International Journal of Future Engineering Innovations

Predicting Treatment Adherence in Patients with Type 2 Diabetes Using Interpretable Machine Learning: A Study among the Iraqi Population

Yahya Qasim Ibrahim Al-Fadhili
Department of Computer Science, College of Education for Pure Science, University of Mosul, 41002, Mosul, Nineveh, Iraq

* Corresponding Author: **Yahya Qasim Ibrahim Al-Fadhili**

Article Info

ISSN (online): 3049-1215
Impact Factor (RSIF): 8.25
Volume: 03
Issue: 04
Received: 23-04-2026
Accepted: 25-05-2026
Published: 27-06-2026
Page No: 11-22

Abstract

Background: Adherence to treatment is required to ensure the best glycemic control and avoidance of complications associated with diabetes among patients with type 2 diabetes. This study aims to predict treatment adherence among Iraqi patients with type 2 diabetes using interpretable machine learning algorithms to identify the key factors influencing adherence.

Methods: This was a cross-sectional study done in 2026 to come up with predictive models of treatment adherence among patients with diabetes in Mosul, Iraq. The study sample used was 850 subjects and comprised demographic, clinical, and Trans-Theoretical Model construct variables. Machine learning algorithms and data analysis were done using the Random Forest, Support Vector Machine (SVM), and Multi-Layer Perceptron (MLP) after preprocessing the data. The dataset was further stratified into training and testing groups and the performance of the model was measured in terms of Accuracy, Precision, Recall, F1-score, and the Area Under the Receiver Operating Characteristic Curve (AUC). Moreover, the SHAP approach was used to determine the most significant predictors of treatment adherence to enhance the interpretability of the model.

Results: The best predictive model out of the developed models was the Random Forest algorithm, with an accuracy of 60.6, a sensitivity (recall) of 82.2, and an F1-score of 71.2 on the test data. Cross-validation of the Random Forest model gave a hyper parameter optimization of 0.676. The interpretability analysis based on the SHAP algorithm revealed that the most significant predictors of adherence to treatment were HbA1c, body weight, age, income, and self-efficacy and perceived barrier constructs.

Conclusion: These results suggest that machine learning algorithms, especially decision tree-based models like Random Forest, with the use of model interpretability strategies, have significant potential in predicting treatment adherence in diabetic patients.

DOI: https://doi.org/10.54660/IJFEI.2026.3.4.11-22

Keywords: Treatment Adherence, Type 2 Diabetes, Iraqi Population, Machine Learning

Introduction

Treatment Adherence as the extent to which a patient adheres to the health behaviors prescribed by the healthcare team, treatment adherence is among the most basic pillars in effective management of chronic diseases and a major predictor of avoiding any secondary complications in the healthcare systems [1]. It is essential in the realisation of the best metabolic control, lowering mortality rates and improving the quality of life of patients. As per the instructions of the World Health Organization (WHO) [2] the treatment of diabetes mellitus as one of the most significant healthcare issues of the current century presupposes an overall dedication of the patient to the medication treatment process, frequent tests of blood glucose level, changes in lifestyle, and adherence to dietary patterns. Treatment adherence is a complex phenomenon that involves extended treatment persistence, health behavior change, frequent doctor visits, and self-management skills during disease fluctuations; as a result, it is regarded as one of the most crucial phenomena to assess the efficacy of treatment in diabetes patients [3].

Over the last few years, though, therapeutic non-adherence has become one of the most acute issues of healthcare systems in every country of the globe, and the rate of poor adherence among diabetic patients is reported to be between 30 and 60 percent in most countries [4, 5]. Non-adherence is directly correlated with higher glycemic variability, high glycated hemoglobin (HbA1c), and early onset of microvascular (retinopathy, nephropathy, neuropathy) and macrovascular (myocardial infarction, and stroke) complications. This eventual results in high rates of preventable hospitalization, and a significant financial strain on the health economy, and a significant decrease in health-related quality of life (HRQoL) [6, 7].

A myriad of factors, such as individual, demographic, and socioeconomic factors, as well as the perceived benefits and barriers of adherence, affect treatment adherence. Although these factors have been identified, the interaction of these factors are too complex, multidimensional and are mostly non-linear and thus the traditional statistical methods are not sufficient to explain and reveal these hidden interactions accurately [8].

Out of the different sub-groups, patients with Type 2 diabetes form one of the groups most vulnerable as far as non-adherence is concerned. Their therapeutic behaviors are under great risk due to the existence of co-morbidities, supportive dependency or environmental stressors. This question gains a two-fold importance in the framework of Iraqi society. The last several years were marked by special conditions of this population with the structural changes, economic issues, blistering changes in lifestyle (physical inactivity and poor nutrition) and the lack of access to primary healthcare services. Stresses related to the high treatment costs, social fears, inability to obtain constant monitoring, stresses of everyday life, and even changes in patterns of social support are all reasons that significantly increase the risk of deteriorating treatment adherence in Iraqi patients [9, 10].

Despite many studies performed in the field, most historical studies have been based on conventional statistical techniques, including linear and logistic regressions, which have a weak ability to determine the complex and non-linear relationships between variables. Additionally, detailed research that will assess the simultaneous presence of clinical, demographic and psychological factors have not been reported in this particular target population. Conversely, machine learning algorithms have become widely used because of their capability to identify complex patterns and dramatically increase predictive power; most of the algorithms do not provide enough interpretability to clinicians, as they are inherently black-box. This limitation is overcome by the strategic application of Interpretable Machine Learning (IML). In addition to creating very precise models, it facilitates the clear and concrete explanation of what each variable contributes, where it is heading, and its importance in the prediction of treatment adherence, making the results very practical in clinical practice [11, 12].

Therefore, the main gap in the knowledge in this area is the lack of an accurate, interpretable, and machine learning-based predictive model to determine the determinants of treatment adherence among diabetic patients in the investigated population. This research intends to address this scientific gap by using the interpretable machine learning algorithms to determine the most important predictors of adherence behaviors besides quantifying the relative predictive contributions of the baseline variables, decisional balance (perceived pros and cons), and self-efficacy depending on the constructs of the Transtheoretical Model (TTM). Based on this, the research question of the given study is as follows:

The question is, to what degree can the variables of the baseline, the decisional balance (cons and pros), and self-efficacy predict adherence to treatment in patients with diabetes mellitus, who use interpretable machine learning algorithms?

