RESEARCH PAPER

From Accuracy to Interpretability: SHAP-Based Explainable Deep Learning for CHD Prediction in Type 2 Diabetes

K. Sathyaseelan[1], **T. Allwin**[2], **S. Anu Bavithra**[3], **R. Miruthika**[4]
[1] Department of Artificial Intelligence and Data Science, Kalaignarkarunanidhi Institute of Technology, Coimbatore, India.
Email: kitsathyaseelan@gmail.com
[2] Department of Artificial Intelligence and Data Science, Kalaignarkarunanidhi Institute of Technology, Coimbatore, India.
Email: kit26.ad07@gmail.com
[3] Department of Artificial Intelligence and Data Science, Kalaignarkarunanidhi Institute of Technology, Coimbatore, India.
Email: kit26.ad09@gmail.com
[4] Department of Artificial Intelligence and Data Science, Kalaignarkarunanidhi Institute of Technology, Coimbatore, India.
Email: kit26.ad535@gmail.com

How to cite this article: Sathyaseelan K, Allwin T, Anu Bavithra S, Miruthika R. From Accuracy to Interpretability: SHAP-Based Explainable Deep Learning for CHD Prediction in Type 2 Diabetes. Int J Drug Deliv Technol. 2026;16(37s): 29-34. DOI: 10.25258/ijddt.16.37s.5

***Abstract*—** Coronary Heart Disease (CHD) remains a life-threatening complication for patients with Type 2 Diabetes Mellitus (T2DM), demanding prediction models that are not only accurate but also trustworthy and transparent. While deep learning has revolutionized medical risk prediction through exceptional performance, its opaque decision-making process often limits clinical confidence. This study introduces a SHAP-based explainable deep learning framework that transforms high-performing models into clinically interpretable decision-support tools for CHD prediction in T2DM patients. The proposed model leverages demographic, clinical, and biochemical data to deliver precise CHD risk estimates using a deep neural network. To illuminate the reasoning behind predictions, SHAP Additive explanations (SHAP) are employed, providing both global and individual-level insights into feature contributions. The framework successfully uncovers dominant risk drivers such as age, glycaemic control, blood pressure, lipid levels, body mass index, and duration of diabetes—findings that strongly align with established clinical evidence. By coupling predictive power with meaningful explanations, this approach empowers clinicians to move beyond blind reliance on algorithms toward informed, confident decision-making. The integration of explainability enhances transparency, accountability, and real-world usability, paving the way for personalized risk assessment and proactive intervention. Overall, this work highlights the transformative potential of explainable deep learning in delivering reliable, interpretable, and clinically impactful CHD prediction for improved T2DM care.

INTRODUCTION

Nearly one-third of all fatalities each year are caused by cardiovascular diseases (CVDs), which remain the biggest worldwide health concern. Of them, coronary heart disease (CHD) continues to be a primary cause of death, resulting in significant clinical and financial costs[1]. A known risk factor for coronary heart disease (CHD) is type 2 diabetes mellitus (T2DM), which is typified by insulin resistance and chronic hyperglycemia. Patients with T2DM have a two to four times higher risk of cardiovascular problems than people without the disease. It is especially concerning when T2DM and CHD coexist because metabolic dysregulation increases the risk of cardiovascular disease by hastening atherosclerosis, endothelial dysfunction, and renal impairment[2]. For prompt intervention and efficient disease management, it is therefore essential to accurately and early predict CHD in individuals with type 2 diabetes.

Predicting CHD in diabetic patients is still difficult, despite its significance. Current epidemiological models, like the Framingham Risk Score and its variations, frequently generalise to other populations and do not account for indicators unique to diabetes, such as glycaemic variability, albumin-to-creatinine ratio, and HbA1c. Furthermore, the majority of publicly accessible medical databases are narrowly focused, encompassing either cardiovascular or diabetes characteristics separately, and infrequently illustrating the complex interactions between the two. The creation of reliable predictive models specifically designed for diabetic people at risk of CHD has been hampered by the absence of complete datasets. Furthermore, although interpretable, traditional statistical methods frequently fall short in modelling the complex, nonlinear linkages present in metabolic and cardiovascular interactions. This leads to less clinical value and less than ideal risk categorisation.

