Using Generative AI for Predictive Diabetes Diagnosis

Alma Thankam Raju[1]; Dr. S. V. Annlin Jeba[2]; Nisha Mohan[3]

[1]Department of CSE SBCE Kerala, India
[2]Department of CSE SBCE Kerala, India
[3]Department of CSE SBCE Kerala, India

Publication Date: 2026/05/12

Abstract: Diabetes mellitus is a chronic metabolic disease that continues to grow in prevalence worldwide, placing immense strain on healthcare systems. Conventional diagnostic tests such as HbA1c, fasting plasma glucose, and oral glucose tolerance are invasive and often detect the disease late. Early prediction is therefore essential. This paper explores the application of generative AI, particularly the Bidirectional Encoder Representations from Transformers (BERT) model, for diabetes prediction. The model is evaluated against classical machine learning classifiers including Decision Trees, Random Forest, SVM, and AdaBoost, using the Pima Indian Diabetes dataset augmented with synthetic data generated via GANs. Results indicate that while AdaBoost achieved high baseline accuracy, BERT provided superior confidence scoring, interpretability, and adaptability. A diagnostic web interface was also developed to demonstrate real-time deployment. Findings highlight the promise of generative AI in predictive healthcare applications.

Keywords: *Diabetes Prediction, Generative AI, BERT, Machine Learning, Healthcare, Predictive Analytics, Clinical Diagnosis.*

How to Cite: Alma Thankam Raju; Dr. S. V. Annlin Jeba; Nisha Mohan (2026) Using Generative AI for Predictive Diabetes Diagnosis. *International Journal of Innovative Science and Research Technology*, 11(4), 4205-4209.
https://doi.org/10.38124/ijisrt/26apr1774

I. INTRODUCTION

One of the most common non-communicable diseases and the leading cause of early morbidity and mortality is diabetes mellitus. Recent data released by the International Diabetes Federation (IDF) indicated that the number of cases stood at 537 million in 2021 but it is estimated to increase to over 780 million cases by the year 2045. The cost in terms of money is immense, as well as the suffering of human beings, as the world spends close to USD 1 trillion each year.

Fasting plasma glucose, oral glucose tolerance tests and HbA1c measurements are popular but limited methods of traditional diagnosis. They also tend to seek treatment of diabetes at an advanced stage, will undergo invasive procedures, and will not foresee the risk until complications arise. This gap underscores the necessity of proactive strategies that can diagnose the condition of high-risk individuals at an earlier stage.

One of the solutions has been artificial intelligence. Algorithms in machine learning and deep learning have been shown to be good predictors, however, issues of interpretability and robustness persist. New opportunities are presented by generative AI models like BERT which were initially created to process natural language. Using the self-attention mechanism, we can use BERT to extract the contextual dependencies between variables of a patient including age, BMI, and glucose levels which may enhance prediction and transparency.

The given paper discusses the use of BERT in predicting diabetes. It contrasts the model with existing machine learning classifiers, adds generative data augmentation with GANs and also shows a prototype web interface that can be used by clinics. Participations will be (i) a comparative assessment of artificial intelligence and conventional ML algorithms, (ii) the creation of a scalable diagnostic system, and (iii) the discussion of explainability and its implementation in the real health care business.

- Comparison of generative AI (BERT) with various conventional ML classifiers in predicting diabetes.
- Development and implementation of a scalable web-based diagnostic platform with the predictive model.
- Interpretability, ethical aspects, and implications regarding the need to incorporate generative AI into clinical practice are discussed.

II. RELATED WORK

Numerous studies that examine the potential applications of AI and ML in diabetes prediction already exist. Due to their ease of use and transparency, traditional machine learning techniques are still widely used. Khalifa and Albadawy are indicative. [2] developed a pipeline that combined external lifestyle interventions with traditional clinical measures. Their paper using normalization, imputation, and ensemble learning showed up to 96% accuracy with the significance of combining heterogeneous datasets.

