensitivity and specificity. This result also resonates with the findings by Chang et al. [11], where interpretable models like Naïve Bayes and decision trees were found to be effective with fewer features.

In summary, the findings of this study confirm the effectiveness of neural and ensemble learning approaches for diabetes classification and reinforce the idea that hybrid models like AMNN can provide robust performance across multiple evaluation metrics. The consistency of these results with prior research supports the continued use and development of such models in medical diagnostic tasks involving structured datasets like PIMA.

Despite these promising results, the study has certain limitations. This study draws on the Pima Indians Diabetes dataset (n = 768), composed solely of adult women of Pima descent. While this dataset provides a well-characterized sample, its narrow demographic scope and modest size may introduce spectrum bias and limit the generalizability of the findings to populations beyond Pima women. Additionally, the dataset is imbalanced, with fewer diabetic cases compared to non-diabetic ones. To address this, the SMOTE method was applied. However, it is important to evaluate how such oversampling techniques might affect the model's ability to generalize to real-world scenarios. Future research could benefit from larger and more balanced datasets to enhance model reliability and robustness.

Another limitation is that the PIMA dataset includes only female participants, which prevents an analysis of gender-based differences in diabetes prediction. Future studies could incorporate datasets that include both male and female individuals to examine potential gender-based variations in diabetes risk factors.

To further improve generalizability, future validation should involve multi-center and demographically diverse cohorts that include different ethnicities, age groups, and clinical settings. In addition, nested cross-validation and external test evaluations are recommended to ensure model stability and reproducibility across heterogeneous populations.

Moreover, this study incorporated both traditional machine learning algorithms and deep learning models, including a novel hybrid approach (AMNN). For future research, exploring more advanced hybrid models, transfer learning strategies, or federated learning techniques could further enhance diabetes classification performance. Additionally, implementing these models in real-world clinical decision support systems could help validate their practical applicability.

5. CONCLUSION

This study aims to improve diabetes classification using various machine learning models, while also identifying the most influential variables through explainable AI (XAI) and feature selection methods. A key innovation of the study is the implementation of the AMNN model, a novel artificial neural network (ANN) architecture that outperformed all other models in diabetes prediction. In addition, the study explores the KAN model, another recent ANN approach, which offers a distinctive design through learnable edge-based activation functions.

AMNN's SHAP-based local explanations can support threshold-based decisions by providing clinicians with patient-specific risk factors. Real-time feature extraction and missing data imputation will be critical for integration with electronic health records (EHRs). In addition, ensuring model interpretability, calibration across diverse patient populations, and seamless workflow integration represent key challenges for clinical adoption. Future implementations could involve embedding AMNN-driven alerts or risk dashboards within clinical decision support systems, enabling clinicians to validate model outputs and adjust thresholds based on institutional or demographic contexts.

The experimental results demonstrate that the AMNN model achieved the highest classification performance, with an accuracy of 75.76%, recall of 75.76%, and F1-score of 76.18%, outperforming other models in overall predictive capability. Although LR (Logistic Regression) recorded the highest AUC-ROC value (0.8347), indicating strong discriminatory power, the overall balanced performance of AMNN across all metrics makes it the most effective model in this study.

Ensemble models like XGBoost and ET (Extra Trees) also showed competitive results, confirming their robustness in medical classification tasks. While the KAN model did not surpass AMNN, it achieved a respectable accuracy of 72.72% and AUC-ROC of 0.8225. Its innovative use of learnable activation functions at the edges instead of traditional fixed-node activations offers greater adaptability and interpretability, marking it as a promising direction for future studies.

In terms of feature importance, both feature selection and XAI techniques consistently highlighted glucose level, BMI, and age as the most significant predictors of diabetes. Additionally, Diabetes Pedigree Function and number of pregnancies were found to contribute meaningfully to classification outcomes, which aligns with established medical knowledge.

Despite these encouraging findings, some limitations remain. The dataset is relatively small, includes only female participants of Pima Indian heritage, and is class-imbalanced, which may limit generalizability. Future research should aim to address these limitations by utilizing larger, more diverse datasets and integrating advanced techniques such as hybrid learning, transfer learning, or federated learning to improve model accuracy and robustness.

In conclusion, the results confirm that machine learning, particularly deep learning and ensemble-based methods like AMNN and XGBoost, hold great potential for early diabetes diagnosis. The integration of XAI techniques enhances model transparency, facilitating trust and interpretability in clinical decision-making. Moreover, innovative models like KAN warrant further exploration, especially in hybrid configurations, to advance diabetes prediction and support personalized healthcare applications.

Building upon the promising results achieved by AMNN and KAN, future research could explore hybrid models that integrate both architectures to leverage their complementary strengths. Specifically, combining the additive-multiplicative capabilities of AMNN with the nonlinear transformation power of KAN could lead to a more robust model that captures even more complex patterns and relationships in data. Additionally, investigating the application of federated learning could expand the models' applicability, enabling decentralized training across multiple institutions or devices while preserving data privacy. Such an approach would facilitate the development of scalable, privacy-preserving solutions, particularly in sensitive fields like healthcare, where data sharing is often limited. Future efforts should also focus on model interpretability in these hybrid settings, ensuring that the combined models remain transparent and understandable to end-users, especially in high-stakes domains such as clinical decision-making.