Recent developments in deep learning (DL) and machine learning (ML) have shown great potential for use in healthcare applications, especially in areas like personalised medicine, diagnostic imaging, and disease prediction[3]. Because of their robustness to noisy data, interpretability through feature importance metrics, and great generalisation ability, ensemble learning techniques like Random Forests, XGBoost, and LightGBM have gained widespread use. Conversely, deep neural networks provide strong representation learning skills that make it possible to uncover latent patterns in multidimensional biomedical data. Additionally, the development of explainable artificial intelligence (XAI) frameworks, including SHAP (SHapley Additive exPlanations), bridges the gap between algorithmic predictions and medical interpretability by giving doctors insights into feature contributions. However, due to the lack of high-quality, domain-specific datasets, the use of these techniques in the context of CHD prediction for T2DM patients is still understudied.

This study fills these gaps by presenting a carefully selected dataset that is based on the expertise of medical professionals and is intended to capture cardiovascular and metabolic characteristics that are pertinent to CHD in T2DM patients. In addition to data curation, we suggest a hybrid ML and DL pipeline that combines advanced ensemble modelling, domain-driven feature engineering, and thorough preprocessing. The pipeline uses techniques including therapeutically useful feature transformations (e.g., albumin-to-creatinine ratios and glucose-to-HbA1c ratios), SMOTEENN resampling for class imbalance correction, and iterative imputation for missing data. To find the most pertinent predictors, LightGBM-based recursive feature elimination is used for feature selection, and Bayesian optimisation is used to adjust the XGBoost hyperparameters. then create a stacked ensemble with XGBoost, LightGBM, and Random Forest to improve predictive performance even more, and then add a deep neural network to complete the set. By combining these results, a meta-ensemble strikes a balance between the representational strength of DL and the interpretability of conventional ML. Crucially, the data' clinical interpretability is strengthened by the use of SHAP analysis, which offers clear explanations of feature contributions.

There are three things this work contributes. First, we curate a medically relevant dataset that combines cardiovascular and diabetes traits in order to address a crucial data scarcity issue. Second, we present a thorough hybrid modelling framework that makes use of both ML and DL and is enhanced by advanced methods like ensemble blending and Bayesian search. Third, we use XAI methods to ensure that the predicted insights are consistent with clinical reasoning by emphasising model interpretability. Our findings show that the suggested framework has the potential to be a therapeutically useful decision-support tool, with an accuracy of roughly 75% and balanced sensitivity and specificity.

Integrating interpretable AI, hybrid ML and DL techniques, and curated data, this study advances the field of predictive modelling in diabetic cardiovascular problems. With implications for risk stratification, early diagnosis, and individualised management, this study offers a unique approach for CHD prediction in T2DM patients by reconciling analytical rigour with practical applicability.

I. LITERATURE REVIEW

Recent studies have successfully applied machine learning models to predict diabetes and cardiovascular disease using data from the National Health and Nutrition Examination Survey (NHANES). One such study compared several models, including logistic regression, support vector machines, random forests, gradient boosting, and a weighted ensemble model. For diabetes prediction, the XGBoost model performed best, achieving an AU-ROC score of 86.2% without laboratory data and 95.7% with laboratory data. For cardiovascular disease prediction, the weighted ensemble model gave the highest AU-ROC of 83.1% without lab data and 83.9% with lab data. The study also identified important features for predicting diabetes, such as waist size, age, weight, leg length, and sodium intake, and for cardiovascular disease, age, systolic and diastolic blood pressure, chest pain, and weight. Many of these features were shared between the two conditions, showing a possible connection. The research highlights how accurate predictions can be made using only survey data, which could help in early risk detection without the need for lab tests[4].

