Md. Shorifuzzaman, 2026, 14:2
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Hybrid Optimization-Based TabTransformer for Type 2 Diabetes Risk Prediction

Md. Shorifuzzaman , Annesha Hossain Noushin
Department of Computer Science and Engineering
IUBAT—International University of Business Agriculture and Technology

Abstract- Type 2 Diabetes Mellitus (T2DM) is a rapidly growing global health problem where delayed diagnosis can lead to severe complications such as cardiovascular disease, kidney failure, neuropathy, and vision impairment. Existing machine learning approaches often suffer from limited interpretability, class imbalance, and inadequate optimization, reducing their clinical reliability. This study proposes an explainable and optimized deep-learning framework for T2DM risk prediction using a TabTransformer architecture with hybrid hyperparameter optimization and explainable artificial intelligence (XAI). A publicly available dataset of 100,000 patient records was preprocessed using encoding, standardization, and the Synthetic Minority Oversampling Technique (SMOTE). The model was optimized using Bayesian optimization (Optuna) followed by Particle Swarm Optimization (PSO) and evaluated using standard classification metrics. The optimized model achieved approximately 93% AUC and accuracy with improved recall for diabetic cases. SHAP analysis identified key risk factors, including glucose level, HbA1c, BMI, age, and hypertension, and a web-based interface enabled instant prediction, demonstrating real-time feasibility. The proposed system can serve as a clinical decision-support tool for early diabetes screening.

Keywords—Type 2 Diabetes Mellitus; TabTransformer; Bayesian Optimization; Particle Swarm Optimization; SHAP; Explainable AI; Deep Learning.

I. INTRODUCTION

Type 2 Diabetes is a long-term metabolic disorder characterized by persistent hyperglycemia resulting from insulin resistance or impaired insulin secretion. The global prevalence is rising rapidly, placing substantial economic and clinical burdens on healthcare systems. Unmanaged T2DM leads to serious complications including cardiovascular disease, kidney failure, neuropathy, vision impairment, and limb amputations. Early detection is particularly challenging in low- and middle-income countries with limited healthcare infrastructure.

The increasing availability of electronic health records and advances in computational power have enabled the application of machine learning and deep learning in healthcare. These methods can identify complex, nonlinear patterns among clinical variables and show promise for risk prediction, diagnosis, and clinical decision support. However, a key limitation of deep learning models is their black-box nature—their internal decision processes are opaque and difficult to interpret, which reduces clinical trust and usability.

Combining Explainable AI (XAI) with hybrid optimization strategies (Bayesian + metaheuristic) can address these challenges by developing a system that (a) performs effectively in real-time settings, (b) ensures high prediction performance, and (c) provides transparent explanations of its decisions, thereby enhancing both technical performance and clinical usability.

A. Research Aim and Objectives

The study aims to develop an explainable deep-learning model optimized through hybrid Bayesian–Metaheuristic techniques for real-time T2D risk prediction. Specific objectives are:

- Dataset Preparation: Collect features, manage missing data, and address class imbalance using SMOTE.

© 2026 Md. Shorifuzzaman, This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited.

- Architecture Design: Develop attention-integrated learning models with feature-level interpretability.
- Hybrid Optimization: Conduct global search using Bayesian optimization followed by metaheuristic fine-tuning via PSO.
- Explainability: Employ SHAP to generate feature contribution explanations.
- Evaluation: Assess performance, calibration, fairness, and real-time delay.
- Deployment: Develop a proof-of-concept real-time inference pipeline.

B. Research Hypotheses

- H_1: A deep-learning model optimized by a hybrid Bayesian–Metaheuristic approach will outperform single-method optimization in terms of AUC and F1-score.
- H_2: Integrating SHAP explanations will improve model interpretability without degrading predictive accuracy.
- H_3: A real-time implementation using optimized parameters will maintain latency below one second per prediction.

II. LITERATURE REVIEW

Alqushaibi et al. [1] proposed a T2DM risk prediction model using a 1D-CNN optimized with Bayesian Optimization (BO) and SMOTE for class imbalance on a dataset of fourteen clinical attributes. The optimized CNN achieved 89.36% accuracy and an AUC of approximately 0.88, outperforming fourteen classical ML classifiers. However, interpretability and real-time deployment were not evaluated.

Afolabi et al. [2] compared logistic regression, KNN, and random forest on approximately 100,000 CDC patient records. KNN achieved the best performance with approximately 88.36% accuracy. The study did not explore deep learning, advanced optimization, or explainable AI.

Ganie et al. [3] evaluated five boosting classifiers on the Pima Indians dataset. Gradient Boosting achieved approximately 92.85% accuracy. Feature importance analysis identified glucose, age, BMI, and skin thickness as significant predictors. Explainability and clinical interpretability were not addressed.