The response to this question can open the doors to the development of intelligent early-screening systems, prediction of patients who are likely to fail treatment before the onset of clinical complications, designing specific educational interventions, as well as efficient allocation of resources. Moreover, the results of the current study can provide an evidence-based scientific framework to the Iraqi Ministry of Health, practicing endocrinologists, health psychologists, and healthcare policy-makers to design a specific supportive program, and provide a new paradigm to future studies of chronic disease management.

Materials and Methods

Design and Sample

The study was a cross-sectional research carried out in 2026 in Mosul, Iraq, in order to create and test an interpretable machine learning model to predict patient treatment adherence in diabetes. A multistage method of sampling was used. During the initial phase, healthcare facilities offering diabetes services and follow-up to patients in various regions of Mosul were chosen randomly. Afterward, patients with known diabetes that were registered and attended regularly at the chosen centers were recruited on the basis of the pre-established eligibility criteria. Overall, 850 patients with diabetes were involved in this study.

The inclusion criteria were a confirmed diagnosis of diabetes, aged 20-65 years, access to the chosen healthcare facilities, and capacity to fill out the instruments of the study, and written informed consent to participate. The exclusion criteria were acute illnesses or chronic disease that might severely limit daily activity (i.e. advanced cardiovascular disease, severe musculoskeletal diseases or other conditions with special medical needs), pregnancy, physician-established medical restrictions and unwillingness or withdrawal at some point of the study.

A structured data collection instrument was used to collect the data of the participants. The former consisted of demographic and clinical variables, such as age, sex, height, weight, family size, income status, educational level, family history of diabetes, gym membership, exercise status, glycated hemoglobin (HbA1c) level, and adherence status to treatment. Heights and weights were recorded and the body mass index (BMI) was determined using the standard formula: weight/height (kg/m 2).

Treatment adherence was the main outcome measure of the research and was measured with the application of the adherence assessment tool. Using the scores obtained, participants were divided into two groups optimal adherence and poor adherence. This factor was thought of as the target variable when developing machine learning models and the other demographic and clinical factors were taken as the input features to train the machine learning models.

The second part of data collection instrument was composed of psychological constructs relying on the Transtheoretical Model (TTM) that assesses cognitive and behavioral variables that relate to self-care behaviors. This chapter had three primary sub-elements perceived barriers (14 items), perceived benefits (29 items), and self-efficacy (9 items). Each of the items was rated on a 4-point Likert scale as 1-4 where higher scores reflected more perceived benefits of health-promoting behaviors, less perceived barriers, and increased self-efficacy in the management of diabetes. Validity and reliability of the instrument were already proved with Content Validity Index (CVI) of 0.78 and Cronbachs alpha coefficient of 0.89.

To create machine learning models, the dataset was first analyzed in terms of data quality, missing elements, and possible inconsistencies. The Synthetic Minority Oversampling Technique (SMOTE) was used on the training data only to eliminate the issue of class imbalance to avoid information leakage and ensure model performance is valid. A stratified train-test split method was then used to split the dataset into training (80) and testing (20) subsets.

The three machine learning algorithms such as the Random Forest (RF), Support Vector Machine (SVM), and Multilayer Perceptron (MLP) were created to predict treatment adherence. The hyperparameter optimization was done by a combination of grid search and cross-validation (5-fold CV). The model performance has been assessed through various classification measures such as Accuracy, Precision, Recall, F1-score and Area Under the Receiver Operating Characteristic Curve (AUC-ROC).

In addition, to enhance the transparency of the model and improve the clinical interpretability, SHapley Additive exPlanations (SHAP) framework was used to measure the contribution of the features, determine the most significant predictors, and describe how the developed machine learning models made a decision.

Variables and Data Collection Instruments

The structured instrument to be used in gathering the data was comprised of two main sections. The demographic and clinical characteristics of the participants were included in the first section. The variables measured in this section were age, sex, height, weight, income status, family size, educational status, family history of diabetes, gym membership, physical activity status, the duration of diabetes and the glycated hemoglobin (HbA1c) level. The height and weight of the participants were taken to determine the Body Mass Index (BMI) based on the following standard formula:

$$BMI = \frac{Weight(kg)}{Height(m)^2}$$

The main outcome variable of the study was treatment adherence which was measured using the corresponding adherence assessment instrument. Through the scores of adherence obtained, the subjects were grouped into two categories: adherent and non-adherent. This is a binary variable that was used as the target variable to develop and evaluate the machine learning models.

Definition of the Target Variable (Treatment adherence)

This paper has established treatment adherence as the target variable to develop machine learning-based predictive models. The adherence status was calculated by the score of the corresponding adherence assessment tool. Based on previously established classification criteria, the participants were divided into two categories: treatment adherent (Class 1) and non-adherent to treatment (Class 0).

In the last dataset, out of 850 participants enrolled, 503 (59.2) out of the patients were considered to be adherent to treatment, and 347 (40.8) were considered non-adherent. This binary outcome variable was used as the dependent variable in the machine learning models. The aim of the algorithms was to be trained on the underlying patterns of the demographic, clinical, and psychological characteristics of patients and predict the likelihood of treatment adherence in people with diabetes.

Data Preprocessing and Machine Learning Framework

Firstly, the raw dataset was analyzed to identify the quality problems such as missing values, outliers, and possible inconsistencies to come up with the predictive models. The variables included in the modeling process were demographic, clinical, laboratory, and psychosocial variables. Following the data cleaning, the predictor set and the target variable (treatment adherence) were ready to develop the model. Since the target classes were not evenly distributed, Synthetic Minority Over sampling Technique (SMOTE) was only used on the training set to alleviate bias in the algorithms towards the major class. Creating synthetic minority samples, SMOTE allowed to increase the balance of classes and to learn meaningful patterns by the models. Notably, the method of SMOTE was used solely on the training set to avoid information leakage to the test set.

After preprocessing, the dataset was split into training and test sets with a stratified traintest split, where the original proportions of classes in the dataset were maintained in both test and training sets. At least 80 percent of the data were used to train the model and 20 percent to test the independent performance. Three machine learning models were created and compared- Random Forest, Support Vector Machine (SVM) and Multilayer Perceptron (MLP).

In order to maximize the performance of the model, hyperparameter tuning was performed with Grid Search + 5 fold cross validation, and the best parameterization chosen based on the maximum AUC ROC achieved during cross validation. Accuracy, Precision, Recall, F1 score, and Area Under the ROC Curve (AUC ROC) was used to determine model performance on the independent test set. Also, to increase the interpretability of the final model, SHAP (SHapley Additive exPlanations) was used to measure the relative importance of the predictors and determine the most influential factors of treatment adherence.