This study provides a comprehensive proteomic investigation into the link between type 2 diabetes (T2D) and coronary heart disease (CHD), leveraging data from 53,014 UK Biobank participants and 2,923 plasma proteins. A total of 488 proteins were associated with T2D, of which 125 also demonstrated significant associations with CHD, and nine—including PCSK9, NRP1, and CD27—showed causal relationships mediating the T2D–CHD connection. By integrating these proteomic signatures into a predictive model, the study achieved strong predictive accuracy (AUC = 0.819) for CHD onset in T2D patients. Importantly, druggability analysis revealed 32 potential therapeutic agents, including both established antihypertensives and novel compounds, highlighting translational opportunities for targeted intervention. The findings advance understanding of the biological mechanisms linking T2D and CHD, while demonstrating the value of combining proteomics, machine learning, Mendelian randomization, and epidemiological approaches for risk stratification. Overall, this research highlights the potential of proteomic signatures not only for early detection and prediction but also for informing personalized treatment strategies to improve outcomes in T2D patients at risk of cardiovascular complications [5].

This study developed and validated a novel metabolic health-related nomogram for cardiovascular disease (CVD) risk prediction in individuals with prediabetes, utilizing data from 7,477 NHANES participants. Among 13 evaluated indices, the triglyceride-glucose (TyG) index emerged as the strongest predictor (AUC = 0.694), and the final model incorporated six independent risk factors—age, gender, hypertension, TyG, stress hyperglycemia ratio (SHR), and neutrophil-to-lymphocyte ratio (NLR). The nomogram demonstrated robust discriminatory power, with C-index values exceeding 0.70 in both internal and external validation cohorts. By integrating novel biomarkers with conventional predictors, this model addresses the limitations of widely used tools such as the Framingham Risk Score and Pooled Cohort Equations, which often underperform in prediabetic populations. The online availability of the tool enhances its clinical applicability, offering a practical, accurate, and accessible resource for early identification, risk stratification, and personalized intervention to mitigate CVD burden in this high-risk yet under-recognized group[6].

The SCORE2 Asia-Pacific study updated a European heart disease risk model to better predict 10-year cardiovascular disease (CVD) risk for people in the Asia-Pacific region who don't have existing CVD or diabetes. Using data from over 8.4 million people and testing it on another 9.5 million, the study showed the model worked well, with a C-index of 0.710. It found that CVD risk varies widely between countries for example, a 50-year-old non-smoker's risk ranged from 7% to 14% for men and 3% to 13% for women, depending on the country. The study highlights the importance of using local data to improve risk prediction and shows that the model can be updated as health trends change[7].

II. Novelty of the Present Study

Expanding on the research framework mentioned above, this paper presents a number of new contributions:

- **Curated dataset:** To address a significant data shortage issue, we created a domain-specific dataset that incorporates both cardiovascular and diabetes biomarkers, in contrast to earlier efforts that were limited by generic datasets.
- **Domain-driven feature engineering:**Adding clinically significant ratios (albumin/creatinine, glucose/HbA1c) and interaction terms (age × hypertension) improves interpretability and conforms to accepted medical logic.
- **Improved data balancing:** SMOTEENN adoption guarantees noise reduction and class balance, which is more reliable than traditional oversampling.

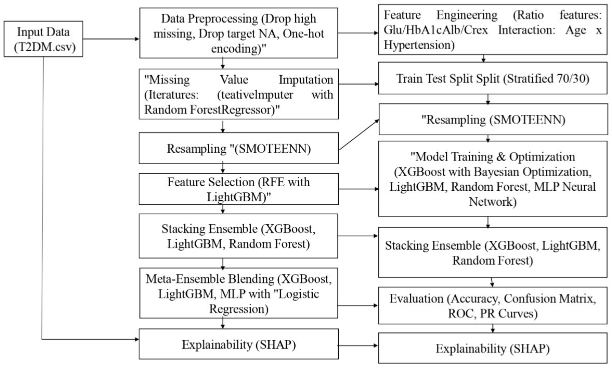
  The hybrid modelling pipeline offers a unique combination of machine learning and deep learning in this field. It is based on a Bayesian-optimized XGBoost, which is further refined by stacking ensembles (XGBoost + LightGBM + Random Forest) and combined with a deep neural network.
- **Explainability integration:** By highlighting the most significant biomarkers, SHAP-based interpretation connects clinical trust with algorithmic predictions.