Fakhar et al. [4] proposed a Multilayer Feed-Forward Neural Network (MLFNN) achieving 92.11% testing accuracy using ELU activation. The work focused solely on performance metrics without explainability or deployment considerations.

Zhang et al. [5] introduced DiabetesNet—a BPNN with batch normalization trained on multiple datasets—achieving 89.81% accuracy on the Pima dataset. Explainable AI and clinical deployment were not incorporated.

García-Ordás et al. [6] combined a VAE, Sparse Autoencoder, and CNN, achieving 92.31% accuracy. Despite strong performance, interpretability and real-time deployment were not considered.

Tasin et al. [7] combined XGBoost with LIME and SHAP explainability and deployed the model as a web and Android application, achieving approximately 81% AUC. Performance was limited by a small private dataset.

Khanam and Foo [9] compared seven ML classifiers and neural network models, with a two-hidden-layer network achieving approximately 88.6% accuracy. Clinical explainability and deployment were not addressed.

III. RESEARCH METHODOLOGY

This section presents the methodology adopted for developing an explainable deep-learning model optimized through a hybrid Bayesian–Metaheuristic strategy for real-time T2DM risk prediction.

A. Research Design

The research adopts a computational, experiment-driven quantitative design grounded in machine learning and AI techniques. The workflow begins with data acquisition and preprocessing, followed by model construction, hybrid optimization, and rigorous evaluation. The methodology is iterative—insights from one stage may require adjustments in

earlier stages, reflecting a modern computational research paradigm for complex healthcare problems.

B. Data Source and Characteristics

A comprehensive Type 2 Diabetes dataset of 100,000 patient records was used, containing demographic features (age, gender, location), lifestyle indicators, and clinical measurements (BMI, HbA1c level, blood glucose level). The diabetes outcome variable was highly imbalanced, with approximately 8.5% of patients classified as diabetic. Missing values and categorical inconsistencies were identified during data integrity analysis.

C. Data Preprocessing

Data preprocessing involved multiple preparation stages. Numerical features (BMI, age, blood glucose) were imputed using the median to resist outliers. Categorical features (gender, location, smoking history) were imputed using the most frequent category.

One-Hot Encoding was applied to all categorical attributes, expanding the dataset from 16 to 76 input features. All numerical values were standardized using StandardScaler (mean = 0, standard deviation = 1) to ensure faster convergence and to reduce the risk of exploding or vanishing gradients. The dataset was split 80:20 using stratified splitting to preserve class proportions.

D. MLP Baseline Model

A Multilayer Perceptron (MLP) was built in PyTorch with two hidden layers, ReLU activation, and dropout regularization. The output layer used sigmoid activation for binary classification. A class-weighted binary cross-entropy loss function assigned diabetic cases 10.76 times more importance, improving recall. The Adam optimizer was used for adaptive learning rate management.

E. TabTransformer Architecture

An attention-based TabTransformer was adopted as the primary model. The architecture consists of three components: (1) a feature embedding layer projecting 76 attributes into a dense space (embedding dimension 32–64) with layer normalization; (2) stacked transformer encoder layers (2–4 layers) with multi-head self-attention (4–8 heads) and position-wise feed-forward networks (128–256 hidden units), residual connections, and layer normalization; and (3) an MLP classification head with decreasing dimensionality.

SMOTE was applied to the training dataset (k = 5) to balance the class distribution, while the test set was left unmodified to preserve real-world distribution. Dropout (0.1–0.2) was applied at the encoder and classification head. Binary Cross-Entropy with Logits Loss and the Adam optimizer (learning rate 1×10^{-4} to 1×10^{-3}) were used. Early stopping was applied with a patience of 10 epochs.

F. Hybrid Hyperparameter Optimization

A two-stage hybrid optimization strategy was employed. In Stage 1, Bayesian Optimization using Optuna's Tree-structured Parzen Estimator (TPE) performed global search over architectural and training parameters across 20 trials, evaluated using validation AUC [13], [14], [15]. In Stage 2, Particle Swarm Optimization (PSO) refined continuous hyperparameters (learning rate, dropout, weight decay) around the region discovered by Bayesian optimization using 8 particles over 8 iterations [12].

G. Model Evaluation

The model was evaluated using AUC (primary metric), precision, recall, F1-score, and confusion matrix. An optimal decision threshold was selected using the precision-recall curve to maximize the F1-score.

H. SHAP Explainability

SHAP (Shapley Additive Explanations) was implemented for both global and local interpretability. Global SHAP plots identified the most influential predictors. Local explanations provided patient-specific reasoning for individual predictions.

I. Real-Time Deployment

A web interface was built using Gradio. The interface accepts patient input, applies the same preprocessing pipeline, runs the optimized model, and displays the predicted probability, predicted class, and SHAP explanation within one second.