Hyperparameter Optimization

Hyperparameter tuning was conducted by a grid search method with five-fold cross-validation to attain optimal performance of every machine learning model. The region beneath the receiver operating characteristic curve (ROC-AUC) of the validation folds was taken as the main parameter of choosing the best hyperparameter settings. Table 1 contains the optimized hyperparameter settings of each model.

Table 1: Optimized hyperparameters of machine learning models

Model	Optimized parameters	Best CV-AUC
Random Forest	n_estimators=300, max_depth=None, min_samples_split=5	n_estimators=300, max_depth=None, min_samples_split=5
SVM	C=50, kernel=RBF, gamma=scale	C=50, kernel=RBF, gamma=scale
MLP	hidden layers=(64,64,32), alpha=0.001, solver=Adam	hidden layers=(64,64,32), alpha=0.001, solver=Adam

Machine Learning Modeling and Evaluation

Three machine learning algorithms, such as Random Forest (RF), Support Vector Machine (SVM), and Multi-Layer Perceptron (MLP), were also created and compared in order to predict treatment adherence. Random Forest algorithm was chosen among the main models because it is highly effective in the nonlinear relationships, complex interactions between predictors and the risk of overfitting is minimized. A Support Vector Machine model that uses a Radial Basis Function (RBF) kernel was applied to determine intricate decision frontiers among adherence categories. Moreover, a Multi-Layer Perceptron neural network was created in order to evaluate its capabilities of learning nonlinear and complicated relationships between the input features and the desired output.

Hyperparameter tuning was done with Grid Search and fivefold cross-validation (5-fold CV) to optimize model performance. The best combination of hyperparameters to fit each model was chosen based on the largest area under the receiver operating characteristic curve (AUC-ROC) in cross-validation.

The evaluation metrics were used to test the performance of the developed models on an independent test dataset and included the following metrics:

- Accuracy
- Precision
- Recall (Sensitivity)
- F1-score
- Area Under the Receiver Operating Characteristic Curve (AUC-ROC)

The F1-score was calculated using the following equation:

$$F1\text{-score} = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Explainable Artificial Intelligence (SHAP Analysis)

SHapley Additive exPlanations (SHAP) was used to improve the interpretability of the developed machine learning models and determine the most important factors in predicting treatment adherence. SHAP offers an interpretable model to quantify the role of each feature in the model prediction and also calculates how far individual variables can add or subtract to the predicted probability of treatment adherence.

The average absolute SHAP values were computed to determine the relative importance of the input features, and the most significant predictors used to make the final model predictions were selected. The SHAP outputs were plotted as bar plots and summary plots and used to give the information about the importance of each feature in general and about the direction and strength of its influence on the model outputs.

Computational and Statistical Environment

The statistical and computational analyses were all done in two complementary software environments. IBM SPSS Statistics version 26 was used to carry out descriptive and inferential statistical analysis. To summarize the demographic, clinical and psychological characteristics of the participants, descriptive statistics such as frequency, percentage, mean, standard deviation, minimum and maximum values were computed. The use of appropriate statistical tests based on the type and measurement scale of the variables was done to determine the relationships between the study variables and to prepare the dataset to be further used in the predictive modeling.

To create and test machine learning-based predictive models, the entire process of data preprocessing, feature selection, and model training, hyperparameter optimization, and performance evaluation were carried out in Python version 3.12. The primary computational packages were Pandas and NumPy to manipulate and preprocess data, Scikit-learn to implement machine learning algorithms and to evaluate models, Imbalanced-learn to deal with class imbalance with Synthetic Minority Over-sampling Technique (SMOTE), and SHAP (SHapley Additive exPlanations) to interpret and quantify predictor effects.

The dataset was split into training and testing subsets randomly after data cleaning and preprocessing in a stratified splitting procedure to maintain the initial class distribution. Standard measures were used to assess the model performance in terms of Accuracy, Precision, Recall (Sensitivity), F1-score, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC). As a control of overfitting and increase the model generalizability, five-fold cross-validation was used in the process of hyperparameter optimization and final model selection. Lastly, SHAP analysis was conducted to explain the results of the chosen model in numerical terms, that is, to estimate the value of each predictor to treatment adherence classification.

Ethical Considerations

The study was carried out with the agreement of the appropriate institutional ethics committee and was carried out with regards to the ethical principles provided in the Declaration of Helsinki on medical and health research with human subjects. All participants were educated on the objectives, procedures, possible benefits and potential limitations associated with the study before being enrolled in the study. All the participants gave written informed consent before participation.

The participants were aware of their right to withdraw at any point without any negative effects. All data collected were sent in confidence and the information of participants was kept in anonymized identification codes in order to protect their privacy. No personal data that could be directly identified on the analytical dataset was provided and only members of the research team, who were authorized, were able to access the raw data.

Data confidentiality and privacy principles were upheld during the computational process and creation of machine learning models. The data gathered were only used in conducting scientific research and all the findings were reported in an aggregate manner without any information that could result in identification of participants. All analysis and processing of data had been adequately recorded so as to

improve reproducibility and scientific validity of the results.
Results

Demographic and Clinical Characteristics of the participants

In total, there were 850 patients with diabetes who were analyzed. Table 1 gives the demographic and clinical profile of the participants. The average age of the participants was 36.64 years old with a standard deviation of 6.45 years. The mean values of height, weight, and HbA1c were 164.52 ± 8.86 cm, 56.51 ± 9.47 kg, and 5.86 ± 0.54%, respectively.

The average of the psychological constructs such as perceived barriers, benefits, and self-efficacy were 39.80, 54.48, and 23.99 with standard deviations of 6.91, 13.33, and 5.38, respectively.

On the target variable, 503 participants (59.2) were observed to be having poor treatment adherence, and 347 participants (40.8) were observed to have optimal treatment adherence. Table 1 shows the distribution of demographic variables such as sex, marital status, family size, educational level, and family history of diabetes.