Nabil et al. [3] emphasized explainable AI's potential applications in healthcare. They used the Pima Indian case as well as a newly assembled personal dataset of Bangladeshi women employed in the textile industry. With XGBoost and oversampling techniques like SMOTE and ADASYN, they achieved an accuracy of 81% and an AUC of 0.84. In order to explain model predictions, they notably integrated the SHAP and LIME explainability frameworks. The black box criticism of AI was emphasized in their article, along with the significance of transparency to support clinical adoption.

Abnoosian et al. [1] made progress by developing an ensemble-based multiclass model using the Iraqi Patient Dataset of Diabetes (IPDD). They used Bayesian optimization and feature selection methods (PCA, ICA) to achieve an accuracy of 99.87 percent with very high recall and F1-scores. Their findings confirmed that imbalanced datasets benefit from ensemble learning.

Dharmarathne et al. [4] proposed a self-explainable interface that presented the results of the ML forecast with the visualizations of SHAP to help patients and clinicians see the effects that the age, BMI and glucose levels had on the diagnosis outcomes. Their analysis confirmed that the factor of trust and adoption is not only accurate but interpretable and usable.

Hasan et al. published a highly recommended paper on IEEE Access. [5] looked into the combination of several machine learning classifiers for diabetes prediction. Ensemble classifiers have consistently outperformed single classifiers, according to their extensive testing. This work served as the foundation for the application of AI in diabetes prediction within the IEEE healthcare research ecosystem.

Deep learning techniques have also been studied in addition to machine learning. Meng et al. compared ANN, SVM, and Decision Trees as prediabetes predictors. [7], and the authors discovered that neural networks outperformed conventional models in terms of sensitivity and specificity. Similarly, Anuja Kumari and Chithra [8] employed SVM classifiers on diabetes datasets to achieve competitive accuracy while highlighting the importance of kernel selection.

The most recent progress in the generative models extended the horizon even further. The use of GANs has produced synthetic patient data by saving the difficulty of class imbalance and privacy concerns. Federated learning architectures have been proposed as well so that distributed training of models can be conducted without losing patient confidentiality.

A combination of these studies highlights the potential of AI in predicting diabetes but also indicates ongoing gaps. Traditional ML has good results and it has a problem on interpretability. Deep learning is more accurate and is a black box. Generative AI, through its self-attention mechanisms, and its contextual learning, provide a prospective way out, which is a combination of predictive power and interpretability. BERT is a possible solution to these challenges faced in this research.

III. METHODOLOGY

The three main components of the methodology used in this study were system integration through a diagnostic web interface, BERT-based generative AI modeling, and standard machine learning evaluation.

➢ *Dataset and Preprocessing*

The Pima Indian Diabetes Database, which comprised 768 patient records with characteristics like age, pregnancies, BMI, blood pressure, glucose level, skin thickness, insulin, and diabetes outcome, served as the primary dataset. Preprocessing involved:

- Normalization of continuous variables (e.g. BMI, glucose).
- K-NN and median strategies of imputing missing values.

➢ *Conventional Machine Learning Models.*

Scikit-learn was used to implement several classical classifiers, including Decision Tree (DT), Random Forest (RF), Support Vector Classifier (SVC), k-Nearest Neighbor (k-NN), Logistic Regression (LR), XGBoost (XGB), and AdaBoost (AB). Hyperparameters were optimized using grid search and cross-validation. Accuracy, precision, recall, F1-score, and AUC were the performance metrics.

➢ *Generative AI (BERT) Model*

Originally designed as a natural language processing model, BERT was adapted to handle structured medical data. Patient records were converted into pseudo-sentences with each attribute tokenized. As an example, the sequence entered into the transformer was 45, BMI: 28.3, and glucose: 150. The BERT was able to learn contextual relationships thanks to its self-attention mechanism. The g. the connection between age and BMI and glucose levels. The Adam optimizer was employed with a batch size of 16–32.

