IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Transformer-Enhanced Ensemble Learning For Interpretable Cardiovascular Disease Prediction

[1]Vaishnavi Kishor Yadav, [2]Dr. Prashant Laxmanrao Paikrao

[1]PG Scholar, Dept. of EXTC, Government College of Engineering, Amravati, India
[2]Assistant Professor, Dept. of EXTC, Government College of Engineering, Amravati, India

Abstract: Cardiovascular diseases remain one of the leading causes of mortality worldwide, making early detection and risk prediction essential for effective clinical intervention and preventive healthcare. In recent years, machine learning techniques have demonstrated significant potential in assisting clinicians by analyzing complex medical datasets and identifying hidden patterns associated with cardiovascular conditions. However, many existing prediction models suffer from limitations such as insufficient feature representation, lack of robustness, and limited interpretability, which restrict their practical adoption in clinical environments. To address these challenges, this study proposes a Transformer-Enhanced Ensemble Learning framework for interpretable cardiovascular disease prediction. The proposed approach integrates transformer-based feature representation with ensemble machine learning models to improve prediction accuracy and model reliability. The transformer encoder captures complex relationships among clinical attributes using attention mechanisms, while the ensemble classification strategy aggregates predictions from multiple classifiers to enhance generalization performance. Furthermore, explainable artificial intelligence techniques are incorporated to provide transparency and interpretability by identifying the most influential clinical risk factors affecting prediction outcomes. Experimental evaluation conducted on benchmark cardiovascular disease datasets demonstrates that the proposed framework achieves superior performance compared with conventional machine learning models in terms of accuracy, precision, recall, F1-score, and area under the ROC curve. The interpretability analysis further reveals clinically relevant features that contribute significantly to cardiovascular risk prediction. The results indicate that the integration of transformer architectures, ensemble learning, and explainable AI provides a reliable and transparent framework for cardiovascular disease prediction and can support clinicians in early diagnosis and decision-making processes.

Index Terms **-** cardiovascular disease prediction; Transformer models; Ensemble learning; Explainable artificial intelligence; Machine learning; Clinical decision support systems; Healthcare analytics.

1. Introduction

Cardiovascular diseases represent a major global health concern, accounting for a significant proportion of mortality and healthcare expenditure. Early diagnosis of heart disease is crucial for reducing complications and improving patient outcomes. With the increasing availability of clinical datasets and electronic health records, machine learning techniques have emerged as valuable tools for predicting cardiovascular risk using patient data [1], [2].

Traditional prediction models such as logistic regression, decision trees, and support vector machines have been widely used in cardiovascular disease prediction tasks. These approaches analyze clinical attributes including cholesterol levels, blood pressure, age, and electrocardiographic indicators to identify patterns associated with cardiovascular risk [4], [5]. While these models provide useful diagnostic assistance, their performance may be limited when handling complex feature interactions present in clinical datasets.

To improve predictive performance, ensemble learning techniques have been introduced to combine multiple classifiers. Ensemble models such as random forests, gradient boosting, and stacking frameworks reduce prediction variance and improve generalization by aggregating outputs from multiple learners [7], [8]. Several studies have demonstrated that ensemble learning improves cardiovascular disease prediction accuracy compared with single-model approaches [9].

Recent advancements in deep learning have introduced transformer architectures that utilize attention mechanisms to capture dependencies among input features. Originally developed for natural language processing, transformers have been increasingly applied to healthcare data analysis due to their ability to model complex relationships within structured datasets [3], [21]. These models enable improved feature representation, which is critical for accurate clinical prediction.

Another key challenge in medical AI systems is interpretability. Clinicians require transparent models that explain how predictions are generated. Explainable AI techniques such as SHAP provide feature-level explanations that help identify the clinical factors influencing model decisions [10], [11].

Motivated by these developments, this study proposes a Transformer-Enhanced Ensemble Learning framework that integrates attention-based feature representation with ensemble classification and explainable AI. The objective is to improve prediction accuracy while maintaining interpretability for cardiovascular disease prediction.