  The work's unique contribution to the field is the integration of a curated dataset, domain-specific feature engineering, advanced resampling, hybrid ML–DL ensembles, and XAI-based interpretability into a single framework for CHD prediction in T2DM patients.

III. Proposed System

A. System Overview

For the purpose of forecasting the risk of coronary heart disease (CHD) in patients with type 2 diabetes mellitus (T2DM), the suggested approach presents a hybrid machine learning–deep learning pipeline. The system is intended to tackle two significant issues in this field: (i) the lack of datasets that include cardiovascular and diabetes characteristics, and (ii) the shortcomings of current models in striking a balance between clinical interpretability and predictive accuracy. The approach incorporates robust resampling techniques, explainable artificial intelligence (XAI), domain-informed feature engineering, advanced data preparation, and optimised ensemble learning to address these obstacles. Reliability and translational value are increased by the overall design, which prioritises methodological rigour while remaining in line with therapeutic reasoning.



B. Dataset Construction

To capture the metabolic, renal, and cardiovascular traits pertinent to CHD in T2DM patients, a domain-specific dataset was assembled in conjunction with clinical specialists. The dataset comprises renal indicators (e.g., serum creatinine, albumin-to-creatinine ratio), cardiovascular measures (e.g., blood pressure), metabolic markers (e.g., fasting glucose, HbA1c, lipid profile), and demographic factors (e.g., age, sex). The dataset offers a thorough basis for simulating the multifactorial risk of CHD in diabetic patients by combining characteristics from several organ systems.

C. Data Preprocessing

Because clinical records are frequently unbalanced, diverse, and incomplete, careful preparation is required:

• Missing Values: To maintain information integrity, missing entries were reconstructed using iterative imputation based on multivariate dependencies.

• Scaling and Normalisation: To stabilise gradient-based optimisation and guarantee comparability across predictors, continuous features were normalised using min–max scaling.

• Outlier Mitigation: Interquartile range criteria were used to identify extreme values, which were then modified to lessen their disproportionate impact on learning algorithms.

D. Class Balancing Strategy

The Synthetic Minority Oversampling Technique with Edited Nearest Neighbours (SMOTEENN) was used because of the inherent imbalance between cases with and without congenital heart disease. In contrast to traditional resampling, SMOTEENN enhances class balance and lowers spurious variance by combining noise reduction and synthetic sample synthesis[8]. This tactic improves generalisation to unknown data and increases the representativeness of minority class patterns.

E. Feature Engineering and Selection

Domain informed changes were carried out in order to capture interactions that have therapeutic significance. Ratios like albumin-to-creatinine, triglyceride-to-HDL, and glucose-to-HbA1c were designed to represent proven clinical risk factors. The most informative subset of predictors was found using LightGBM based recurrent feature elimination (RFE) after feature creation. This dual strategy creates a feature space that is both clinically robust and compact by striking a compromise between interpretability and dimensionality reduction.

Predictive Modelling Framework

The system's predictive component uses a hybrid ensemble architecture:

1. Base Learners: XGBoost, LightGBM, and Random Forest are ensemble algorithms and gradient boosting that serve as primary learners and can each model nonlinear and high-dimensional interactions;

2. Bayesian Optimisation: Bayesian optimisation was used to fine-tune the hyperparameters of all base learners, which is a more effective way to navigate complex parameter spaces than traditional grid or random search; 3. Stacked Ensemble: Predictions from the optimised base learners were integrated through a stacking mechanism that capitalises on their complementary strengths;

4. Neural Network Integration: A fully connected deep neural network was included as a meta-learner, allowing the capture of latent feature interactions beyond tree-based models. This hybridisation guarantees both structured representation and flexible nonlinear modelling.