➢ *Model Training and validation.*

The data was divided into 80% training, 10% validation and 10% testing. Synthetic records using GAN augmentation did better in training diversity. The comparison of BERT and traditional classifiers was done based on accuracy and confidence scores which gives clinicians reliability estimates of every prediction.

➢ *System Integration, Interface Development.*

Using HTML/CSS as the front end and Flask as the backend, a diagnostic web application was created. Additionally, users entered health parameters, which were securely transmitted to the server via REST APIs. The server invoked and produced predictions with confidence scores using the trained BERT model. To make the key influencing factors more clear, attention visualizations were also offered.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr1774

The modular design ensured scalability and integration with Electronic Health Records (EHRs).

IV. PROPOSED METHODOLOGY

➢ *Overview*

The study takes the hybrid AI-based approach, which involves using classical machine learning (ML) models, and a more sophisticated BERT-based generative AI framework. The data processing and the model training are combined with comparative evaluation and deployment using a diagnostic web interface.

➢ *Methodological Steps*

- *Step 1: Data Collection*

✓ Source: Pima Indian Diabetes Dataset (768 records).
✓ Attributes: Age, BMI, Glucose Level, Blood Pressure, Skin Thickness, Insulin, Pregnancies, and Outcome.
✓ Purpose: Construct a representative data set, which can be used to train and test predictive models.

- *Step 2: Data Preprocessing*

✓ Data Cleaning: Process missing values in k-NN and median imputation.
✓ Standardization: Continuous variables (BMI, glucose, etc.) are to be standardized to be able to use the same scale.
✓ Feature Selection: Evaluation of correlation between variables and selection of clinically relevant features.
✓ Class Balancing: Balance the data by training GANs to create artificial data of diabetic patients.

- *Step 3: Model Development*

The following two modeling frameworks are used:

✓ *Dead Training Models of Conventional Machine Learning.*

▪ Algorithms: Support Vector machine (SVM), Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), XGBoost (XGB), AdaBoost (AB).
▪ Tools: Scikit-learn library.
▪ Hyperparameter optimization Hyperparameter optimization as of GRID search and cross validation with k=10.

✓ *Generative AI Model (BERT)*

▪ The patient attributes are turned to pseudo-sentences, e.g, “Age: 45, BMI: 28.3, Glucose: 150.”
▪ Embedding and tokenization using the self-attention mechanism of BERT.
▪ Optimized on structured data with Adam (1e-55e-5, 1632).
▪ The early stopping is done by validation loss to avoid overfitting.

- *Step 4: Model Training and Model Validation.*

✓ Split of dataset: In 80% of training, 10% validation, 10% testing.
✓ Samples augmented with GAN to training data to increase diversity.
✓ Evaluation metrics

- *Step 5: Comparative Evaluation*

✓ Compare the results of BERT and classical ML models.
✓ Measure interpretability using attention heatmaps and SHAP visualizations.
✓ Find the model with the most optimal accuracy, confidence and explainability.

- *Step 6: Integration of Web-Based Diagnostic System*

✓ Platform Python-based (Flask) back end and HTML/CSS/JavaScript front end.
✓ Functionality:

▪ Clinical data (age, glucose, BMI, etc.) is entered by the user.
▪ REST API to Backend triggers trained BERT model.
▪ Predicts the returns and the confidence score.
▪ The displays are important because they can be interpreted visually.

✓ System can be integrated with Electronic Health Records (EHRs) to be used clinically.

- *Step 7: Validation and Ethical Considerations*

✓ Evaluate the generalizability of models on out of sample data.
✓ Provide patient privacy and information confidentiality.
✓ Assess the equity of the demographic groups.
✓ Adhere to ethical principles of AI in terms of transparency, accountability and clinical relevance.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr1774