CONFLICT OF INTEREST

The authors stated that there are no conflicts of interest regarding the publication of this article.

CRediT AUTHOR STATEMENT

Şeyda Demirel Tatlı: Methodology, Software, Formal Analysis, Writing – Original Draft, Writing – Review & Editing, Visualization. **Kürşad Aytekin**: Methodology, Writing – Review & Editing, **Melih Agraz:** Conceptualization, Methodology, Software, Formal Analysis, Writing – Original Draft, Writing – Review & Editing, Visualization.

REFERENCES

[1] Temurtas H, Yumusak N, Temurtas F. A comparative study on diabetes disease diagnosis using neural networks. Expert Syst Appl 2009; 36(4): 8610-8615.

[2] Başer BÖ, Yangın M, Sarıdaş ES. Classification of diabetes mellitus with machine learning techniques. Süleyman Demirel University, J Nat Appl Sci 2021; 25(1): 112-120.

[3] Kaggle. PIMA Indians diabetes database [Internet]. 2024. Accessed December 2024. Available from: https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database?select=diabetes.

[4] Kayaer K, Yıldırım T. Medical diagnosis on Pima Indian diabetes using general regression neural networks. In: Proceedings of the international conference on artificial neural networks and neural information processing (ICANN/ICONIP) 2003; 181–184.

[5] Karatsiolis S, Schizas CN. Region based Support Vector Machine algorithm for medical diagnosis on Pima Indian Diabetes dataset. In: 2012 IEEE 12th International Conference on Bioinformatics & Bioengineering (BIBE) 2012; 139-144.

[6] Yangın G. Application of XGBoost and decision tree based algorithms on diabetes data (Master's thesis). Institute of Science, Mimar Sinan Fine Arts University, İstanbul, Turkey, 2019; Available from: https://hdl.handle.net/20.500.14124/1152

[7] Sankar Ganesh PV, Sripriya P. A comparative review of prediction methods for Pima Indians Diabetes dataset. Comput Vis Bio-Inspired Comput 2019; 735-750.

[8] Lakhwani K, Bhargava S, Hiran KK, Bundele MM, Somwanshi D. Prediction of the onset of diabetes using artificial neural network and Pima Indians Diabetes dataset. In: 2020 5th IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE) 2020; 1-6.

[9] Patra R, Khuntia, B. Analysis and prediction of Pima Indian Diabetes Dataset using SDKNN classifier technique. In: IOP Conference Series: Materials Science and Engineering. IOP Publishing 2021; 1070(1): p. 012059.

[10] Mousa A, Mustafa W, Marqas RB, Mohammed SH. A comparative study of diabetes detection using the Pima Indian Diabetes database. J Duhok University 2023; 26(2): 277-288.

[11] Chang V, Bailey J, Xu QA, Sun Z. Pima Indians diabetes mellitus classification based on machine learning (ML) algorithms. Neural Comput Appl 2023; 35(22): 16157-16173.

[12] Farsana KS, Poulose A. Hybrid convolutional neural networks for PIMA Indians diabetes prediction. In: 2024 Fifteenth International Conference on Ubiquitous and Future Networks (ICUFN). IEEE. 2024; 268-273.

[13] Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 2002; 16: 321-357.

[14] Douzas G, Bacao F, Last F. Improving imbalanced learning through a heuristic oversampling method based on k-means and SMOTE. Inf Sci 2018; 465: 1-20.

[15] Kursa MB, Rudnicki WR. Feature selection with the Boruta package. J Stat Softw 2010; 36: 1-13.

[16] Kursa MB, Jankowski A, Rudnicki WR. Boruta–a system for feature selection. Fundam Inform 2010; 101(4): 271-285.

[17] Zhao Z, Anand R, Wang M. Maximum relevance and minimum redundancy feature selection methods for a marketing machine learning platform. In: 2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA). IEEE. 2019; 442-452.

[18] Peng H, Long F, Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans Pattern Anal Mach Intell 2005; 27(8): 1226-1238.

[19] Chen XW, Jeong JC. Enhanced recursive feature elimination. In: Proceedings of the Sixth International Conference on Machine Learning and Applications (ICMLA). IEEE. 2007; 429-435.

[20] Breiman L. Random forests. Mach Learn 2001; 45: 5-32.

[21] Agraz M. Comparison of feature selection methods in breast cancer microarray data. Med Rec 2023; 5(2): 284-9.

[22] Agraz M, Deng Y, Karniadakis GE, Mantzoros CS. Enhancing severe hypoglycemia prediction in type 2 diabetes mellitus through multi-view co-training machine learning model for imbalanced dataset. Sci Rep 2024; 14(1): 22741.

[23] Ribeiro MT, Singh S, Guestrin C. "Why should I trust you?" Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 2016; 1135-1144.

