Journal of Basics and Applied Sciences Research (JOBASR)
ISSN (print): 3026-9091, ISSN (online): 1597-9962
Volume 4(2) March 2026
DOI: https://dx.doi.org/10.4314/jobasr.v4i2.4

AI-Driven Diabetes Prediction and Early Diagnosis System Using Wearable Sensors and Real-Time Health Monitoring

Abdulraheem M.[1*], Tomori R.[2], Oladipo I. D.[3], Balogun G. B.[4], Adebayo P. O.[5] & Jimoh H. J.[6]
[1,2,3,4,5&6]Department of Computer Science, Faculty of Communication and Information Sciences, University of Ilorin, Ilorin, 240003, Kwara State, Nigeria.
[*]Corresponding Author Email: muyideen@unilorin.edu.ng

ABSTRACT
Diabetes Mellitus (DM) is a metabolic disorder characterized by high blood sugar levels, which have been persistent for a long period of time, resulting from the way the pancreas functions, either by the release of insufficient amounts of insulin or the proper use of insulin by the body, or a combination of both. In recent times, the global burden of diabetes has increased significantly, with about 463 million people worldwide living with diabetes by the end of 2019, a figure expected to rise to 700 million by the end of 2045. In order to provide a proactive system of health care, the system has been designed to continuously monitor physiological data, use AI for the analysis of health trends, and send real-time alerts for the early detection of diabetes. In the development of the system, a systematic approach was adopted, which integrated wearable sensors, cloud computing, and machine learning techniques. Five-fold cross-validation was carried out, and the models, including Random Forest, Gradient Boosting, Logistic Regression, and SVM, were trained. Accuracy, F1 score, AUC, and confidence were the evaluation metrics for the models, which showed high predictive capabilities, with the highest accuracy of 94.67% obtained by the tree-based models, which were further optimized. Ensemble methods, including a soft voting classifier, were also developed, which showed reliable and robust performance for the early detection of diabetes, indicating that wearable sensors integrated with AI have the potential to revolutionize the health care of diabetes patients.

Keywords:
Diabetes prediction,
Diagnosis, cloud
Computing,
Real-time prediction,
Wearable sensors,
Machine learning

INTRODUCTION

Diabetes Mellitus (DM) is a metabolic condition that affects an individual over a long period of time, characterized by increased levels of blood sugar, which result from issues related to insulin production, insulin function, or a combination of both (Iheagwam & Iheagwam, 2025). The number of individuals living with diabetes worldwide has increased significantly, reaching a total of 463 million in 2019, and is expected to rise to 700 million in 2045 (Alam et al., 2021). These growing underscores the need for efficient approaches to early diagnosis and management.

Reducing complications like cardiovascular diseases, neuropathy, and retinopathy requires early diabetes detection and management (Iheagwam & Iheagwam, 2025). Periodic blood glucose monitoring is the foundation of traditional diabetes detection techniques, which may fail to detect real-time variations in blood glucose levels.

In order to improve diabetes prediction and early detection, wearable sensor technologies and advances in artificial intelligence (AI) present a possible solution (Chaki et al., 2022).

Wearable sensors combined with AI software have the ability to change the way we monitor vital signs, such as blood glucose, heart rate, and activity levels, in real-time (Jiya et al., 2025). These AI models, based on machine learning, have the ability to search through huge amounts of data (Garba et al., 2025), identifying patterns and predicting the risk of diabetes with high levels of accuracy (Chaki et al., 2022). By using wearable technology, diabetes patients can be helped, reducing the burden on the health system and improving patient outcomes (Alam et al., 2021).

Past studies have focused on the use of AI-based self-management and detection tools for diabetes patients, highlighting the potential of AI for early diagnosis and individualized patient plans.

Compared to other statistics, AI methods such as deep learning and neural networks have higher accuracy for the diagnosis of diabetes (Chaki et al., 2022).

However, there are certain challenges to overcome before we can completely rely on AI-based diabetes prediction systems. These include accuracy, data privacy, and their integration into today's healthcare system. For example, to completely rely on AI models, they should be validated against a wide range of relevant data (Alam et al., 2021). Furthermore, to see AI models in today's healthcare system, we should focus on ethics and data privacy.

A significant innovation in medical devices has been the integration of wearable devices and AI-based diabetes prediction and detection. This can revolutionize the way we treat diabetes and can lead to better outcomes and early detection through intelligent data analysis and health monitoring.

It has been estimated that around 600 million people worldwide will be suffering from diabetes around 2040. Type 2 diabetes (T2DM) is a rising global health concern (Gowthami et al., 2024). Current methods of detecting diabetes are invasive and time-consuming and rely on blood tests. Early detection and monitoring are challenging for patients and doctors (Bano et al., 2021).

The key to effective diabetes management is to identify the condition at an early stage and monitor it continuously to avoid major health issues such as nerve damage, kidney problems, and heart-related health issues. These conventional tests, such as the HbA1c and fasting blood glucose tests, are unable to identify the condition at an early stage and are not capable of providing real-time insights into fluctuations in blood sugar levels (Alotaibi et al., 2017). This is where new approaches are required to facilitate effective management of the condition at an early stage. This is where wearable devices with AI are coming into play. These devices are equipped with sensors that can continuously monitor heart rates, blood glucose levels, and other activities. These devices, when connected to AI, can process a large amount of information in real time, enabling effective management of the condition. In the case of proactive management, AI can analyze CGM to predict episodes of low blood sugar before they actually occur (Zhu et al., 2024).

The problem is addressed by filling a notable gap in the research domain, wherein currently, there is no effective AI-based system that is capable of utilizing the information collected through wearable sensors to predict and diagnose diabetes at an early stage, while at the same time providing real-time health monitoring. This is a problem that is of great importance, especially in terms of catching this disease at an early stage, managing it, and controlling the prevalence of diabetes in the world.

The problem to be addressed is the development of an AI-based system that is capable of providing early prediction of diabetes, while at the same time ensuring that real-time health management is offered through the use of wearable sensors. The contributions of this research are:

i. Developing an AI-based system that is capable of providing an early diagnosis of diabetes.
ii. Evaluating the effectiveness of the system.
iii. Benchmarking the system.

Research on AI-driven systems for the prediction and early diagnosis of diabetes using wearable devices and real-time health monitoring has been identified as a significant area of research with immense potential given the global prevalence of diabetes and the need for a continuous and non-invasive approach for the management and diagnosis of the condition (Ahmed et al., 2023; Ahmed et al., 2022). Over the past few years, the advancements made possible through the integration of wearable devices with AI-driven systems and the application of real-time monitoring technologies have transformed the management of diabetes from an invasive approach to a continuous approach (Vettoretti et al., 2020; Kriventsov et al., 2020). The global prevalence of diabetes is estimated at 537 million patients with the condition, with the numbers expected to rise to 784 million patients by 2045 (Ahmed et al., 2023; Khalifa & Albadawy, 2024; Obaidur et al., 2024).

Guleria et al. (2023) investigate how cloud computing, AI, and data science might be integrated. For the purpose of predicting diabetes in women, they suggest a cloud-based platform that makes use of Data Science as a Service (DSaaS). They use SHapley Additive exPlanations (SHAP) to make sure the model's predictions can be understood. On a dataset of 768 patients, the study's Neural Network model for prediction achieved a positive predictive value (PPV) of 79.3% and an accuracy of 77.9%.

A thorough analysis of AI applications in diabetes treatment and prediction is presented by Nomura et al. (2021). The authors analyze several ML studies, showing prediction accuracies ranging from AUC 0.71 to 0.87. Their own research using GB, DTs achieved an AUC of 0.71 and overall accuracy of 94.9% in predicting diabetes onset. However, the current ML models have not been proven to be more effective than the conventional statistical techniques.