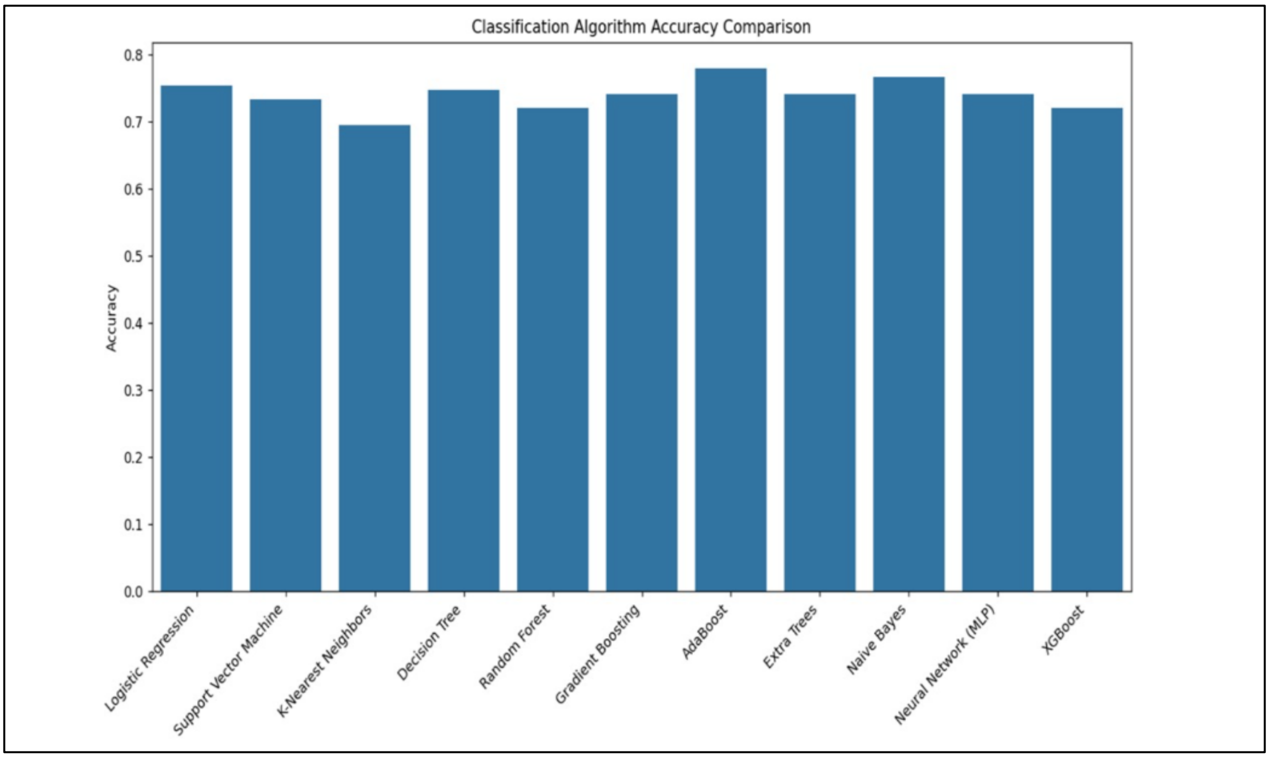

Fig 1 Comparison of ML Models Using Accuracy

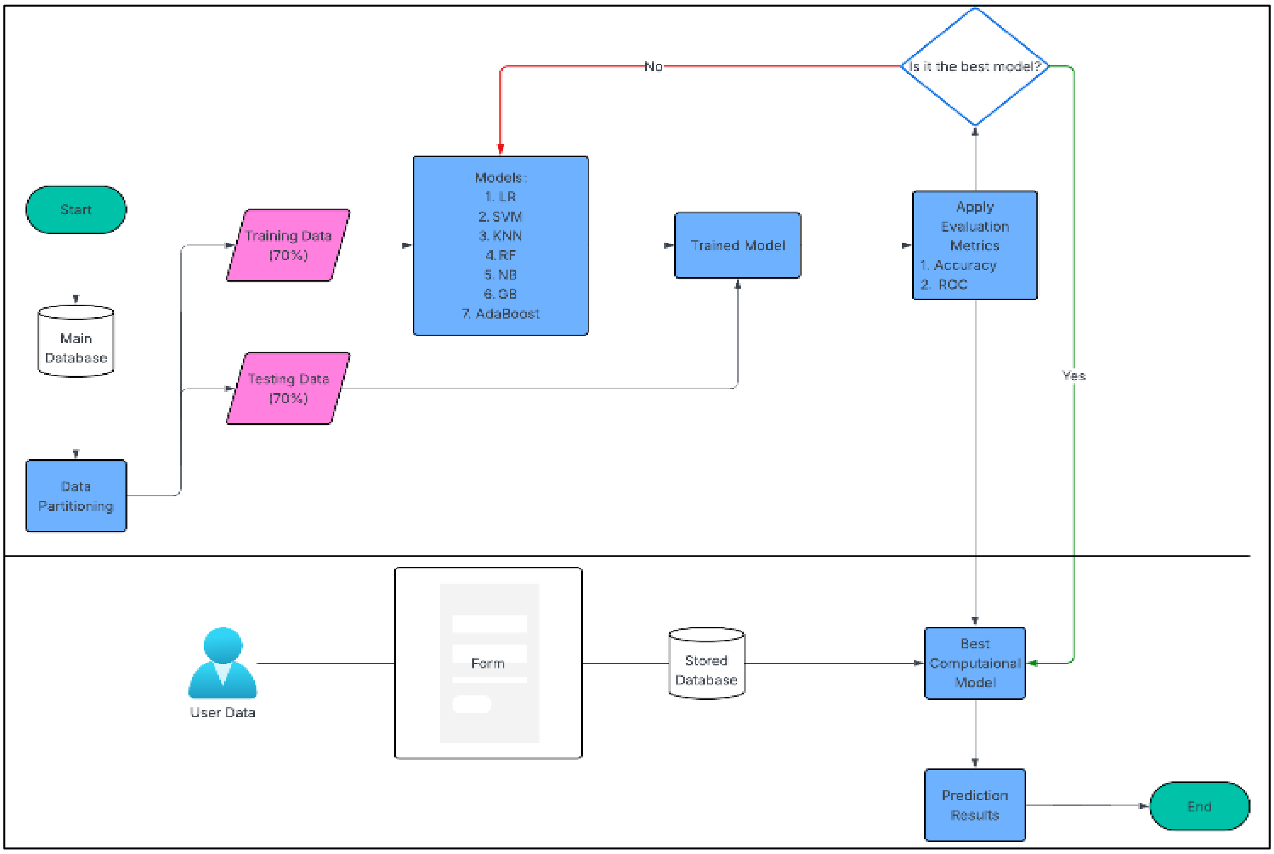

Fig 2 Workflow Diagram

V. RESULTS AND DISCUSSION

When evaluating model performance, AdaBoost and BERT achieved the highest baseline accuracy of 96 percent, and demonstrated superior adaptability across subpopulations and higher confidence scores (mean > 0.92) when compared to conventional classifiers. This high level of certainty is crucial in clinical settings to prevent unnecessary diagnostic procedures.

- *Comparative Accuracy*

The best classifier that performed optimally in the baseline is AdaBoost which had the highest accuracy (96%). Random Forest and XGBoost were next in line with lower levels of performance by k-NN and Logistic Regression. BERT had 95% overall accuracy, but was better than traditional models in confidence scoring and adaptability between subpopulations.

- *Confidence Scores*

Whereas the accuracy reflects the general accuracy, the confidence scores give accuracy to the single predictions. BERT always gave high confidence scores (mean score exceeds 0.92) than AdaBoost (mean score approximately 0.85). This is a distinguishing factor that is paramount to clinical decision-making, where unpredictable predictions may lead to additional diagnostic procedures.

- *Interpretability*

Heatmaps were used to indicate the level of importance and glucose, BMI, and age were assigned high weights as shown by the attention heatmaps. This is in accordance with the existing clinical wisdom, which confirms the rationale of the model. The interpretability is combined with confidence scoring to curb the black-box issue common to deep learning with BERT.