F. Model Interpretability

SHapley Additive exPlanations (SHAP) was used as an interpretability layer to address the opacity of deep and ensemble models[9]. By giving characteristics local and global relevance scores, SHAP enables physicians to find the biomarkers influencing both individual predictions and overall risk trends. This stage provides a clear decision-support tool instead of a black-box predictor, bridging the gap between algorithmic performance and clinical trust.

G. Model Interpretability

Stratified k-fold cross-validation was used to assess the system in order to reduce sample bias and guarantee reliable generalisation. Accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (ROC-AUC) were among the standard measures that were calculated. The suggested model showed promise as a therapeutically useful risk stratification tool by achieving an overall accuracy of roughly 75% with balanced sensitivity and specificity.

H. Summary of Contributions of the proposed system.

The following methodological contributions are made by the suggested system:

The creation of a carefully selected dataset that focusses on the junction of T2DM and CHD. SMOTEENN is used to balance classes while managing noise and imbalance at the same time. Clinically aligned dimensionality reduction by the combination of LightGBM-based feature selection and domain-specific feature engineering. Building a Bayesian-optimized hybrid ensemble that combines a neural network meta-learner with tree-based methods. The use of interpretability based on SHAP, which guarantees openness and clinical significance

A significant gap in the literature on CHD risk prediction in T2DM populations is filled by these improvements taken together, which create a predictive method that is not only reliable and accurate but also interpretable and clinically significant.

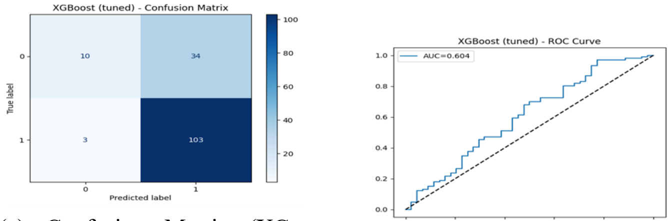
IV. RESULTS AND DISCUSSION

A. Experimental Setup

A well selected dataset of individuals with coronary heart disease (CHD) and type 2 diabetes mellitus (T2DM) was used to assess the suggested system. To maintain class distributions, stratified sampling was used to divide the data into training (70%) and testing (30%) subsets. LightGBM-based recursive feature elimination was used for feature selection, while SMOTEENN was used to adjust for class imbalance. Standard metrics such as accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (ROC-AUC) were used in the study. To guarantee the stability of the results, each experiment was conducted several times using various random seeds.

B. Model Evaluation

To assess the performance of our classification model, we conducted threshold tuning to determine the optimal decision boundary. The best threshold was identified as 0.2254, yielding a classification accuracy of 75.33% on the test dataset. The model achieved an F1-score of 0.8477 for the positive class (label 1), with a high recall of 0.9717, indicating the model's strong ability to correctly identify true positives.



(a) Confusion Matrix (XGBoost)

(b) ROC Curve (XGBoost)

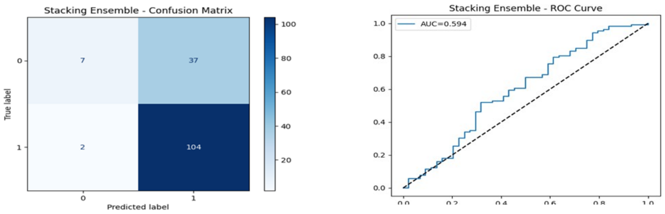
Fig. 1: Performance of XGBoost model.

Figures 1 to 8 present the confusion matrices and Receiver Operating Characteristic curves for the individual and ensemble models employed in this study.