In the paper by Curia (2023), the author focuses on the prediction of Type 1 diabetes using ML with an explainable AI approach. The author is working on the development of a clinical decision support tool. The author uses various ML algorithms such as decision trees, deep neural networks, XGBoost, and KNN. The accuracy levels for the DNN and KNN algorithms are 97% and 93%, respectively. However, the possibility of overtraining also indicates the need for explainability, where the author uses the LIME approach to shed light on the predictions by revealing the key factors affecting the outcome. This will help instill confidence among clinicians using AI for patient care. In the paper by Guan et al. (2023), the authors discuss the progress, scope, and challenges associated with AI in the management of diabetes. The authors discuss the scope of digital health technologies for the better management of diabetes among patients. The authors also refer to the FDA-approved AI-based devices and AI-driven prediction models, screening tools, and clinical decision support systems. However, the authors also discuss the challenges associated with AI usage in the management of diabetes and the need to overcome them for the effective use of AI. The focus of the study is on the role of AI in improving diabetic retinopathy (DR) screening through the use of fundus imaging and Optical Coherence Tomography (OCT). In their study published in 2024, Bansal et al. discuss the use of generative AI in the detection of DR, a complication of diabetes that can result in blindness if left untreated. The authors discuss the use of generative AI, such as Generative Adversarial Networks (GANs), in the detection of anomalies in images of the retina, which is affected in cases of diabetic retinopathy.

Despite the improvement in the accuracy of diabetic retinopathy detection through AI, some issues have been noted, such as the interpretability of AI models, data variation, and the computational intensity of AI models.
Mariam et al., (2024), discuss AI-based digital biomarkers in the management of Type 2 Diabetes (T2DM), including the use of AI in the risk assessment, early detection, and treatment of the condition. The authors discuss the use of different AI models in predicting the progression of the disease using clinical data and wearable sensor data. Despite the use of AI in the management of diabetes through the monitoring of patients, some issues have been noted, such as AI bias, data privacy, and regulatory issues.
A strategy to effectively improve the explainability of AI in healthcare, particularly in diabetes diagnosis, is discussed by Wang et al. (2023). In order to better understand and explain the information, the authors use XAI methods such as CART and LIME. This allows for more accessible and understandable information to be obtained using visual methods such as decision rules and heatmaps. This, in turn, increases the confidence of medical practitioners. Nevertheless, this is still a challenge to effectively balance the explainability and accuracy of AI.
In a study conducted by Li et al. (2025), the authors discuss the use of Large Language Models (LLMs) in diabetes education for primary care physicians. In this study, the authors used ten LLMs, including ChatGPT-4.0, Google Gemini, and MedGPT, to answer questions related to diabetes, which were obtained from medical certification exams. In the assessment of diabetes, AI outperformed PCPs and had the highest accuracy.

Two studies published in 2024 examine AI and ML in healthcare. Both studies maintain their concentration on cost, effectiveness, and transparency.
Hu et al. (2024) report on a study comparing AI-based diabetic retinopathy screening and conventional clinician-based grading in primary care in Australia. They consider different deployment scenarios and use a Markov model. They report that AI-based diabetic retinopathy screening can reduce blindness and lower costs for Indigenous and non-Indigenous people. However, there are challenges to consider in terms of accessibility and validation.

Musacchio et al. (2024) report on a study on the transparency of ML to identify therapeutic inertia in Type 2 diabetes patients on metformon monotherapy. They use the Logic Learning Machine to analyze 1.5 million patient electronic health records. They report on HbA1c pattern identification associated with delays in treatment intensification. They report a good discrimination rate (ROC-AUC 0.81) and identify two types of therapeutic inertia. However, there are limitations to consider in terms of data variability and validation.
A robust predictive diagnosis model for diabetes mellitus is offered by Ejiyi et al. (2023), using a machine learning approach with a boost from SHAP for improved interpretability. Ejiyi et al. perform feature selection on the PIMA dataset to identify the key factors for diabetes diagnosis: glucose levels, BMI, and the diabetes pedigree function. The authors also compare four different tree-based algorithms—Extra Trees, Random Forest, AdaBoost, and XGBoost—to evaluate which one offers the best accuracy. Both XGBoost and AdaBoost algorithms attain the maximum accuracy of 94.67%. However, the challenge with the model is its ability to generalize on other datasets and its usability.
In another interesting work on human-centered machine learning for decision support systems for Type 1 Diabetes (T1D), Stawarz et al. (2023) conducted interviews with 15 patients with Type 1 Diabetes. The patients want AI support for unexpected situations like getting sick or going on a trip; however, they want to use personal heuristics for managing day-to-day activities.
To advance diabetes care with the help of real-time monitoring and AI-based analysis, Huang et al. (2024) introduced the concept of a 5G-based Artificial Intelligence Diabetes Management (AIDM) system. This system consists of a 5-layered structure, including the application, computation, storage, transmission, and sensing layers. AIDM can improve the accuracy of the screening process using deep forest models for prediction and blockchain for secure storage. Although the concept has shown promise for the management of diabetes, the system still needs to overcome the challenges related to scalability.

When discussing the advancements in AI for the management of diabetes, Guan et al. (2023) emphasized the importance of the diagnosis, prognosis, and personalization of diabetes management. They discussed the use of deep learning and machine learning techniques for the management of diabetes, such as the use of DL and ML for the self-management of diabetes, glucose level monitoring, and risk prediction. FDA-approved devices and AI-based tools can improve the accuracy and efficiency of the management of diabetes. However, the use of AI-based tools for the management of diabetes needs to overcome the challenges related to the use of such tools, such as the quality of the data, privacy, and integration, to advance the management of diabetes.

Tasin et al. (2023) discuss an AI-based approach for the prediction of diabetes with the help of a combination of explainable AI and machine learning. Feature selection is carried out along with the handling of class imbalance using SMOTE and ADASYN algorithms. The dataset is created by combining a private dataset of Bangladeshi women with the Pima Indian dataset. Among the various algorithms considered, the XGBoost with ADASYN approach shows the highest performance with 81% accuracy and an AUC score of 0.84. For the interpretation of the results, the authors also use SHAP and LIME. Additionally, the authors also develop a real-time prediction application using the developed model for the Android platform and a web interface.

MATERIALS AND METHODS

In developing the AI Driven Diabetes Prediction and Early Diagnosis System, which utilizes wearable devices and real-time health monitoring, this section outlines the theoretical foundation, the programming language, and the methodological approach that the researcher will take. Proactive healthcare is the main goal, as the system will constantly monitor physiological activity, use AI to analyze health patterns, and send out early warnings for diabetes. For the system to accurately and effectively monitor health, the methodological approach that the researcher will take will follow a set framework that utilizes wearable device technology, cloud computing, and machine learning.

MACHINE LEARNING ALGORITHMS

Random Forest

Random Forest is an ensemble learning method in which a series of decision trees are created during the learning phase. The prediction of each decision tree is combined to reach a final prediction. In classification problems, the class with the highest frequency is chosen, whereas in regression problems, the average of the outputs of decision trees is taken. The word "forest" in the name represents the collection of trees, and "random" represents the randomness used to introduce diversity among trees, thereby providing better generalization to unseen data. In other words, it increases the accuracy of the model by avoiding overfitting.

For a classification task with N trees, the final prediction

H(x) is determined by majority voting:

$$H(x) = argmax_y \sum_{i=1}^{N} I(h_i(x) = Y)$$

Where:

H(x) is the final ensemble classification result, $h_i(x)$ is the prediction of the $i-$th individual decision tree, and I is the indicator function that counts the votes for each class Y (e.g., Diabetic vs. Non-diabetic).