The main contributions of this study include:

- Integration of transformer-based feature representation with ensemble learning
- Development of an interpretable cardiovascular disease prediction model
- Performance evaluation against conventional machine learning methods

2. Related Work

Machine learning techniques have been widely applied to cardiovascular disease prediction using clinical datasets. Early research primarily focused on supervised learning models such as logistic regression, decision trees, and support vector machines for predicting cardiovascular risk [2], [5]. These approaches demonstrated promising results and served as the foundation for automated diagnostic systems.

Several studies have proposed machine learning frameworks integrated with clinical decision-support systems. For example, Fitriyani et al. developed the HDPM model to assist physicians in cardiovascular diagnosis using patient data and machine learning techniques [2]. Similar studies explored various classification algorithms for analyzing clinical attributes associated with heart disease risk [4].

To improve prediction accuracy, ensemble learning techniques have been introduced to combine multiple classifiers. Ensemble approaches such as stacking and blending have shown improved robustness and generalization performance in cardiovascular disease prediction tasks [8], [9]. Hybrid machine learning models have also been proposed to capture complex relationships between clinical features and disease outcomes [7].

Deep learning models have further expanded the capabilities of medical prediction systems. Neural network-based approaches have been applied to medical imaging and clinical datasets to identify disease patterns and improve prediction accuracy [18], [21]. More recently, transformer architectures have been introduced for healthcare analytics due to their ability to capture feature dependencies through attention mechanisms [3].

Another important research direction involves improving model interpretability. Explainable AI techniques provide insights into prediction outcomes and highlight key clinical factors influencing disease risk. Several studies have incorporated explainability mechanisms to improve transparency in cardiovascular prediction systems [11], [12].

Despite these advancements, many existing models either lack interpretability or fail to capture complex feature interactions in structured clinical datasets. Therefore, combining transformer-based feature representation with ensemble learning provides a promising approach for improving both predictive accuracy and model transparency.

3. Proposed Methodology

This study proposes a Transformer-Enhanced Ensemble Learning framework for interpretable cardiovascular disease prediction. The framework integrates transformer-based feature representation with ensemble classification and explainable artificial intelligence techniques. The overall architecture consists of four stages: data preprocessing, transformer-based feature representation, ensemble prediction, and model interpretation.

Figure 1 shows the architecture of the proposed Transformer-Enhanced Ensemble Learning framework for cardiovascular disease prediction. Clinical data are first preprocessed, then passed through a transformer encoder to extract feature representations. These features are used by multiple classifiers in an ensemble model to generate the final disease prediction, while an explainable AI module provides interpretation of the prediction results.

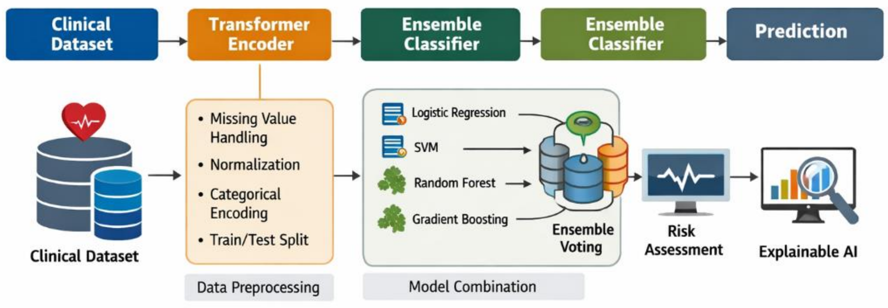

Fig.1. Proposed Transformer-Enhanced Ensemble Learning Framework for Cardiovascular Disease Prediction

3.1 Data Preprocessing

The cardiovascular disease dataset consists of structured clinical attributes including age, sex, chest pain type, cholesterol level, resting blood pressure, electrocardiographic results, and maximum heart rate. Before model training, several preprocessing steps are applied to improve data quality and model performance.

These steps include:

- Handling missing values
- Removing inconsistent records
- Normalizing numerical attributes
- Encoding categorical variables

Feature normalization ensures that attributes with different scales do not bias the learning process. After preprocessing, the dataset is divided into training and testing subsets for model evaluation.

3.2 Transformer-Based Feature Representation

To capture complex relationships between clinical attributes, the proposed framework utilizes a transformer encoder. Transformers employ an attention mechanism that allows the model to assign importance weights to different features during prediction.