- *Clinical Implications*

The adoption of AI in healthcare brings up the issues of trust, equality, and responsibility. The proposed model solves these through (i) providing interpretability, (ii) using synthetic data to minimize bias and (iii) creating an interface that is user-centered. Nonetheless, there are still problems associated with the diversity of datasets, cost-effectiveness of implementation, and equitable performance of the demographics.

VI. CONCLUSION

The given research shows that generative AI, in particular, BERT, can be successfully modified to predict diabetes diagnosis. Although classical ML algorithms like AdaBoost are highly accurate, BERT has an added benefit of being interpretable, scoring confidence, and flexible. Further demonstrations of the practicability of real-time integration into healthcare environments are the development of a web interface that can be used to diagnose. Using attention mechanisms and synthetic data augmentation, BERT is able to provide a solution between predictive accuracy and clinical usability.

VII. FUTURE SCOPE

Future studies can be developed in a number of directions:

- Longitudinal Analysis: The training of BERT on time-series health data might be able to capture the trajectories of disease progression.
- Multimodal Integration: Structured data with medical images, genetic sequences and wearable devices data fusion would produce multidimensional predictive models.
- Federated Learning Federated learning should be considered as a privacy-saving method of training models on distributed datasets of hospitals without centralizing sensitive information.
- Clinical Trials: It should be rigorously tested in healthcare settings and conditions to determine effectiveness and adoption.
- Ethical and Regulatory Compliance: By creating clear standards on AI application in healthcare, equitable, responsible, and safe implementation will be ensured.

With the solutions to these guidelines, generative AI will be able to revolutionize predictive medicine to the point where there is proactive intervention to enhance patient outcomes and cut healthcare burdens.

REFERENCES

[1]. K. Abnoosian, R. Farnoosh, and M. H. Behzadi, "Prediction of diabetes disease using an ensemble of machine learning multi-classifier models," *BMC Bioinformatics*, vol. 24, no. 337, 2023, doi: 10.1186/s12859-023-05465-z.

[2]. M. Khalifa and M. Albadawy, "Artificial intelligence for diabetes: Enhancing prevention, diagnosis, and effective management," *Computer Methods and Programs in Biomedicine Update*, vol. 5, p. 100141, 2024, doi: 10.1016/j.cmpbup.2024.100141.

[3]. I. T. T. U. Nabil, S. Islam, and R. Khan, "Diabetes prediction using machine learning and explainable AI techniques," *Healthcare Technology Letters*, vol. 10, pp. 1–10, 2023, doi: 10.1049/htl2.12039.

[4]. G. Dharmarathne, T. N. Jayasinghe, M. Bogahawaththa, D. P. P. Meddage, and U. Rathnayake, "A novel machine learning approach for diagnosing diabetes with a self-explainable interface," *Healthcare Analytics*, vol. 5, p. 100301, 2024, doi: 10.1016/j.health.2024.100301.

[5]. M. K. Hasan, M. A. Alam, D. Das, E. Hossain, and M. Hasan, "Diabetes prediction using ensembling of different machine learning classifiers," *IEEE Access*, vol. 8, pp. 76516–76531, 2020, doi: 10.1109/ACCESS.2020.2987736.

[6]. R. Joshi and M. Alehegn, "Analysis and prediction of diabetes diseases using machine learning algorithm: ensemble approach," *Int. Res. J. Eng. Technol.*, vol. 4, no. 10, pp. 2395–2402, 2017.

[7]. X.-H. Meng, Y.-X. Huang, D.-P. Rao, and Q. Liu, "Comparison of three data mining models for predicting diabetes or prediabetes by risk factors," *Kaohsiung J. Med. Sci.*, vol. 29, pp. 93–99, 2013.

[8]. V. AnujaKumari and R. Chithra, "Classification of diabetes disease using support vector machine," *Int. J. Eng. Sci.*, vol. 3, no. 2, pp. 1797–1801, 2013.