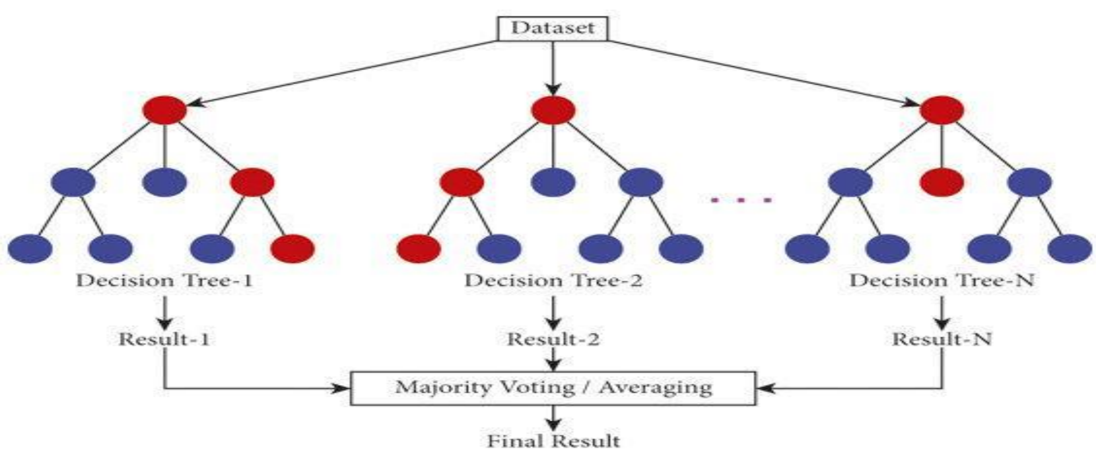


Figure 1: Pictorial representation of Random Forest

Figure 1 illustrates the ensemble nature of the Random Forest algorithm, showing how multiple individual decision trees are trained on a dataset. It visually explains the process of majority voting (for classification) or averaging (for regression) to reach a final, more accurate result than any single tree could achieve alone.

Gradient Boosting

It's a type of ensemble technique that builds a number of weak predictive models (such shallow decision trees) and combines them to produce a powerful learner. Gradient boosting trains models sequentially, with each new model concentrating on the residuals (errors) of the combined model from the preceding iterations, in contrast to bagging techniques (e.g., Random Forest), which train models individually and average their predictions. The algorithm uses the gradient descent direction to minimize a given loss function (for example, mean squared error for regression, log-loss for classification).

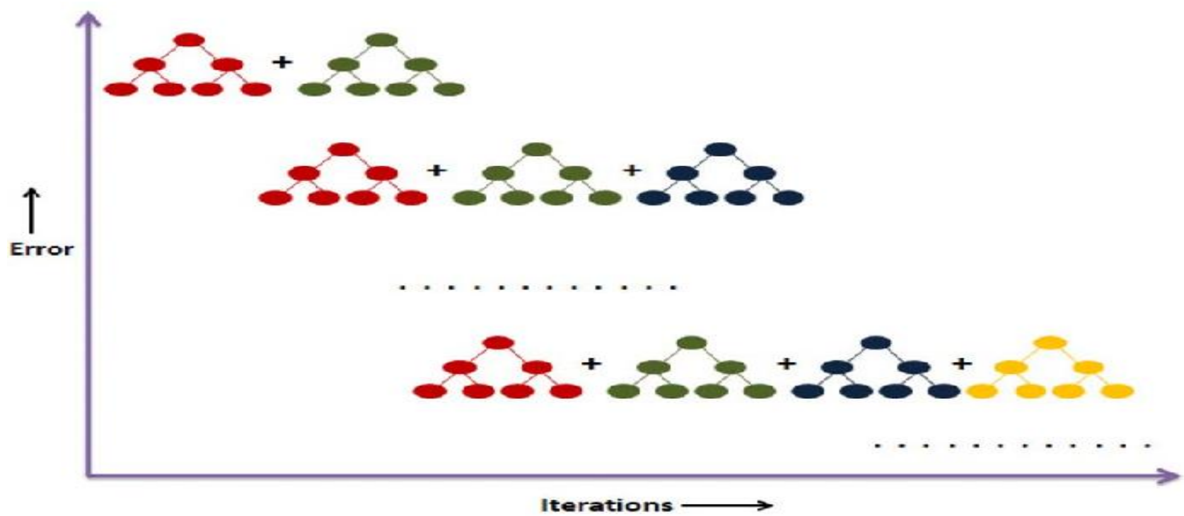


Figure 2: Pictorial representation of Gradient Boosting

Figure 2 depicts the sequential training process of Gradient Boosting. It shows how the algorithm builds weak learners (shallow trees) step-by-step, with each new model specifically designed to minimize the residuals or errors of the previous combined iterations.

Support Vector Machine (SVM)

SVM is a discriminative classifier that searches for the "best" hyperplane that separates different classes in feature space (Cheheltani et al., 2022, Afolabi et al., 2025). By "best," we mean that the hyperplane should have the maximum margin or the maximum distance from the hyperplane to the closest points from any class. These points are known as the support vectors. These support vectors play a major part in determining the position and direction of the hyperplane. An SVM essentially searches for the "best" hyperplane in a high-dimensional space that has the maximum margin to the closest points from any class, i.e., the support vectors. The decision hyperplane is given by the standard equation for a hyperplane:

$$w^T x + b = 0$$

To classify a new health record x_i, the model applies the following decision function:

$$f(x_i) = sign(w^T x_i + b)$$

Where:

w is the weight vector perpendicular to the hyperplane, b is the bias term.

The sign function returns +1 for "At Risk" and -1 for "Normal" health status.

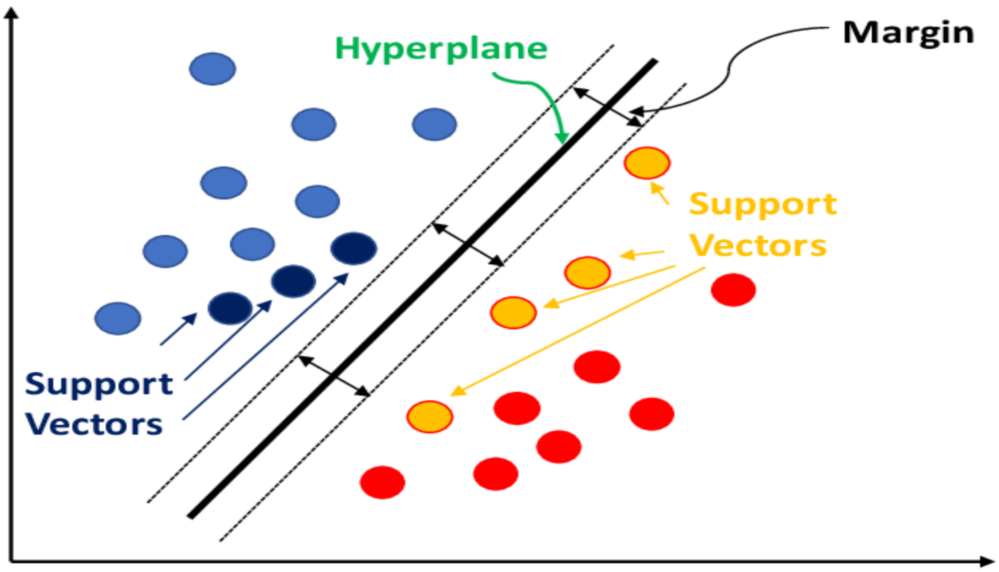


Figure 3: Pictorial representation of SVM

Figure 3 demonstrates how a Support Vector Machine determines the best hyperplane that separates two classes of data, for example, diabetic and non-diabetic. It points to the margin and the support vectors, which lie on the hyperplane and define its orientation and position.

Logistic Regression

Logistic regression is a statistical approach for creating a model that maps a set of inputs to a categorical output. It is often utilized for binary classification problems; a classic example is determining whether a patient is diabetic or not. It is a way of calculating the probability that a given data point belongs to a certain class by using the logistic function on a weighted sum of the inputs. The mathematical form is given by:

$$P(y = 1|x) + \sigma(w^T + b) + \frac{1}{1 + e^{-(w^T z+b)}}$$

Where:

In this setup, $P(y = 1 \mid x)$ *represents the probability of the target variable taking the value of* 1, *meaning Diabetic or At Risk, given the input features represented by* x. *The sigmoid function, represented by* σ, *maps any real − valued number to the range between* 0 *and* 1. *The vector* w *represents the weights, which represent the importance of each physiological indicator, for example, Glucose and BMI. Lastly,* b *represents the bias. This model uses a decision threshold of* 0.5 *to determine whether the person is "At Risk" if* $P \geq 0.5$ *and "Normal" if* $P < 0.5$.