The attention mechanism is defined as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \qquad □□□$$

Where:

- Q represents the query matrix
- K represents the key matrix
- V represents the value matrix
- K^T denotes the transpose of the key matrix
- d_k represents the dimension of the key vectors

The term QK^T computes the similarity between the query and key vectors. The scaling factor $\sqrt{d_k}$ prevents extremely large values that may destabilize the softmax function during training. The softmax function converts similarity scores into normalized attention weights, which determine the importance of each feature. These weights are then multiplied with the value matrix V to generate the final attention-based feature representation.

Each patient record can be represented as a feature vector:

$$F = f(x_1, x_2, x_3, \ldots, x_n) □2□$$

where $x_1, x_2, \ldots, x_n$ represent the clinical attributes and FF denotes the encoded feature representation generated by the transformer encoder.

Figure 2 illustrates the Transformer Feature Representation Module used to extract meaningful features from the clinical dataset. First, the input clinical attributes are converted into numerical vectors through feature embedding. These embedded features are then processed using a self-attention mechanism, where query (Q), key (K), and value (V) matrices compute relationships between different features. Finally, the transformer generates an encoded feature representation, which captures dependencies among clinical attributes and is used as input for the ensemble classification model.

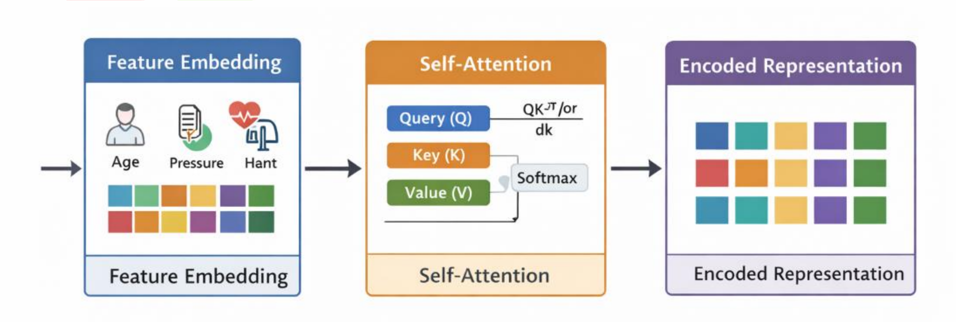

Fig.2. Transformer Feature Representation Module

3.3 Ensemble Classification

To improve prediction accuracy and model robustness, the encoded features generated by the transformer module are used to train multiple machine learning classifiers. The proposed framework employs an ensemble learning strategy, where several classifiers contribute to the final prediction.

The ensemble includes models such as:

- Logistic Regression
- Support Vector Machine

- Random Forest
- Gradient Boosting

The final prediction is obtained by aggregating outputs from individual classifiers. The ensemble prediction can be expressed as:

$$P_{ensemble} = \frac{1}{N}\sum_{i=1}^{N} P_i \quad (3)$$

where

- P_i represents the prediction generated by the i^thith classifier
- N represents the number of classifiers in the ensemble.

This aggregation reduces prediction variance and improves generalization performance.

Figure 3 illustrates the Ensemble Prediction Framework used in the proposed model. Multiple classifiers—Logistic Regression, Support Vector Machine (SVM), Random Forest, and Gradient Boosting—are trained using the extracted feature representations. Each classifier produces a prediction, and these outputs are combined in a voting layer. The voting mechanism aggregates the individual predictions to generate the final cardiovascular disease prediction, indicating either the presence or absence of disease.