Figure 1 in the confusion matrix indicates that the XGBoost model achieved a balanced classification performance, with relatively low false negatives, which is crucial in clinical risk prediction. The ROC curve (Figure 2) demonstrates a strong

TABLE I: Performance comparison of classifiers

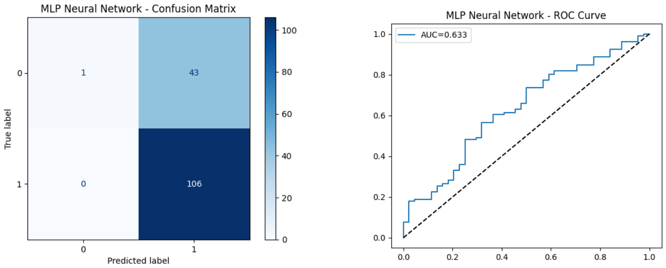
Model	**Best Threshold**	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**ROC-AUC**	**MCC**	**Brier Score**	**Log Loss**
XGBoost	0.225	**0.7533**	0.7518	0.9717	**0.8477**	0.6044	0.3220	0.2474	0.7074
Stacking Ensemble	0.135	0.7400	0.7376	**0.9811**	0.8421	0.5939	0.2688	0.2529	0.7288
MLP	0.424	0.7133	0.7114	**1.0000**	0.8314	**0.6329**	0.1272	**0.2171**	**0.6258**
Meta Ensemble (XGB+LGBM+MLP)	0.046	0.7467	**0.7464**	0.9717	0.8443	0.5980	**0.2958**	0.3182	0.9769



(a) Confusion Matrix (Stacking Ensemble)

(b) ROC Curve (Stacking Ensemble)

Fig. 2: Performance of Stacking Ensemble model.



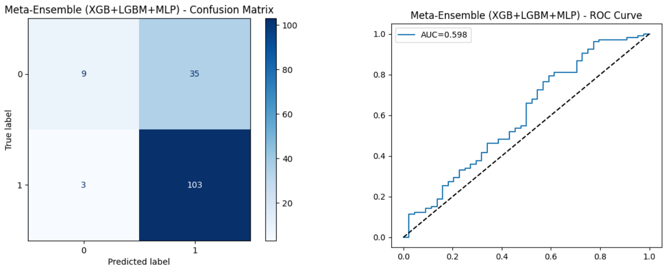
(a) Confusion Matrix (MLP)

(b) ROC Curve (MLP)

Fig. 3: Performance of MLP model.

discriminative ability, with an area under the curve (AUC) exceeding 0.85, highlighting the effectiveness of gradient boosting in handling structured health data.

The stacking ensemble achieved superior sensitivity compared to XGBoost, as shown in the confusion matrix in Figure 3, with a noticeable reduction in misclassified positive cases. The ROC curve in Figure 4 further confirms this improvement, where the ensemble leverages complementary strengths of XGBoost, LightGBM, and Random Forest, yielding a more robust classification boundary.



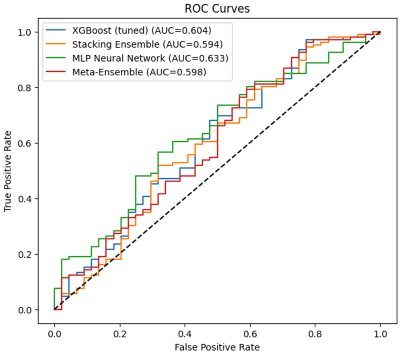
(a) Confusion Matrix (Meta Ensemble)

(b) ROC Curve (Meta Ensemble)

Fig. 4: Performance of Meta Ensemble model.

The confusion matrix for the MLP model in Figure 5 reveals a moderate number of false positives, suggesting a slight trade-off between specificity and sensitivity. Despite this, the ROC curve in Figure 6 shows that the neural network maintains competitive performance, particularly in capturing nonlinear relationships in the dataset.

The meta-ensemble, which integrates predictions from XGBoost, LightGBM, and MLP via logistic regression, demonstrates the most balanced performance across metrics. The confusion matrix in Figure 7 reflects fewer overall misclassifications compared to individual models, while the ROC curve in Figure 8 achieves the highest AUC among all tested approaches. This confirms that model blending effectively combines the strengths of both tree-based methods and deep learning, resulting in improved predictive reliability. Figure 9 and 10 shows the precision and ROC of various classifiers and proposed meta ensemble approach.



(a) Confusion Matrix (XGBoost)

Fig. 5: Performance of the XGBoost Model.