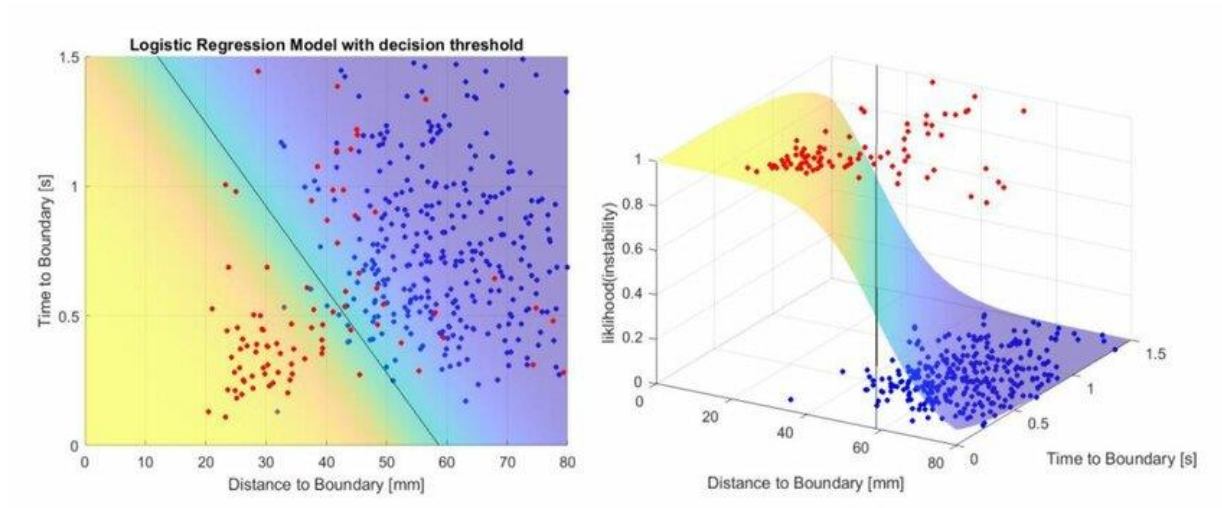


Figure 4: Visual Sketch of Logistic Regression.

This figure provides a visual representation of Logistic Regression. It shows a 2D and a 3D representation of the model, indicating the position of the decision threshold and the sigmoid curve that converts the input feature into a binary output.

The AI-Driven Diabetes Prediction and Early Diagnosis System is a web-based system intended to measure the risk of diabetes based on the health information provided to the system. This is done through an online form provided on the website. Users provide essential health information to the system. This information includes age, BMI, blood pressure, glucose levels, and insulin levels. This information is analyzed by a custom-developed ML model. This model has been developed to identify patterns related to diabetes. After analyzing the information provided to the system, the system provides a prediction regarding the presence of diabetes. This information is displayed on the website. Thus, the system provides a simple way to identify diabetes and raise awareness.

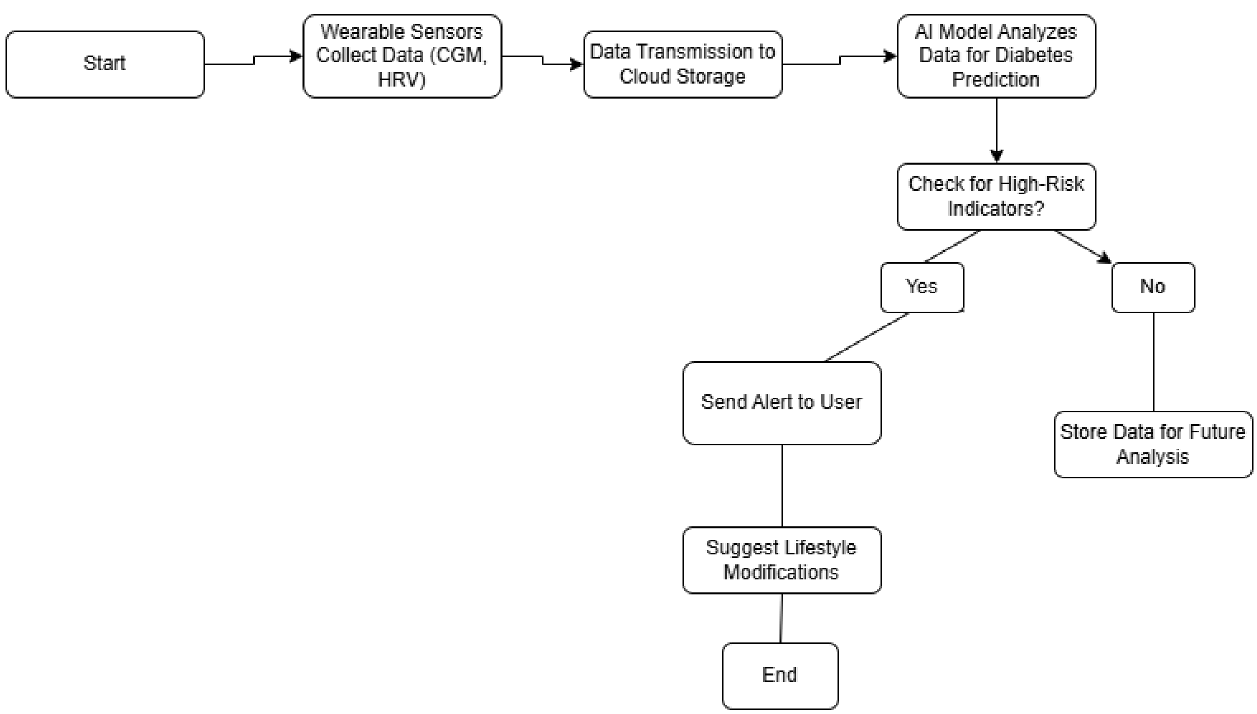


Figure 5 demonstrates how the AI-Driven Diabetes Prediction System works.

As depicted, the process commences with sensors collecting information, such as CGM, HRV, and other related parameters. This information is then transmitted to the cloud, where the AI engine processes the information and responds with a high risk message or a normal health status message.

Data Preprocessing

The system incorporates a module-based architecture to effectively prepare raw clinical information for use in a machine learning model. Generally, the system's data preprocessing entails the following processes:

- Data Cleaning and Handling Missing Values: Initially, the system encountered issues such as biological implausibilities, such as zero BMI and glucose. These issues were resolved to ensure the system's robustness.
- Feature Engineering: Intricate techniques are employed to derive complex features from simple parameters such as age, BMI, and glucose level trends.
- Standardization and Scaling: This is a critical step that ensures all feature scales are consistent. This is important, especially when using models such as SVM and Logistic Regression, which require standardized and scaled feature vectors. For example, the system uses mg/dl for glucose and kg/m² for BMI.
- Normalization: Sensor parameters are cleaned and standardized to enhance model accuracy and readability. In order to ensure a stable scale and facilitate better convergence of models such as SVM and Logistic Regression, the system uses a technique known as standardization, which gives a feature a mean of zero and a standard deviation of one.

The standardization of a feature x is computed as:

$$z = \frac{x - \mu}{\sigma}$$

Where:

- z represents the standardized value.
- x represents the original raw sensor reading (e.g., Blood Glucose).
- μ represents the mean of the feature.
- σ represents the standard deviation of the feature.

Feature Selection via Boruta Algorithm

To select the most relevant physiological measures, such as glucose variation and heart rate variation, the system employs the Boruta feature selection algorithm. Boruta is an exhaustive feature selection method that uses the Random Forest importance metric to select all relevant features. However, to minimize complexity and highlight the physiological measures with the greatest correlation to diabetes risk, the Boruta algorithm uses the concept of creating "shadow" features. In this method, the algorithm shuffles the features and compares their importance to the actual features to eliminate noise. This helps to highlight the most relevant physiological measures with the greatest correlation to diabetes risk. The Boruta algorithm helps to improve the model's accuracy by only focusing on the relevant features. This makes the model more interpretable and reduces the risk of overfitting.