Table 2: Demographic and clinical characteristics of participants (N=850)

Variable	Category/Unit	n (%) / Mean ± SD
Age (years)	Mean ± SD	36.64 ± 6.45
Height (cm)	Mean ± SD	164.52 ± 8.86
Weight (kg)	Mean ± SD	56.51 ± 9.47
Income	Mean ± SD	2.37 ± 0.84
HbA1c (%)	Mean ± SD	5.86 ± 0.54
Barrier score	Mean ± SD	39.80 ± 6.91
Benefit score	Mean ± SD	54.48 ± 13.33
Self-efficacy score	Mean ± SD	23.99 ± 5.38
Sex	man	420 (49.4%)
	women	430 (50.6%)
Marital status (mch)	YES	445 (52.4%)
	NO	405 (47.6%)
Number of family members (nfm)	1	5 (0.6%)
	2	12 (1.4%)
	3	69 (8.1%)
	4	292 (34.4%)
	5	301 (35.4%)
	6	105 (12.4%)
	≥7	66 (7.8%)
Education level	Level 1(bisavad)	110 (12.9%)
	Level 2(sikl)	266 (31.3%)
	Level 3(diplom)	239 (28.1%)
	Level 4(lisans)	145 (17.1%)
	Level 5(foqe lisans)	69 (8.1%)
	Level 6(doktora)	21 (2.5%)
Family history of diabetes	YES	202 (23.8%)
	No	648 (76.2%)
Treatment adherence (Target variable)	Poor adherence (0)	503 (59.2%)
	Good adherence (1)	347 (40.8%)

Target Variable Distribution and Class Balancing

The pre-intervention analysis of the target variable distribution showed the disproportion of treatment adherence classes. In particular, 503 participants (59.2%) were identified as the ones with the optimal treatment adherence, and 347 participants (40.8%), as having the poor treatment adherence. Even though the extent of imbalance between the classes was not that high, the Synthetic Minority Oversampling Technique (SMOTE) was only applied to the training data to minimize the possibility of the model being biased toward the majority.

Following the SMOTE application, the training set was balanced in terms of classes, with 402 samples per class. This equal training sample was then employed in training and testing the machine learning models.

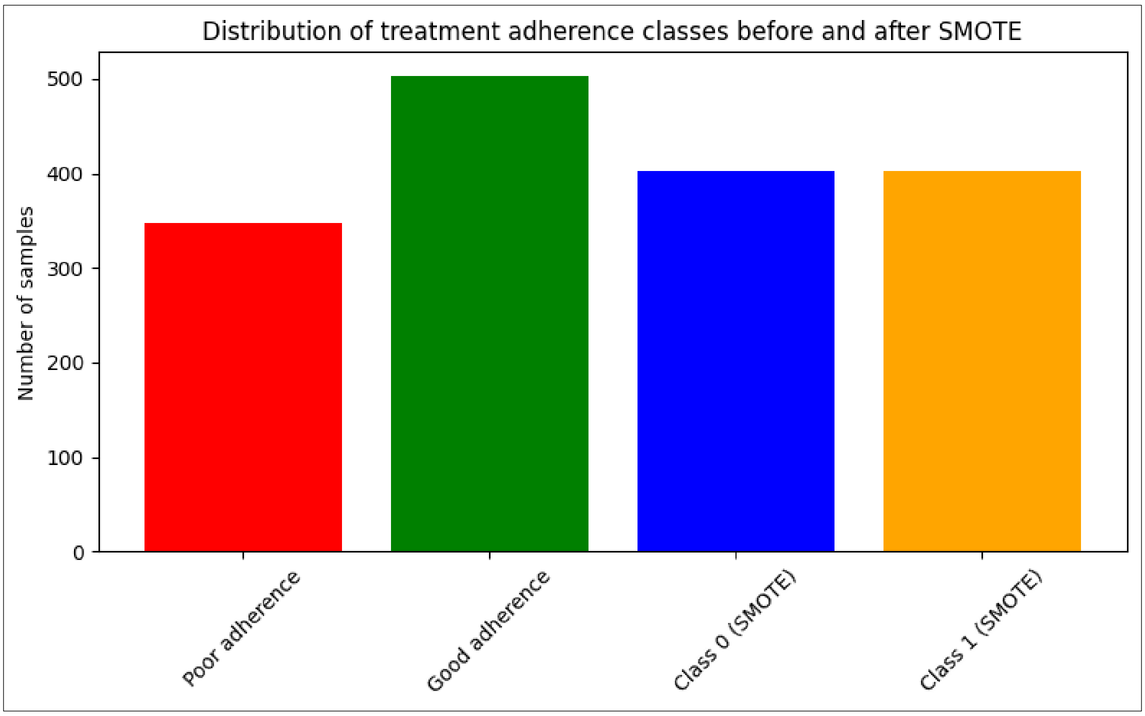

Fig 1: Distribution of treatment adherence classes before and after SMOTE

Performance of Machine Learning Models

After data preprocessing, the dataset was stratified and split into training (80%) and testing (20%) subsets, using stratified splitting approach. Three machine learning algorithms (RF, SVM, and Multi-Layer Perceptron (MLP)) were created and tested. Accuracy, Precision, Recall, F1-score and Area Under the Receiver Operating Characteristic Curve (AUC-ROC) were used to estimate model performance.

Table 2 shows that the Random Forest model has the highest overall Recall and F1-score than any of the assessed algorithms. The RF model achieved a Recall of 82.18% and an F1-score of 71.24%. All the developed models in terms of performance metrics, Accuracy, Precision, Recall, F1-score, AUC-ROC values are summarized in Table 3.

Table 3: Final performance comparison of machine learning models on the test set

Model	Accuracy	Precision	Recall	F1-score	AUC
Random Forest	0.606	0.629	0.822	0.712	0.516
SVM	0.565	0.626	0.663	0.644	0.526
MLP	0.547	0.618	0.624	0.621	0.591

The results revealed that the Random Forest (RF) model had the best overall performance in predicting treatment adherence and had a more balanced performance in classifying the poor and optimal adherence groups than the other algorithms that were tested.

Machine Learning Models Confusion Matrix Analysis

Confusion matrices were also used to assess the performance of the developed machine learning models in terms of classification in predicting treatment adherence. The three tested algorithms (Random Forest (RF), Support Vector Machine (SVM), and Multi-Layer Perceptron (MLP)) produced confusion matrices to show the distribution of correctly and incorrectly classified cases between the two treatment adherence groups (poor adherence and optimal adherence).

The confusion matrices were very informative in terms of the number of true and false classifications of each model and were utilized in order to evaluate the class-specific prediction performance. Figure 2 shows the results of the confusion matrix analysis.