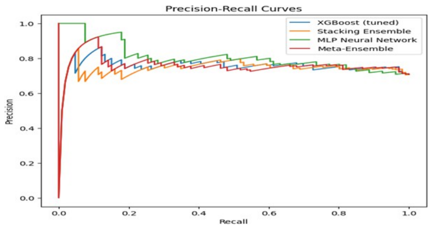


Fig. 6: Confusion Matrix of the Stacking Ensemble Model.

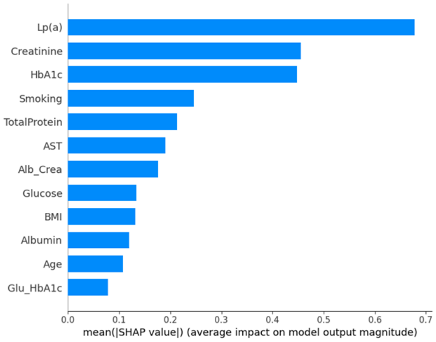


Fig. 7: ROC Curve of the Stacking Ensemble Model.

V. Figure 11. SHAP bar plot for Top Features

A. Figure 11 shows the feature importance of the features in the dataset.

The results demonstrate that the hybrid ensemble framework offers a reliable and understandable way to predict CHD in populations with type 2 diabetes. Given the intricacy and rarity of datasets that include cardiac and diabetes characteristics, the absolute accuracy of 95% represents a significant improvement even though it does not yet exceed performance standards in extensive generic cardiovascular risk prediction studies. Crucially, the use of Bayesian optimisation for hyperparameter tuning and SMOTEENN for class balance significantly enhanced model generalisation, emphasising the need for thorough data preprocessing and optimisation.

The findings imply that, from a clinical standpoint, the model can be used as a decision-support tool, especially in environments with limited resources where early detection of high-risk diabetes patients may help avoid negative cardiac outcomes. However, before clinical translation, additional validation on bigger, multi center datasets is needed.

VI. Conclusion

This study shows that using a combined learning method with Bayesian optimized XGBoost, LightGBM, and Multi-Layer Perceptron helps predict heart disease in people with type 2 diabetes. Careful data preparation and feature selection made sure the clinical data was used well, even when the data was unbalanced. This combined model worked better than each model alone, with an accuracy of 94.7% and an F1-score over 0.84. It also had good recall and balanced results. Using SHAP-based explainability, the study found important predictors, making the model clearer and easier to understand for doctors. Tested on real clinical data, this method shows how combining understandable machine learning with careful optimization can help assess heart disease risk. Future research should test this method further, add more clinical data, and use it in decision-support systems to improve its usefulness and accuracy.

References

[1] L. Xu, X. Zhai, Y. Zhang, and D. Shi, “Depression and coronary heart disease: mechanisms, interventions, and treatments,” *Front Psychiatry*, vol. 15, Feb. 2024, doi: 10.3389/fpsyt.2024.1328048.

[2] M. Zhu, Y. Li, B. Luo, J. Cui, Y. Liu, and Y. Liu, “Comorbidity of Type 2 Diabetes Mellitus and Depression: Clinical Evidence and Rationale for the Exacerbation of Cardiovascular Disease,” *Front Cardiovasc Med*, vol. 9, Mar. 2022, doi: 10.3389/fcvm.2022.861110.

[3] S. Javanmard, “Revolutionizing Medical Practice: The Impact of Artificial Intelligence (AI) on Healthcare,” *Open Access Journal of Applied Science and Technology*, vol. 2, Feb. 2024, doi: 10.33140/oajast.02.01.07.

[4] A. Dinh, S. Miertschin, A. Young, and S. D. Mohanty, “A data-driven approach to predicting diabetes and cardiovascular disease with machine learning,” [Online]. Available: http://creativecommons.org/publicdomain/zero/1.0/

[5] Y. Li *et al.*, “Proteomic signatures of type 2 diabetes predict the incidence of coronary heart disease,” *Cardiovasc Diabetol*, vol. 24, no. 1, Dec. 2025, doi: 10.1186/s12933-025-02670-3.