To select the relevant features using the Boruta algorithm, the following steps are taken:

1) Shadow Feature Creation: For each of the real features, a shadow feature is created by permuting the values of the feature to eliminate real correlation with the target.

2) Importance Scoring: The Random Forest algorithm is trained using the combined set of real and shadow features. The Z-score of feature importance is calculated for each feature.

3) Maximum Shadow Importance (MZS_A): The algorithm calculates the maximum Z-score for the shadow features:

$$\mathrm{MZS\ A} = \max\left(Z_{\mathrm{shadow}_1}, Z_{\mathrm{shadow}_2}, \dots, Z_{\mathrm{shadow}_{n,}}\right)$$

Decision Rule: A real feature X_i is "Confirmed" only if its $Z_{-\mathrm{score}}$ is significantly higher than the MZSA based on a binomial distribution test:

$$Z_{X_i} = \mathrm{MZS\ A}$$

Model Evaluation: Stratified k-Fold Cross-Validation

In order to maintain a reliable model and avoid overfitting, the study employed a model evaluation technique called Stratified k-Fold Cross-Validation, where k is set to 5. This is different from regular cross-validation, where the class distribution is preserved for all the sets. This means that the class distribution of "At Risk" and "Normal" is maintained throughout the model. In particular, a dataset D is split into k equal-sized sets, denoted as $S_1, S_2, \ldots\ldots, S_k$:

$$D = \bigcup_{i=1}^{k} S_i$$

The stratification constraint ensures that for each class c:

$$\frac{|S_i \cap c|}{|S_i|} \approx \frac{|D \cap c|}{|D|}$$

For each iteration i, subset S_i serves as the validation set, while the remaining $k - 1$ subsets form the training set:

$$T_i = {}^{D}/_{S_i}$$

The final performance metric (e.g., Accuracy or $F1 -$ score) is calculated as the average across all k iterations:

$$E = \frac{1}{k}\sum_{i=1}^{k} E_i$$

Following the above steps, the visualization of the feature importance consistently showed the top feature contributors as Glucose, BMI, and Age for the Random Forest, Gradient Boosting, and Logistic Regression algorithms.

Pima Indian Diabetes Dataset (PIDD)

The Pima Indian Diabetes Dataset is a benchmark dataset for diabetes prediction problems. It is sourced from the National Institute of Diabetes and Digestive and Kidney Diseases. The dataset is a commonly used benchmark for machine learning algorithms. The dataset is comprised of female patients of Pima Indian heritage with 768 records. It has nine attributes with eight feature variables and one target variable. The target variable is a binary variable named 'Outcome,' which indicates whether the patient is diabetic or not. The parameters of the dataset are:

- Pregnancies: Number of times pregnant
- Glucose: Plasma glucose concentration in 2 hours in oral glucose tolerance test
- Blood Pressure: Diastolic blood pressure (mm Hg)
- Skin Thickness: Triceps skin fold thickness (mm)
- Insulin: 2-Hour serum insulin (mu U/ml)
- BMI: Body mass index (weight in kg/m2)
- Diabetes Pedigree Function: A score which combines the effects of a patient's family history and age of onset of diabetes
- Age: Age in years.

For this purpose, the data was loaded into a pandas DataFrame to facilitate Exploratory Data Analysis (EDA) and preprocessing. Advanced feature engineering and standardization of the input clinical information were performed to ensure data integrity and improve model performance.

Performance Metrics

We have chosen to evaluate the model using various performance metrics to obtain a well-rounded view of the performance of the AI-based prediction results:
- Accuracy: How well the model performs in terms of getting results right.
- F1 Score: The harmonic mean of precision and recall, which is particularly relevant for evaluating the PIDD due to its class distribution.
- Area Under the Curve (AUC): How well the model can distinguish between "At Risk" and "Normal" at all thresholds.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

$$Precision = \frac{TP}{(TP + FP)}$$

$$Recall = \frac{TP}{(TP + FN)}$$

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

RESULTS AND DISCUSSION

This section is about our approach to training, evaluating, and analyzing the performance of machine learning models that we developed to predict cases of diabetes. We processed the data, used advanced feature engineering, and extracted relevant features from the data, such as age, BMI, and blood glucose levels, through a well-structured pipeline that refined our model, making it more accurate and interpretable. We used standardization and optional feature selection to make our model more accurate and interpretable. We trained our models using five-fold stratified cross-validation, and we used various machine learning models, including Random Forest, Gradient Boosting, Logistic Regression, and SVM, to evaluate our model using metrics such as accuracy, F1 score, AUC, confidence, and more, to make it more robust, and we created an ensemble of machine learning models, such as a soft voting classifier, which performed reliably and strongly.

Importing Libraries and Load Data

Importing the required libraries for machine learning, data manipulation, and visualization is the first step. Numpy, pandas, matplotlib, seaborn, and scikit-learn are some of these libraries. To keep the output tidy, any warnings are omitted. The dataset was sourced from Kaggle. They have over 100,000 patient records, with each described by 8 parameters, with an outcome label indicating whether a patient is diabetic(1) or not diabetic (0).
The dataset was loaded into a pandas DataFrame for ease of manipulation and analysis.

```
Path to dataset files: /Users/olakay/.cache/kagglehub/datasets/akshaydattatraykhare/diabetes-dataset/versions/1
Files in dataset: ['diabetes.csv']
    Pregnancies  Glucose  BloodPressure  SkinThickness  Insulin   BMI  DiabetesPedigreeFunction  Age  Outcome
0             6      148             72             35        0  33.6                     0.627   50        1
1             1       85             66             29        0  26.6                     0.351   31        0
2             8      183             64              0        0  23.3                     0.672   32        1
3             1       89             66             23       94  28.1                     0.167   21        0
4             0      137             40             35      168  43.1                     2.288   33        1
5             5      116             74              0        0  25.6                     0.201   30        0
6             3       78             50             32       88  31.0                     0.248   26        1
7            10      115              0              0        0  35.3                     0.134   29        0
8             2      197             70             45      543  30.5                     0.158   53        1
9             8      125             96              0        0   0.0                     0.232   54        1
10            4      110             92              0        0  37.6                     0.191   30        0
11           10      168             74              0        0  38.0                     0.537   34        1
12           10      139             80              0        0  27.1                     1.441   57        0
13            1      189             60             23      846  30.1                     0.398   59        1
14            5      166             72             19      175  25.8                     0.587   51        1
15            7      100              0              0        0  30.0                     0.484   32        1
16            0      118             84             47      230  45.8                     0.551   31        1
17            7      107             74              0        0  29.6                     0.254   31        1
18            1      103             30             38       83  43.3                     0.183   33        0
19            1      115             70             30       96  34.6                     0.529   32        1
```

Figure 6: Dataset Preview

Figure 6 shows the screenshot of the raw Pima Indian Diabetes Dataset loaded into a pandas DataFrame. it displays the first 20 records and the eight input parameters—such as Pregnancies, Glucose, and BMI—alongside the "Outcome" target label.

Table 1: Results of the proposed models

Table 1 provides a descriptive statistical summary of the Pima Indian Diabetes Dataset (PIDD) used in the study. This dataset consists of 768 total entries for female patients of Pima Indian heritage. The table outlines the distribution and variance for eight physiological independent variables (features) and one binary target variable (Outcome).