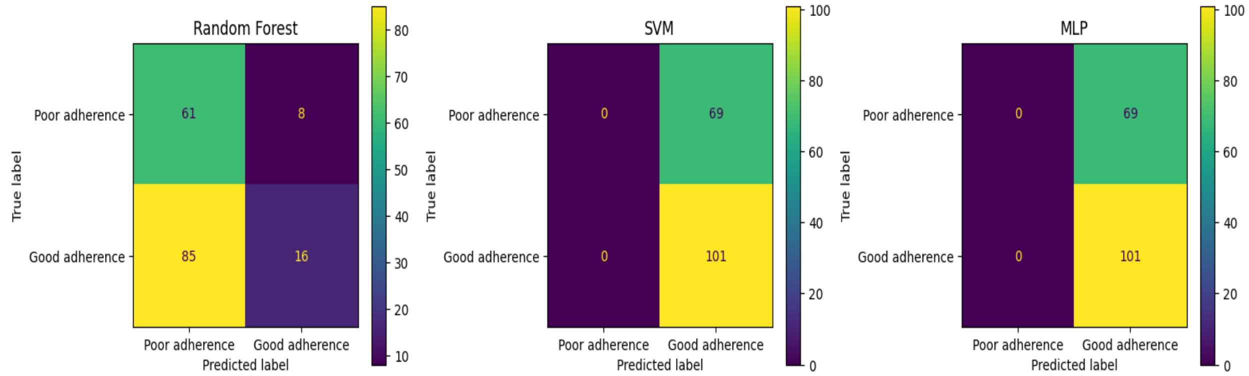

Fig 2: Confusion matrices of Random Forest, SVM, and MLP models for treatment adherence classification

Hyperparameter Optimization and Model Optimization

The Hyperparameters were optimized through the use of a Grid Search that had a 5-fold cross-validation. The findings showed various optimization performance between the assessed machine learning models in predicting performance of treatment adherence.

The best setting of the Random Forest (RF) model was: 300 trees, no maximum depth, and five samples minimum to split a node, with a cross-validation AUC (CV-AUC) of 0.675. Support Vector Machine (SVM) model with RBF kernel gave its best performance with C = 50 and gamma = scale, having a CV-AUC = 0.684. In the case of the Multi-Layer Perceptron (MLP) model, the best architecture was three hidden layers (64, 64, and 32 neurons), a ReLU activation and the Adam optimizer with a CV-AUC of 0.591. Table 3 presents the optimized hyperparameter settings and the corresponding cross-validation AUC of all the models.

Table 3: Optimized hyperparameters of machine learning models

Model	Optimized parameters	CV-AUC
Random Forest	n_estimators=300, max_depth=None, min_samples_split=5	0.675
SVM	C=50, Kernel=RBF, Gamma=scale	0.684
MLP	Hidden layers=(64,64,32), ReLU, Adam, Alpha=0.001	0.591

Comparison of Models and ROC Curve Analysis

In order to further assess the discriminative performance of developed models and compare predictive ability of the models on treatment-adherence classification, receiver operating characteristic (ROC) curves and associated area under the curve (AUC) of all three algorithms were determined. The analysis of the independent test data using ROC revealed that the Multi-layer perceptron (MLP) model has the highest value of AUC among the tested models with AUC of 0.591. By contrast, the RF and Support Vector Machine (SVM) models also exhibited comparable results in terms of the AUC, similar to values near 0.52.

All three machine learning models have their ROC curves provided in Figure 3, which shows the variations in the discriminative performance of the tested algorithms.

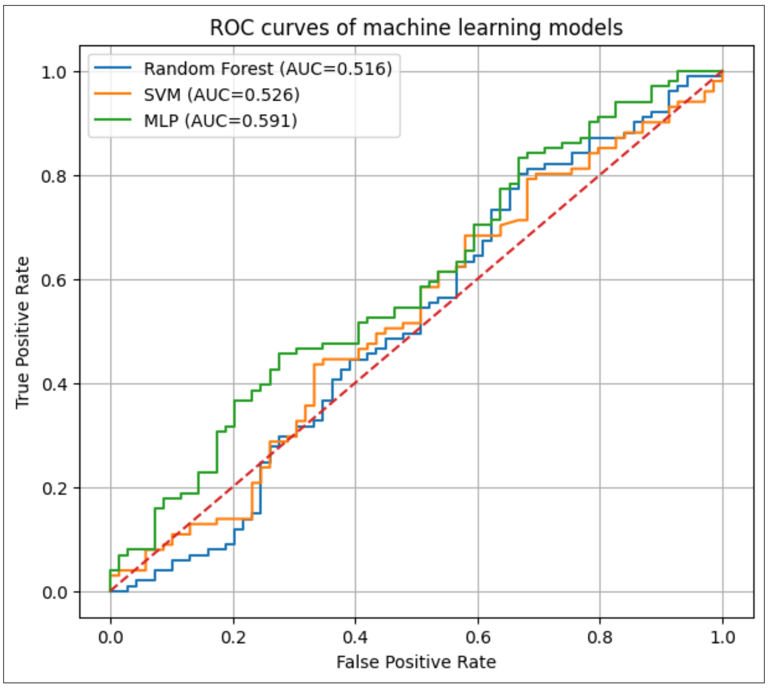

Fig 3: ROC curves of Random Forest, SVM, and MLP models for treatment adherence classification

SHAP Interpretability of a Model

SHapley Additive exPlanations (SHAP) analysis was used to explore the decision-making patterns of the machine learning model and the most influential predictors of treatment adherence. The values of SHAP are used to measure the contribution of each feature to the model output, as well as give an interpretable measure of the importance and direction of influence of features.

The clinical, demographic, and psychological characteristics considered as shown in the SHAP summary plot (Figure 4) had different contributions to the model predictions. The size of the contribution of each feature was calculated by its SHAP value.

According to the average values of the SHAP, the HbA1c was found to be the strongest predictor as it contributed the most to the model output. Body weight, age, and income were other significant predictors and were among the most contributing features. Moreover, some of the self-efficacy items and perceived barriers were found among the predictors that were influential. The entire list of feature importance in terms of SHAP values is shown in Table 4.

Figure 5 depicts the feature importance plot as determined by the mean absolute SHAP values which shows the relative contribution of the predictors that make up the final model.