[24] Kuhn M, Johnson K. Applied predictive modeling. New York: Springer. 2013.

[25] Ho TK. Random decision forests. In: Proceedings of the 3rd International Conference on Document Analysis and Recognition 1995; 278-282.

[26] Tatlıdil H. Applied multivariate statistical analysis. Academy Printing House, Ankara, 2002; 167.

[27] Kalaycı Ş. SPSS applied multivariate statistics techniques. Ankara, Asil Publishing, 2016.

[28] Rumelhart DE, Hinton G. E, Williams RJ. Learning representations by back-propagating errors. Nature 1986; 323(6088): 533-536.

[29] Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat 2001; 1189-1232.

[30] Chen T, Guestrin C. Xgboost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 2016; 785-794.

[31] Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. Mach Learn 2006; 63: 3-42.

[32] Fix E, Hodges JL. Discriminatory analysis: nonparametric discrimination, small sample performance. Air University, USAF School of Aviation Medicine 1952.

[33] John GH, Langley P. Estimating continuous distributions in Bayesian classifiers. In: Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence 1995; 338-345.

[34] Liu Z, Wang Y, Vaidya S, Ruehle F, Halverson J, Soljačić M, Hou TY, Tegmark M. KAN: Kolmogorov-Arnold networks. arXiv preprint 2024; 19756.

[35] Higgins C. Diagnosing diabetes: blood glucose and the role of the laboratory. Br J Nurs 2001; 10(4): 230-236.

[36] Barber TM, Kyrou I, Randeva HS, Weickert MO. Mechanisms of insulin resistance at the crossroad of obesity with associated metabolic abnormalities and cognitive dysfunction. Int J Mol Sci 2021; 22(2): 546.

[37] Naito H, Kaga H, Someya Y, et al. Fat accumulation and elevated free fatty acid are associated with age-related glucose intolerance: Bunkyo Health Study. J Endocr Soc 2023; 8(2): bvad164.

[38] Nyakairu Doreen G. The Impact of Genetic History on the Risk of Developing Type II Diabetes. Res Output J Biol Appl Sci 2024; 4(1): 51-57.

[39] Diaz-Santana MV, O'Brien KM, Park YM, Sandler DP, Weinberg CR. Persistence of risk for type 2 diabetes after gestational diabetes mellitus. Diabetes Care 2022; 45(4): 864-870.

[40] Cubillos G, Monckeberg M, Plaza A, Morgan M, Estevez PA, Choolani M, Kemp MW, Illanes SE, Perez CA. Development of machine learning models to predict gestational diabetes risk in the first half of pregnancy. BMC Pregnancy Childbirth 2023; 23(1): 469.

[41] Kumar M, Ang LT, Ho C, et al. Machine Learning-Derived Prenatal Predictive Risk Model to Guide Intervention and Prevent the Progression of Gestational Diabetes Mellitus to Type 2 Diabetes: Prediction Model Development Study. JMIR Diabetes 2022; 7(3): e32366.

[42] Lai H, Huang H, Keshavjee K, Guergachi A, Gao X. Predictive models for diabetes mellitus using machine learning techniques. BMC Endocr Disord 2019; 19(1): 1-9.

[43] Agraz M, Goksuluk D, Zhang P, Choi BR, Clements RT, Choudhary G, Karniadakis GE. ML-GAP: machine learning-enhanced genomic analysis pipeline using autoencoders and data augmentation. Frontiers in Genetics 2024; 15: 1442759.

[44] Mak KK, Wong YH, Pichika MR. Artificial intelligence in drug discovery and development. Drug discovery and evaluation: safety and pharmacokinetic assays 2024; 1461-1498.

[45] Yolcu U, Egrioglu E, Aladag ÇH. A new linear & nonlinear artificial neural network model for time series forecasting. Decision Support Systems 2013; 54(3): 1340-1347.

[46] Valenca M, Ludermir T. Multiplicative-additive neural networks with active neurons. In IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No. 99CH36339) IEEE, 1999; 6, 3821-3823.

[47] Tatlı ŞD, Yakut SG. Determination of factors affecting university students' happiness levels through decision trees analysis. Journal of Awareness 2024; 9(2), 237-250.

[48] Kolmogorov AN. On the representations of continuous functions of many variables by superposition of continuous functions of one variable and addition. In: Dokl. Akad. Nauk USSR 1957; 953-956.

[49] Ong KL, Stafford LK, McLaughlin SA, Boyko EJ, Vollset SE, Smith AE, Brauer M. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. The Lancet 2023; 402(10397), 203-234.

[50] Topşir A, Güler F, Çetin E, Burak MF & Agraz M. Thyroid disease classification using generative adversarial networks and Kolmogorov-Arnold network for three-class classification. BMC Medical Informatics and Decision Making, 2025, 25(1), 284.

[51] Agraz M, Mantzoros C & Karniadakis GE. ChatGPT-Enhanced ROC Analysis (CERA): A shiny web tool for finding optimal cutoff points in biomarker analysis. PLOS ONE, 2024; 19(4), e0289141.