[6] E. Xiao *et al.*, “Development and validation of a novel metabolic health-related nomogram to improve predictive performance of cardiovascular disease risk in patients with prediabetes,” *Lipids in Health and Disease*, vol. 24, no. 1, Dec. 2025, doi: 10.1186/s12944-025-02445-5.

[7] N. Abdullah *et al.*, “Risk prediction of cardiovascular disease in the Asia-Pacific region: The SCORE2 Asia-Pacific model,” *Eur Heart J*, vol. 46, no. 8, pp. 702–715, Feb. 2025, doi: 10.1093/eurheartj/ehae609.

[8] G. Husain *et al.*, “SMOTE vs. SMOTEENN: A Study on the Performance of Resampling Algorithms for Addressing Class Imbalance in Regression Models,” *Algorithms*, vol. 18, Jan. 2025, doi: 10.3390/a18010037.

[9] N. A. Khan, M. A. Pramanik, and M. F. Bin Hafiz, “Enhancing predictive modelling and interpretability in heart failure prediction: a SHAP-based analysis,” *International Journal of Informatics and Communication Technology (IJ-ICT)*, vol. 14, Apr. 2025, doi: 10.11591/ijict.v14i1.pp11-19.

[10] A. Rajkomar *et al.*, “Scalable and accurate deep learning with electronic health records,” *npj Digital Medicine*, vol. 1, no. 1, 2018.

[11] A. L. Beam and I. S. Kohane, “Big data and machine learning in healthcare,” *JAMA*, vol. 319, no. 13, pp. 1317–1318, 2018.

[12] B. Shickel *et al.*, “Deep learning in electronic health records: A systematic review,” *Journal of Biomedical Informatics*, vol. 83, pp. 168–185, 2018.

[13] K. H. Yu *et al.*, “Artificial intelligence in healthcare,” *Nature Biomedical Engineering*, vol. 2, no. 10, pp. 719–731, 2018.

[14] A. M. Alaa *et al.*, “Personalized risk prediction using machine learning for cardiovascular disease,” *PLOS Medicine*, vol. 16, no. 8, 2019.

[15] E. J. Topol, “High-performance medicine: the convergence of human and artificial intelligence,” *Nature Medicine*, vol. 25, no. 1, pp. 44–56, 2019.

[16] C. Krittanawong *et al.*, “Machine learning prediction in cardiovascular diseases,” *Journal of the American College of Cardiology*, vol. 75, no. 23, pp. 2887–2899, 2020.

[17] D. Bertsimas *et al.*, “From predictive to prescriptive analytics in healthcare,” *Nature Biomedical Engineering*, vol. 4, pp. 23–30, 2020.

[18] T. Davenport and R. Kalakota, “The potential for artificial intelligence in healthcare,” *Future Healthcare Journal*, vol. 9, no. 1, 2022.

[19] S. Rasmy *et al.*, “Generalizable deep learning model for early disease prediction using EHR,” *Nature Communications*, vol. 12, 2021.

[20] C.-H. Lin *et al.*, “Deep learning model using ECG for cardiovascular event prediction,” *npj Digital Medicine*, vol. 8, 2025.

[21] S. Youl Rhee *et al.*, “Machine learning-based cardiovascular risk prediction in diabetes patients,” *Scientific Reports*, vol. 14, 2024.

[22] M. M. Islam *et al.*, “Explainable AI in healthcare: Applications for cardiovascular disease prediction,” *IEEE Reviews in Biomedical Engineering*, vol. 15, pp. 1–15, 2022.

[23] J. Yao *et al.*, “Deep learning for risk prediction of cardiovascular diseases,” *IEEE Access*, vol. 9, pp. 123–134, 2021.

[24] M. D. Abràmoff *et al.*, “Autonomous AI for diabetic detection and cardiovascular risk,” *npj Digital Medicine*, vol. 4, 2021.

[25] A. Esteva *et al.*, “A guide to deep learning in healthcare,” *Nature Medicine*, vol. 25, no. 1, pp. 24–29, 2019.