Table 5: SHAP-based ranking of the most influential predictors of treatment adherence using the Random Forest model

Rank	Predictor	Description	Mean absolute SHAP value
1	HbA1c	Glycemic control indicator	0.0239
2	Weight	Body weight	0.0151
3	Age	Patient age	0.0150
4	Income	Socioeconomic status indicator	0.0112
5	db4	Diabetes self-management item 4	0.0095
6	education	Maternal educational level	0.0086
7	db11	Diabetes self-management item 11	0.0078
8	db2	Diabetes self-management item 2	0.0076
9	db23	Diabetes self-management item 23	0.0070
10	se5	Self-efficacy item 5	0.0069
11	db38	Diabetes self-management item 38	0.0066
12	db36	Diabetes self-management item 36	0.0065
13	db3	Diabetes self-management item 3	0.0064
14	db6	Diabetes self-management item 6	0.0062
15	Barrier total	Perceived barriers score	0.0062

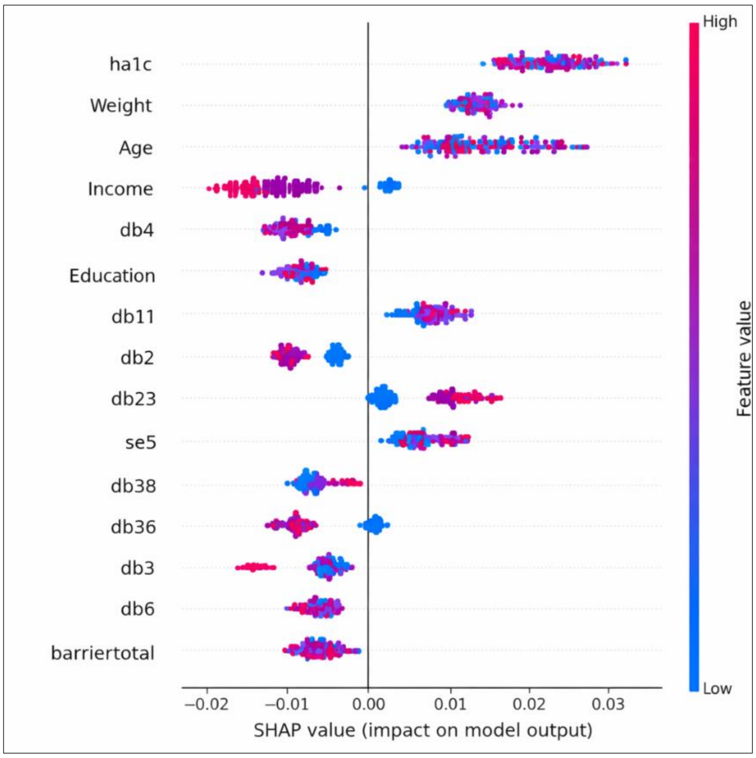

Fig 4: SHAP summary plot illustrating the impact of each predictor on the model’s output

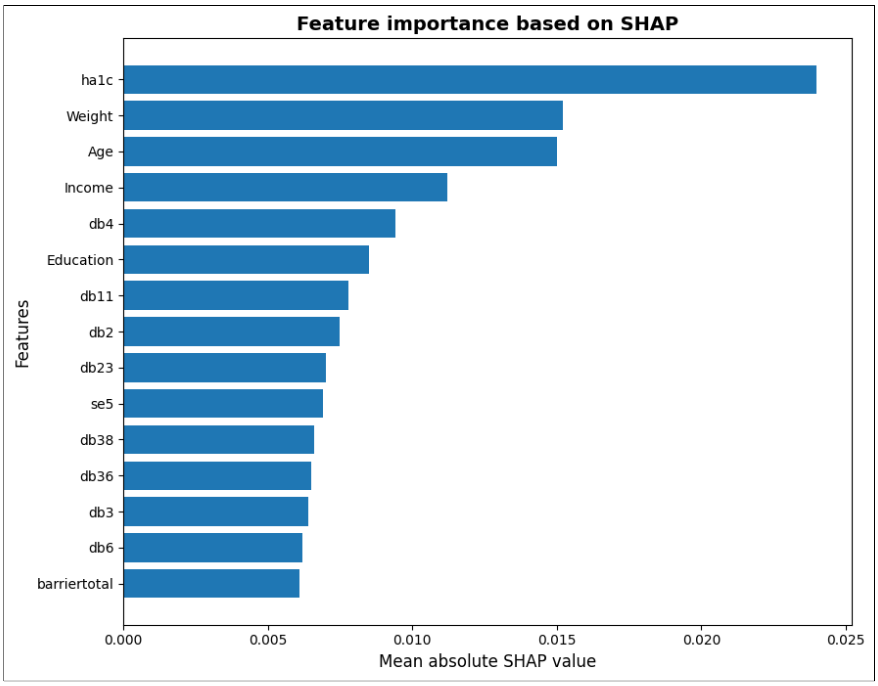

Fig 5: Feature importance based on mean absolute SHAP values

Discussion

The current research was intended to create and test machine learning models to predict the adherence to treatment in patients with diabetes. The results showed that machine learning models are able to detect intricate patterns underlying treatment adherence behaviors by incorporating demographic, clinical, and psychological variables together. The most promising balance of the sensitivity, F1-score, and AUC was realized with the Random Forest model, which showed the best overall performance following the optimization of the hyper parameters. These results imply that the interconnections between the variables affecting the treatment adherence are largely nonlinear and comprise complex correlations between numerous variables, which can not be effectively represented with the help of traditional statistical methods, including logistic regression, which presupposes the linear relationships between the variables [13, 14]. The results of the current research are aligned with the recent data on the use of machine learning in diabetes management. According to these studies, it has been found that tree-based algorithms, especially the Random Forest and Gradient Boosting, are always more effective compared to traditional statistical methods and a variety of other machine learning algorithms in a wide scope of diabetes-related prediction problems, such as glycemic control, diabetes complications, and treatment adherence [15, 16]. This better performance has been explained by the fact that they can model nonlinear relationships, can capture complex interactions among predictors, are able to withstand noisy data, and they are less sensitive to multicollinearity. The current results also confirm the existence of the belief that treatment adherence is a multifactorial behavioral outcome with complex relationships between biological, psychological and social outcomes, and therefore nonlinear machine learning methods are very suitable in predictive modeling.