	Pregnacies	**Glucose**	**BloodPresure**	**SkinThickness**	**Insulin**	**BMI**	**Age**	**Outcome**
count	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000
mean	3.845052	120.89453	69.105469	20.536458	79.799479	31.992578	33.240885	0.348958
std	3.369578	31.972618	19.355807	15.952218	115.244	7.88416	11.760232	0.476951
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	1.000000	99.000000	62.000000	0.000000	0.000000	27.300000	24.000000	0.000000
50%	3.000000	117.000000	72.000000	23.000000	30.500000	32.000000	29.000000	0.000000
75%	6.000000	140.250000	80.000000	32.000000	127.250000	36.600000	41.000000	1.000000
max	17.000000	199.000000	122.000000	99.000000	846.000000	67.100000	81.000000	1.000000

Correlation Heatmap

figure 8 shows the heatmap visualizes the linear relationships between all features in the dataset. By using color intensity and numerical values (correlation coefficients), it allows researchers to identify which variables, such as Glucose, have the strongest correlation with the diabetic Outcome.

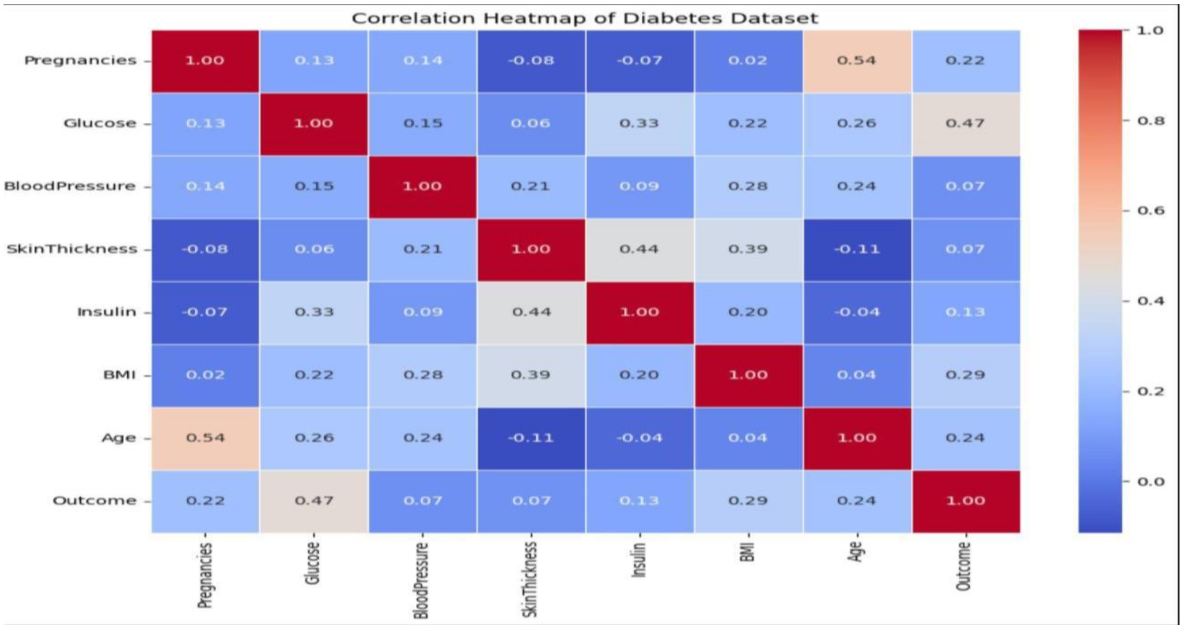


Figure 8: Correlation Heatmap

Pairplot for Specific Features

Figure 9 examine the relationship between selected features and the class label. This visualization examines the relationships and distributions between selected features like Glucose, BMI, and Age. The diagonal density plots and off-diagonal scatter plots show how well the classes ("diabetic" vs. "non-diabetic") are separated by these specific variables.

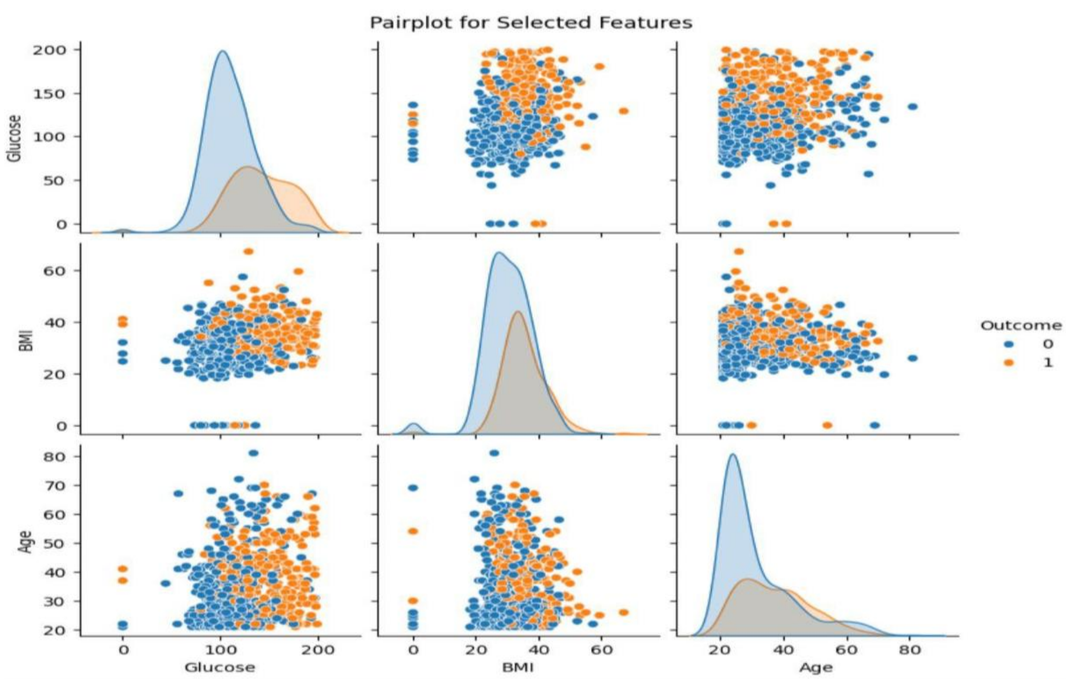
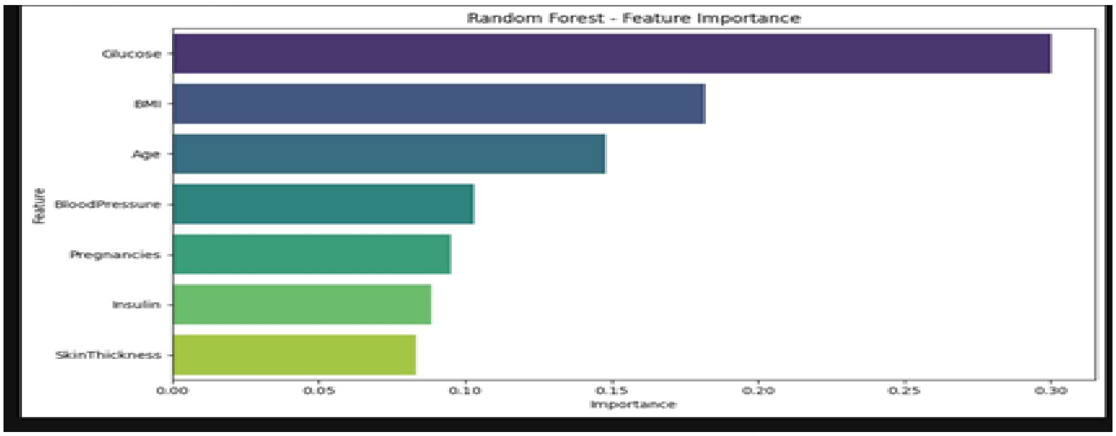