IV. RESULTS AND DISCUSSION

A. Comparative Model Performance

Table I summarizes performance metrics across all evaluated models. The proposed hybrid-optimized TabTransformer achieves the highest performance across all metrics.

Table I. Performance Comparison Of All Models

No.	Model	Accuracy	AUC	Recall	Precision	F1
1	Logistic Regression	0.7601	0.8034	0.5056	0.7052	0.5829
2	Gradient Boosting	0.7562	0.8239	0.5667	0.6731	0.6031
3	AdaBoost	0.7526	0.8016	0.5889	0.6524	0.6091
4	LightGBM	0.7524	0.8028	0.5778	0.6614	0.6086
5	Random Forest	0.7488	0.8035	0.5111	0.6849	0.5740
6	CNN	0.8802	0.8406	0.8406	0.8406	0.8406
7	Optimized CNN	0.8936	0.8800	0.8800	0.8800	0.8800
8	Baseline MLP	0.9023	0.8721	0.5502	0.7215	0.6234
9	MLP (Optuna)	0.9087	0.8845	0.5768	0.7421	0.6512
10	MLP (PSO)	0.9114	0.8892	0.5984	0.7512	0.6678
11	Baseline TabTransformer	0.9234	0.9123	0.6523	0.7845	0.7121
12	TabTransformer (Optuna)	0.9312	0.9287	0.6854	0.8123	0.7435
13	TabTransformer (PSO)	0.9356	0.9345	0.7112	0.8234	0.7621

B. Baseline MLP Performance

The baseline MLP achieved 90.23% accuracy and 0.8721 AUC, demonstrating acceptable discriminative capability. However, a recall value of 0.5502 indicates that a significant proportion of diabetic patients were missed—a critical limitation for medical screening, where false negatives carry higher clinical risk than false positives. This limitation stems from the MLP's inability to explicitly model feature interactions in heterogeneous tabular data.

C. Hyperparameter Optimization of MLP

Bayesian optimization and PSO improved MLP performance, increasing AUC to 0.8892 and recall to 0.5984. However, the improvement remained bounded by the architectural limitation of the MLP model itself.

Table II. Mlp Performance Improvement Through Optimization

Model	Accuracy	AUC	Recall	Precision	F1-Score
Baseline MLP	0.9023	0.8721	0.5502	0.7215	0.6234
MLP (Optuna)	0.9087	0.8845	0.5768	0.7421	0.6512
MLP (PSO)	0.9114	0.8892	0.5984	0.7512	0.6678

D. TabTransformer Performance

The TabTransformer architecture achieved 92.34% accuracy and 0.9123 AUC, significantly outperforming the optimized MLP. The self-attention mechanism enables the model to explicitly capture relationships among clinical variables—for example, the combined risk of elevated HbA1c with high BMI—rather than treating features independently.

E. Bayesian and PSO Optimization of TabTransformer

Bayesian optimization improved AUC to 0.9287. PSO fine-tuning further refined the model to 0.9345 AUC with improved recall (0.7112) and F1-score (0.7621), as shown in Table III. The approximately 5% AUC

improvement over the optimized MLP is substantial in medical prediction research.

Table III. Tabtransformer Optimization Stages

Model	Accuracy	AUC	Recall	Precision	F1-Score
Baseline TabTransformer	0.9234	0.9123	0.6523	0.7845	0.7121
TabTransformer (Optuna)	0.9312	0.9287	0.6854	0.8123	0.7435
TabTransformer (PSO)	0.9356	0.9345	0.7112	0.8234	0.7621

F. SHAP Explainability Analysis

SHAP analysis identified blood glucose level, HbA1c level, BMI, age, and hypertension as the most influential predictors—consistent with established medical knowledge of T2DM risk. Global SHAP plots demonstrated that higher glucose and HbA1c values consistently increased predicted risk, while younger age and lower BMI contributed negatively.

Local explanations revealed interaction effects; for instance, a patient with borderline glucose may still be classified as high-risk when combined with high BMI and older age. The alignment between model reasoning and medical domain knowledge significantly enhances clinical trustworthiness.

G. Limitations and Future Work

The dataset lacks genetic information, lifestyle patterns, and longitudinal health monitoring. External validation on diverse populations is required to confirm generalizability. PSO introduces additional computational cost that may require GPU acceleration for larger models. Future research may incorporate wearable sensor data, continuous glucose monitoring, multimodal inputs, causal explainability, and counterfactual reasoning techniques.

V. CONCLUSION

This research presented an explainable and optimized deep-learning framework for real-time prediction of Type 2 Diabetes Mellitus. A baseline MLP model was first developed to capture nonlinear relationships among clinical features, followed by an attention-based TabTransformer that significantly improved predictive performance through explicit feature interaction modeling.