Even though the optimized version of the Random Forest model only showed a moderate discriminative performance (cross-validated AUC = 0.675), its relatively high recall suggests that it can be used to reasonably well identify patients with sufficient treatment adherence. Clinically, predictive models that are being used in screening should be based on sensitivity rather than overall accuracy since their main goal is to identify as many individuals who can be helped as possible through a timely intervention. Thus, although the proposed model is moderately accurate, it has a significant clinical use in the early risk stratification and personalized intervention planning.

Among the most significant results of the research, it was possible to highlight the determination of HbA1c as the best predictor of adherence to treatment. This finding is in line with the earlier research that found that there is a strong correlation between ideal glycemic regulation and increased medication adherence and self-management practices. Patients with lower HbA1c tend to adhere better to prescribed medications, healthier eating habits, more active, and do regular blood sugar monitoring [17, 18]. On the contrary, high HbA1c might indicate the summative impact of less than ideal adherence practices and thus serves as a composite measure of the level of involvement of patients in long-term management of diabetes. The fact that HbA1c is a key element in the prediction model underlines its importance as a metabolic biomarker as well as a behavioral surrogate of overall disease management.

In addition to clinical predictors, SHAP analysis indicated that age, body weight, income, educational attainment, perceived barriers, and self-efficacy are significant in predicting treatment adherence. Such results support the multidimensional aspect of treatment adherence and show that the successful management of diabetes cannot be confined to biomedical factors only [19]. Self-efficacy has always been a very powerful determinant of lasting health behaviors according to behavioral theories and especially the Trans-Theoretical Model (TTM) and Social Cognitive Theory. People who are more confident in their skills to control diabetes are more likely to stick to medications, change their lifestyles to be healthier and actively participate in self-management activities. Similarly, perceived barriers have been known to play a significant role in impeding the treatment adherence process since patients with more financial, psychological, or practical perceived barriers have less tendency to adhere to the prescribed treatment plans consistently [20-22]. The great role of these psychological concepts in the current model implies that the inclusion of behavioral variables in addition to clinical predictors significantly increases the predictive power of machine learning models in the treatment of diabetes.

On the same note, the significance of income and educational level highlights the role of social determinants of health with regard to treatment adherence. The findings of many epidemiological research works indicate that low socioeconomic status is linked to the low access to medical services, expensive medicines, low health literacy, and worse compliance with prescribed treatment. These results reinforce the increasing understanding of the fact that socioeconomic and behavioral determinants must be included in precision medicine and digital health approaches along with traditional clinical variables. Combining these multidimensional factors makes machine learning models give a more detailed and clinically significant depiction of the patient adherence behavior [23, 24].

The analysis of the tested algorithms showed that Random Forest was always more accurate than Support Vector Machine (SVM) and Multi-Layer Perceptron (MLP). The given finding aligns with the prior research that assesses the effectiveness of machine learning methods in diabetes-related prediction tasks. Despite the impressive performance of deep neural networks when they are trained on very large datasets, tree-based ensemble algorithms tend to be more robust, stable, and generalizable in research with comparatively small sample sizes and characterized clinical data. As a result, the choice of algorithms would be determined by the properties of the data available instead of the complexity of the model, and more advanced algorithms do not always provide high-quality predictive results in clinical studies [25, 26].

One of the key advantages of the current research is the combination of predictive machine learning and explainable artificial intelligence (XAI). The black-box nature of most predictive algorithms is one of the main obstacles to implementing artificial intelligence in clinical practice, as clinicians cannot trust them, and they are not easily adopted in everyday practice. Using SHAP, the current study was able to measure direction and magnitude of each variable contribution to individual predictions, greatly enhancing model transparency and interpretability. This interpretable architecture boosts the trust clinicians have in model outputs and makes it easier to integrate machine learning models into clinical decision support systems. Finally, the integration of precise prediction and open interpretation is a crucial milestone on the way to safe, reliable, and clinically viable application of artificial intelligence in the personalized management of diabetes.

Limitation

Although the methodological innovations of the current study have led to the creation of a predictive model that can be interpreted, it has a number of limitations that can be noted when interpreting the findings.

To begin with, the cross-sectional character of the study does not allow the accurate and clear determination of the causal relationships after a period of time between the predictive variables (self-efficacy and decisional balance) and behaviors of adherence to the treatment. Second, measurement of adherence behaviors and psychological variables was based on self-reported questionnaires, so it is possible that they were more prone to cognitive biases, including recall bias or social desirability bias. Third, the limited scope of the study was because of the limitations of the operation that could only be done within certain healthcare centers in Iraq, which would question the overall generalizability of the results to the whole population residing in other cultural and geographical settings of the country. Lastly, the machine learning models used in the current study were trained with single-point baseline data, and the availability of continuous and longitudinal electronic health records (EHR) would have been more representative of dynamic shifts in behavior and clinical variability throughout the course of the study in the modeling process.

Conclusion

The results of this research show that machine learning algorithms, specifically decision tree-based models like Random Forest and model interpretability methods have significant potential in the prediction of treatment adherence in diabetes patients. A combination of clinical, sociodemographic, and psychological data in intelligent predictive models offers a promising alternative to identifying persons at risk of low rates of treatment adherence. The SHAP analysis also helped to increase the transparency of the model, as it was used to determine the most influential predictors, which can increase its potential use in clinical practice. Practically, the designed model can be used as a preliminary screening tool to determine high-risk patients in terms of poor treatment adherence. Early diagnosis of these people may help to carry out special educational work, behavior therapy, and online health care before the onset of severe diabetes complications. With the growing worldwide incidence of diabetes and the necessity to manage the disease over a long period, predictive models based on artificial intelligence can be used as a complementary tool in predicting individual, patient-centred care and clinical decision-making. The proposed model is to be validated in future research with the help of larger, multicenter, and longitudinal research to make it more general. Moreover, the use of other clinically relevant variables, including medication refill history, continuous glucose monitoring, lifestyle habits, and electronic health records, can also be used to improve the model performance. Prospective intervention studies should also explore how explainable machine learning models are practically applied in clinical decision support systems, and how they affect patient adherence and clinical outcomes, along with the use of healthcare resources in the future.

Acknowledgements

The authors wish to thank most of the patients with diabetes that volunteered to take part in this research. This research was made possible by their precious time, collaboration and readiness to share their health experiences.

Consent for publication
Not Applicable

Competing interests
The authors state that there are no conflicts of interest.

Funding
The authors have not received any financial support for the research.