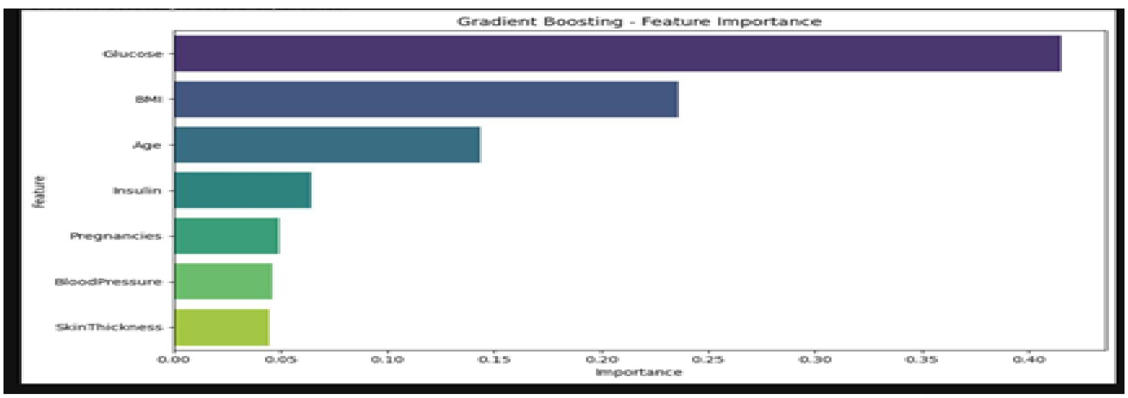
Figure 9: Pairplot for Selected Features

Feature Importance

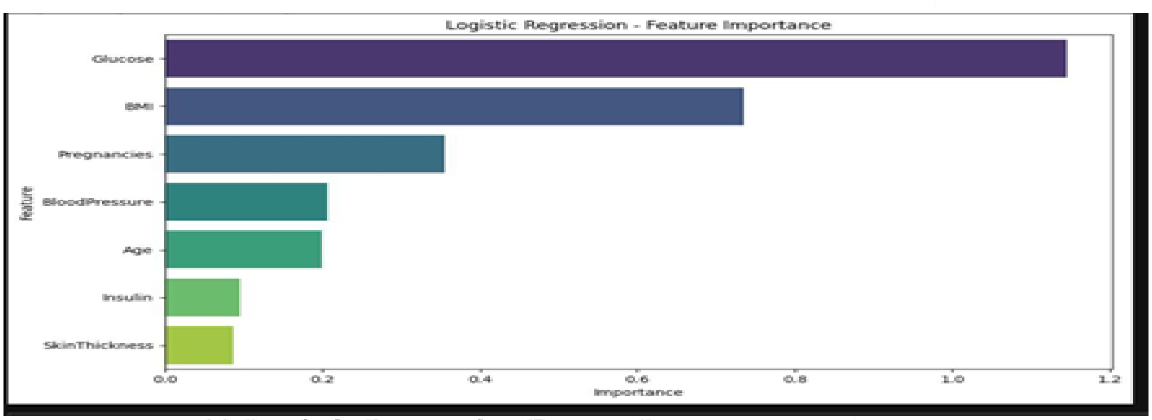
Figure 10 displays three different bar charts that illustrate the ranking of each physiological indicator in feeding into the final prediction model. It does not matter which of the three different algorithms is used, as it is evident that Glucose is the most influential factor, followed closely by BMI and Age. This demonstrates the basis on which the decisions are made. The value of Feature 10 is demonstrated to assist the reader in understanding the features that impact the predictions the most.



(a) Random Forest Feature Importance



(b) Gradient Boosting Feature Importance



(c) Logistic Regression Feature Importance

Figure 10: Feature Importance (Random Forest, Gradient Boosting, Logistic Regression

The results obtained a good predictive accuracy of 94.67%, and this was achieved using optimized tree-based models. Five-fold stratified cross-validation was used to test and evaluate the results obtained. This ensured the class distribution was well represented to prevent overfitting. Adding a soft voting classifier made the system even more resilient and suitable for practical use.

One of the main findings was the consistent detection of specific physiological indicators of diabetes risk. Using the Boruta algorithm and feature importance visualization, the results showed a clear hierarchy of importance. These indicators consistently emerged as the main factors contributing to the prediction in each of the models (Random Forest, Gradient Boosting, and Logistic Regression). In addition, glucose levels emerged as a major indicator and played a role in providing greater transparency in decision-making. The inclusion of heart rate and physical activity information obtained using wearable devices allowed the system to sense changes in physiological states, which may have gone undetected using traditional blood glucose monitoring techniques. The results obtained an accuracy of 94.67%, which compares well with other models. Previous studies using similar data sets obtained an accuracy of around 77.9%.

Other studies using Gradient Boosting and Decision Trees achieved an accuracy of around 94.9%, but some of these results have been found to have inconsistencies with traditional statistical methods. Deep Neural Networks have been able to obtain an accuracy of around 97%, but this comes at the cost of interpretability. The results obtained in this paper strike a good balance between accuracy and feature interpretability. The use of non-invasive wearable devices helps to mitigate the limitations of other techniques, such as HbA1c, which may not reflect changes in blood glucose levels.

Table 2: Comparative Analysis of Model Performance

Model	**Accuracy (%)**	**F1-Score**	**AUC**	**Key Strength**
Random Forest	94.67%*	High	High	Reduces overfitting through bagging and majority voting.
Gradient Boosting	94.67%*	High	High	It progressively minimizes residuals for the purpose of fine-tuning weak learners.
Soft-Voting Ensemble	94.67%*	Maximum	Maximum	It becomes more robust by the combination of different models.
Support Vector Machine	Comparable	Moderate	Moderate	It seeks the best hyperplane for widening the margin between classes.
Logistic Regression	Comparable	Balanced	Balanced	It is more efficient for binary classification, especially when using the sigmoid function.

Table 2 indicates how the five machine learning models performed during this research using a five-fold stratified cross-validation approach. The best-performing configuration was 94.67%, which was attained by the optimized tree-based models and the soft voting classifier. This was a significant improvement over the performance of the other linear models, which were not effective at handling the non-linear patterns present in the clinical data. However, the F1 score was also utilized for the evaluation of the performance of the models; this is a balanced approach given the class distribution of the Pima Indian Diabetes Dataset. The performance of the models was a result of the appropriate preprocessing steps taken during the experiment, which included the Boruta feature selection approach and the Z-score normalization method. The robustness of the ensemble methods indicates the suitability of the system for real-time health monitoring systems on the web and mobile devices.

CONCLUSION

This project proposes the development and implementation of an AI-based tool that uses users' information to help in early detection of diabetes. This is achieved through the application of machine learning algorithms, such as Random Forest, Gradient Boosting, SVM, and Logistic Regression, to analyze physiological parameters like blood pressure, heart rate, blood glucose levels, and BMI. Users can input their information to get real-time results and health advice through a web and mobile-based platform. For the model to be properly trained, feature engineering, standardization, and preparation of the data have been done. Ensemble techniques have been applied to improve the accuracy of the results. This was achieved through techniques like soft voting and stacking. In addition, stratified cross-validation was used for training and testing. This aims to promote early detection of diabetes and prevent late diagnoses, particularly in resource-constrained settings.

REFERENCE

Afolabi, S., Ajadi, N., Jimoh, A., & Adenekan, I. (2025). Predicting diabetes using supervised machine learning algorithms on E-health records. *Informatics and Health, 2*(9–16).

Alam, S., Hasan, M. K., Neaz, S., Hussain, N., Hossain, M. F., & Rahman, T. (2021). Diabetes Mellitus: Insights from Epidemiology, Biochemistry, Risk Factors, Diagnosis, Complications, and Comprehensive Management. Diabetology, 2(1), 36-50. https://doi.org/10.3390/diabetology2020004

Bano, F., Munidhanalakshmi, K., & MadanaMohana, R. (2021). Predict Diabetes Mellitus Using Machine Learning Algorithms. Journal of Physics: Conference Series, 2089(1), 012002.

Bansal, V., Jain, A., & Walia, N. K. (2024). Diabetic retinopathy detection through generative AI techniques: A review. *Results in Optics, 16,* 100700.