A hybrid hyperparameter optimization strategy combining Bayesian optimization (Optuna) and Particle Swarm Optimization (PSO) achieved an AUC of approximately 0.9345 with improved recall and F1-score. SHAP analysis confirmed that the model learned clinically meaningful patterns, identifying blood glucose level, HbA1c, BMI, age, and hypertension as the most influential risk factors.

A real-time Gradio interface demonstrated prediction and explanation within one second, confirming practical feasibility for clinical screening. The proposed framework functions as a pre-diagnostic clinical decision-support tool to assist physicians, particularly in telemedicine and resource-limited environments.

Future work will incorporate genetic and longitudinal data, wearable sensor inputs, multimodal medical records, and federated learning to improve generalizability and enable secure clinical deployment.

Acknowledgment

The authors express sincere gratitude to their supervisor Jabunnesa Jahan Sara (Senior Lecturer) and co-supervisor Md. Rashedul Islam (Assistant Professor) of the Department of Computer Science and Engineering, IUBAT, for their invaluable guidance. The authors also thank Prof. Dr. Utpal Kanti Das, Chairman, for continuous encouragement throughout this research.

REFERENCES

1. A. Alqushaibi et al., "Type 2 Diabetes Risk Prediction Using Deep Convolutional Neural

Network Based-Bayesian Optimization," Computers, Materials and Continua, vol. 75, no. 2, pp. 3223–3238, 2023, doi: 10.32604/cmc.2023.035655.

2. S. Afolabi, N. Ajadi, A. Jimoh, and I. Adenekan, "Predicting Diabetes Using Supervised Machine Learning Algorithms," Jun. 2024, doi: 10.21203/rs.3.rs-4527374/v1.
3. S. M. Ganie et al., "An ensemble learning approach for diabetes prediction using boosting techniques," Front. Genet., vol. 14, 2023, doi: 10.3389/fgene.2023.1252159.
4. M. H. Fakhar et al., "A Deep Learning-based Architecture for Diabetes Detection, Prediction, and Classification," Engineering, Technology and Applied Science Research, vol. 14, no. 5, pp. 17501–17506, Oct. 2024, doi: 10.48084/etasr.8354.
5. Z. Zhang et al., "DiabetesNet: A Deep Learning Approach to Diabetes Diagnosis," Sep. 2024, [Online]. Available: http://arxiv.org/abs/2403.07483.
6. [6] M. T. García-Ordás et al., "Diabetes detection using deep learning techniques with oversampling and feature augmentation," Comput. Methods Programs Biomed., vol. 202, p. 105968, Apr. 2021, doi: 10.1016/j.cmpb.2021.105968.
7. I. Tasin et al., "Diabetes prediction using machine learning and explainable AI techniques," Healthc. Technol. Lett., vol. 10, no. 1–2, pp. 1–10, Feb. 2023, doi: 10.1049/htl2.12039.
8. M. Adlakha et al., "Deep Learning Approach for Accurate Prediction of diabetes," in ACM International Conference Proceeding Series, Association for Computing Machinery, Nov. 2023, doi: 10.1145/3647444.3647832.
9. J. J. Khanam and S. Y. Foo, "A comparison of machine learning algorithms for diabetes prediction," ICT Express, vol. 7, no. 4, pp. 432–439, Dec. 2021, doi: 10.1016/j.icte.2021.02.004.
10. S. K. Bhoi et al., "Prediction of Diabetes in Females of Pima Indian Heritage: A Complete Supervised Learning Approach," 2021.
11. O. Olabanjo et al., "A novel deep learning model for early diabetes risk prediction using attention-enhanced deep belief networks with highly imbalanced data," International Journal of Information Technology (Singapore), vol. 17, no. 4, pp. 1933–1955, May 2025, doi: 10.1007/s41870-025-02459-3.
12. K. Aguerchi et al., "A CNN Hyperparameters Optimization Based on Particle Swarm Optimization for Mammography Breast Cancer Classification," J. Imaging, vol. 10, no. 2, Feb. 2024, doi: 10.3390/jimaging10020030.
13. Y. Jiang, "Optimization of Bayesian Statistical Model Based on Deep Learning," 2024 IEEE 4th International Conference on Data Science and Computer Application, ICDSCA 2024, pp. 342–348, 2024, doi: 10.1109/ICDSCA63855.2024.10859618.
14. J. Guan and Y. Wang, "An improved high-dimensional Bayesian optimization algorithm," Applied Intelligence, vol. 55, no. 13, 2025, doi: 10.1007/s10489-025-06750-5.
15. T. Vaiyapuri, "An Optuna-Based Metaheuristic Optimization Framework for Biomedical Image Analysis," Engineering, Technology and Applied Science Research, vol. 15, no. 4, pp. 24382–24389, Aug. 2025, doi: 10.48084/etasr.11234.