References
1. Dunbar-Jacob J, Zhao J. Medication adherence measurement in chronic diseases: a state-of-the-art review of the literature. Nurs Rep. 2025;15(10):370.
2. Humphrey L, Kansagara D, Qaseem A, Centor RM, DeLong DM, Gantzer HE, *et al*. World Health Organization guidelines on medicines for diabetes treatment intensification: commentary from the American College of Physicians High Value Care Committee. Ann Intern Med. 2018;169(6):398-400.
3. Salamat N, Zulkifli N, Wong Y, Karuppannan M, Wahab M. Multifaceted interventions to enhance patients' adherence to self-monitoring blood glucose: a systematic review. Prim Care Diabetes. 2026;20(3):283-291.
4. Piragine E, Petri D, Martelli A, Calderone V, Lucenteforte E. Adherence to oral antidiabetic drugs in patients with type 2 diabetes: systematic review and meta-analysis. J Clin Med. 2023;12(5):1981.
5. Highton P, Funnell M, Gupta P, Zaccardi F, Lim L, Seidu S, *et al*. Improving medication adherence in type 2 diabetes: strategies for better clinical and economic outcomes. Diabetologia. 2026;69(3):541-556.
6. Evans M, Engberg S, Faurby M, Fernandes J, Hudson P, Polonsky W. Adherence to and persistence with antidiabetic medications and associations with clinical and economic outcomes in people with type 2 diabetes mellitus: a systematic literature review. Diabetes Obes Metab. 2022;24(3):377-390.
7. Sendekie A, Netere AK, Kasahun AE, Belachew EA. Medication adherence and its impact on glycemic control in type 2 diabetes mellitus patients with comorbidity: a multicenter cross-sectional study in Northwest Ethiopia. PLoS One. 2022;17(9):e0274971.
8. Rhudy C, Johnson J, Perry C, Bumgardner C, Wesley M, Fardo D, *et al*. Machine learning approaches to predicting medication nonadherence: a scoping review. Int J Med Inform. 2025;204:106082.
9. Al-Chawishli S, Dizaye K, Azeez S. Measuring diabetic medication adherence and factors that lead to non-adherence among patients in Erbil. Cureus. 2024;16(9):e70397.
10. Mohammed A, Yaseen N, AlWaeli D. Analyzing medication adherence patterns among type 2 diabetes patients in Thi-Qar, Iraq: a cross-sectional study. J Diabetes Res. 2025;2025:6659722.
11. Kiran M, Xie Y, Anjum N, Ball G, Pierscionek BK, Russell D. Machine learning and artificial intelligence in type 2 diabetes prediction: a comprehensive 33-year bibliometric and literature analysis. Front Digit Health. 2025;7:1557467.
12. Khamis A, Abdul F, Dsouza S, Sulaiman F, Khyreim C, Siddig M, *et al*. Appraisal of clinical explanatory variables in subtyping of type 2 diabetes using machine learning models. J Clin Med. 2025;14(18):6548.
13. Chen Y, Nguyen P, Chien C, Hsu M, Liou D, Yang H. Machine learning-based prediction of medication refill adherence among first-time insulin users with type 2 diabetes. Diabetes Res Clin Pract. 2024;207:111033.
14. Battineni G, Sagaro GG, Chinatalapudi N, Amenta F. Applications of machine learning predictive models in the chronic disease diagnosis. J Pers Med. 2020;10(2):21.
15. Ganie SA, Pramanik P, Bashir Malik M, Mallik S, Qin H. An ensemble learning approach for diabetes prediction using boosting techniques. Front Genet. 2023;14:1252159.
16. Ghafoor B, Zaccardi F, Khunti K, Shabnam S. Comparative predictive accuracy of machine learning versus traditional statistical methods for diabetes-related complications: a systematic review and analysis. Diabet Med. 2026:e70398.
17. Krapek K, King K, Warren SS, George KG, Caputo DA, Mihelich K, *et al*. Medication adherence and associated hemoglobin A1c in type 2 diabetes. Ann Pharmacother. 2004;38(9):1357-1362.
18. Almomani MH, Al-Tawalbeh SM. Glycemic control and its relationship with diabetes self-care behaviors among patients with type 2 diabetes in Northern Jordan: a cross-sectional study. Patient Prefer Adherence. 2022;16:449-465.
19. Deng M, Yang R, Zheng X, Deng Y, Jiang J. Artificial intelligence in diabetes care: from predictive analytics to generative AI and implementation challenges. Front Endocrinol (Lausanne). 2025;16:1620132.
20. Obirikorang Y, Acheampong E, Anto EO, Afrifa-Yamoah E, Adua E, Taylor J, *et al*. Nexus between constructs of social cognitive theory model and diabetes self-management among Ghanaian diabetic patients: a mediation modelling approach. PLoS Glob Public Health. 2022;2(7):e0000736.
21. Mostafavi F, Alavijeh FZ, Salahshouri A, Mahaki B. The psychosocial barriers to medication adherence of patients with type 2 diabetes: a qualitative study. Biopsychosoc Med. 2021;15(1):1.
22. Zhang B, Kalampakorn S, Powwattana A, Sillabutra J, Liu G. A transtheoretical model-based online intervention to improve medication adherence for Chinese adults newly diagnosed with type 2 diabetes: a mixed-method study. J Prim Care Community Health. 2024;15:21501319241263657.
23. Hill-Briggs F, Adler NE, Berkowitz SA, Chin MH, Gary-Webb TL, Navas-Acien A, *et al*. Social determinants of health and diabetes: a scientific review. Diabetes Care. 2021;44(1):258-279.
24. Chen Y, Zhou X, Bullard KM, Zhang P, Imperatore G, Rolka DB. Income inequalities in diagnosed diabetes prevalence among US adults, 2001-2018. PLoS One. 2023;18(4):e0283450.
25. Tan K, Seng JJB, Kwan YH, Chen Y, Zainudin SB, Loh D, *et al*. Evaluation of machine learning methods developed for prediction of diabetes complications: a

systematic review. J Diabetes Sci Technol. 2023;17(2):474-489.
26. Shojaee-Mend H, Velayati F, Tayefi B, Babaee E. Prediction of diabetes using data mining and machine learning algorithms: a cross-sectional study. Healthc Inform Res. 2024;30(1):73-82.

How to Cite This Article
Karam A, Al-Fadhili YQI. Prediction of treatment adherence among Iraqi patients with type 2 diabetes using interpretable machine learning algorithms. Int J Future Eng Innov. 2026;3(4):11-22. doi:10.54660/IJFEI.2026.3.4.11-22.