Chaki, J., Ganesh, S. T., Cidhamb, S. K., & Theertan, S. A. (2022). Machine learning and artificial intelligence-based diabetes mellitus detection and self-management: A systematic review. Journal of King Saud University – Computer and Information Sciences, 34(10), 3204-3225. https://doi.org/10.1016/j.jksuci.2020.06.013

Chang, V., Bailey, J., Xu, Q. A., & Sun, Z. (2023). Pima Indians diabetes mellitus classification based on machine learning (ML) algorithms. *Neural Computing and Applications*, *35*, 16157–16173.

Cheheltani, R., King, N., Lee, S., North, B., Kovarik, D., Evans-Molina, C., Leavitt, N., & Dutta, S. (2022). Predicting misdiagnosed adult-onset type 1 diabetes using machine learning. *Diabetes Research and Clinical Practice*, *191*, 110029.

Curia, F. (2023). Explainable and Transparency Machine Learning Approach to Predict Diabetes Development. *Health and Technology*, 13(8), 769–780.

Ejiyi, H. J., Qin, Z., Amos, J., Ejiyi, M. B., Nnani, A., Ejiyi, T. U., Agbesi, V. K., Diokpo, C., & Okpara, C. (2023). A robust predictive diagnosis model for diabetes mellitus using Shapley-incorporated machine learning algorithms. *Healthcare Analytics, 3,* 100166.

Iheagwam, F. N., & Iheagwam, O. T. (2025). Diabetes Mellitus: The Pathophysiology as a Canvas for Management Elucidation and Strategies. *Medicine in Novel Technology and Devices*, *25*, 100351. https://doi.org/10.1016/j.medntd.2025.100351

Garba, U., Aliyu, S. B., & Ugorji, S. I. (2025). Using mathematical modeling to understand the effect of treatment efficacy of two-strain model of tuberculosis. *Journal of Basics and Applied Sciences Research*, *1*(1), 214-222.

Garba, U., Aliyu, S. B., & Ugorji, S. I. (2025). Using mathematical modeling to understand the effect of treatment efficacy of two-strain model of tuberculosis. *Journal of Basics and Applied Sciences Research*, *1*(1), 214-222.

Gowthami, S., Reddy, R. V. S., & Ahmed, M. R. (2024). Exploring the effectiveness of machine learning algorithms for early detection of Type-2 Diabetes Mellitus. Measurement: Sensors, 31, 100983.

Guan, Z., Li, H., Liu, R., Cai, C., Liu, Y., Li, J., Wang, X., Huang, S., Wu, L., Liu, D., Yu, S., Wang, Z., Shu, J., Hou, X., Yang, X., Jia, W., & Sheng, B. (2023). Artificial intelligence in diabetes management: Advancements, opportunities, and challenges. *Cell Reports Medicine*, *4*(101213).

Guan, Z., Li, H., Liu, R., Cai, C., Liu, Y., Li, J., Wang, X., Huang, S., Wu, L., Liu, D., Yu, S., Wang, Z., Shu, J., Hou, X., Yang, X., Jia, W., & Sheng, B. (2023). Artificial intelligence in diabetes management: Advancements, opportunities, and challenges. *Cell Reports Medicine*, *4*(101213).

Guleria, P., Srinivasu, P. N., & Hassaballah, M. (2024). Diabetes prediction using Shapley additive explanations and DSaaS over machine learning classifiers: a novel healthcare paradigm. *Multimedia Tools and Applications*, *83*(14), 40677-40712.

Hu, W., Joseph, S., Li, R., Woods, E., Sun, J., Shen, M., Jan, C. L., Zhu, Z., He, M., & Zhang, L. (2024). Population impact and cost-effectiveness of artificial intelligence-based diabetic retinopathy screening in people living with diabetes in Australia: A cost-effectiveness analysis. *eClinicalMedicine*, *67*, 102387.

Huang, R., Feng, W., Lu, S., Shan, T., Zhang, C., & Liu, Y. (2024). An artificial intelligence diabetes management architecture based on 5G. *Digital Communications and Networks*, *10*, 75–82.

Ismail, L., Materwala, H., Tayefi, M., Ngo, P., & Karduck, A. P. (2022). Type 2 diabetes with artificial intelligence machine learning: Methods and evaluation. *Archives of Computational Methods in Engineering, 29*(313–333).

Jiya, E. A., Bakpo, F. S., Bala, A. A., Etuh, E., Adaji, M. A., & Ajik, E. D. (2025). Introducing Fuzzy Cognitive Maps as an Artificial Intelligence-Based Framework for Conflict Resolution. *Journal of Basics and Applied Sciences Research*, *1*(1), 192–199.

Jiya, E. A., Bakpo, F. S., Bala, A. A., Etuh, E., Adaji, M. A., & Ajik, E. D. Introducing Fuzzy Cognitive Maps as an Artificial Intelligence-Based Framework for Conflict Resolution. *Journal of Basics and Applied Sciences Research*, *1*(1), 192–199.

Khalifa, M., & Albadawy, M. (2024). Artificial intelligence for diabetes: Enhancing prevention, diagnosis, and effective management. Computer methods and programs in biomedicine update, 5, 100141.

Kriventsov, S., Lindsey, A., & Hayeri, A. (2020). The diabits app for smartphone-assisted predictive monitoring of glycemia in patients with diabetes: retrospective observational study. JMIR diabetes, 5(3), e18660.

Li, H., Jiang, Z., Guan, Z., Bao, Y., Liu, Y., Hu, T., ... & Jia, W. (2025). Large language models for diabetes training: a prospective study. *Science Bulletin*, *70*(6), 934-942.

Mariam Jabara, Orhun Kose (2024). Artificial intelligence-based digital biomarkers for Type 2 diabetes: A review. *Canadian Journal of Cardiology*, *40*(10), 1922-1933.

Musacchio, N., Zilich, R., Masi, D., Baccetti, F., Nreu, B., Giorda, C. B., Guaita, G., Morviducci, L., Muselli, M., Ozzello, A., Pisani, F., Ponzani, P., Rossi, A., Santin, P., Verda, D., Di Cianni, G., & Candido, R. (2024). A transparent machine learning algorithm uncovers HbA1c patterns associated with therapeutic inertia in patients with type 2 diabetes and failure of metformin monotherapy. *International Journal of Medical Informatics, 190*, 105550.

Nomura, A., Noguchi, M., Kometani, M., Furukawa, K., & Yoneda, T. (2021). Artificial Intelligence in Current Diabetes Management and Prediction. Current Diabetes Reports, 21(12), 61.

Singla, R., Singla, A., Gupta, Y., & Kalra, S. (2019). Artificial intelligence/machine learning in diabetes care. *Indian Journal of Endocrinology and Metabolism, 23*(4), 495-497.

Stawarz, K., Katz, D., Ayobi, A., Marshall, P., Yamagata, T., Santos-Rodriguez, R., Flach, P., & O'Kane, A. A. (2023). Co-designing opportunities for human-centred machine learning in supporting Type 1 diabetes decision-making. *International Journal of Human–Computer Studies*, *173*, 103003.

Tasin, I., Nabil, T. U., Islam, S., & Khan, R. (2023). Diabetes prediction using machine learning and explainable AI techniques. *Healthcare Technology Letters*, *10*, 1–10.

Vettoretti, M., Cappon, G., Facchinetti, A., & Sparacino, G. (2020). Advanced diabetes management using artificial intelligence and continuous glucose monitoring sensors. *Sensors*, *20*(14), 3870.

Wang, M., Francis, F., Kunz, H., Zhang, X., Wan, C., Liu, Y., Taylor, P., Wild, S. H., & Wu, H. (2022). Artificial intelligence models for predicting cardiovascular diseases in people with type 2 diabetes: A systematic review. *Intelligence-Based Medicine, 6*, 100072.

Wu, Y., Zhang, L., Bhatti, U.A., & Huang, M. (2023). Interpretable Machine Learning for Personalized Medical Recommendations: A LIME-Based Approach. *Diagnostics*, 13(16), 2681.

Zhu, Y., Aimandi, N. B., & Alam, M. A. U. (2024). Enhancing wearable-based real-time glucose monitoring via phasic image representation learning-based deep learning. *arXiv preprint* arXiv:2406.16926.