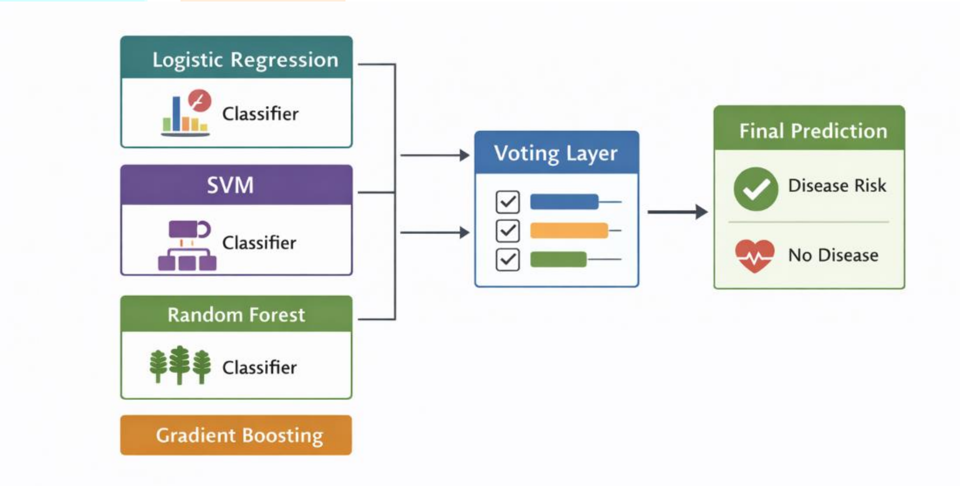

Fig.3. Ensemble Prediction Framework

3.4 Explainable AI Analysis

Interpretability is an important requirement for healthcare prediction systems. Therefore, the proposed framework integrates Explainable Artificial Intelligence (XAI) techniques to analyze model predictions.

The SHAP (Shapley Additive Explanations) method is used to compute feature importance scores and identify the clinical attributes that contribute most significantly to the prediction outcome. This enables healthcare practitioners to understand how patient characteristics influence the model's decision.

Figure 4 presents the SHAP Feature Importance Visualization, which explains how different clinical features influence the prediction of cardiovascular disease. The plot ranks features based on their contribution to the model output. Each point represents the impact of a feature value on the prediction, where red indicates high feature values and blue indicates low feature values. Features such as chest pain type (cp), thalassemia (thal), number of major vessels (ca), ST depression (oldpeak), and cholesterol (chol) show strong influence on the prediction results. This visualization helps interpret the model by identifying the most significant risk factors contributing to cardiovascular disease prediction.

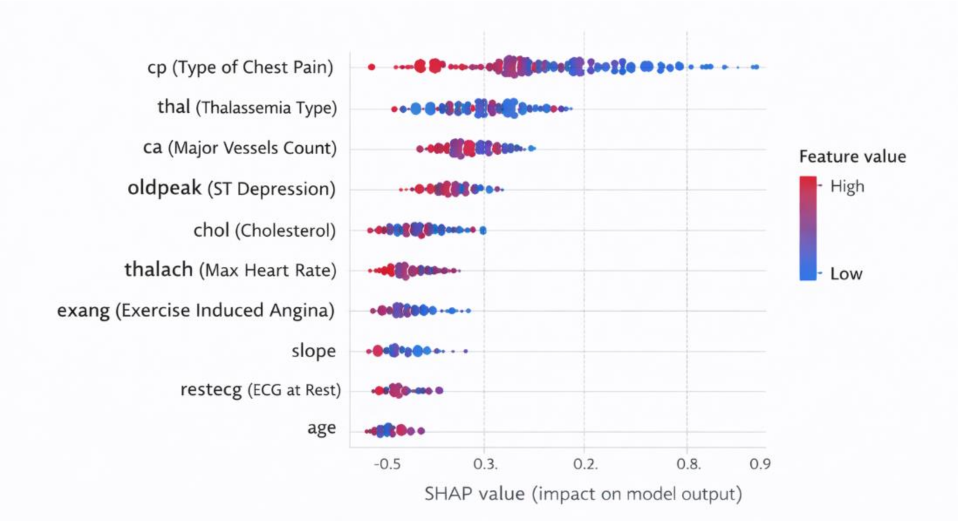

Fig.4. SHAP Feature Importance Visualization

The Explainability module enhances transparency and increases trust in the AI-based prediction system by providing insights into the factors associated with cardiovascular disease risk.

4. Experimental Setup and Performance Evaluation

This section describes the experimental configuration, dataset preparation, and evaluation criteria used to assess the proposed Transformer-Enhanced Ensemble Learning framework for cardiovascular disease prediction.

4.1 Experimental Environment

The proposed framework was implemented using a Python-based machine learning environment. Python libraries were used for data preprocessing, model training, and evaluation. Scikit-learn was used to implement machine learning models, while TensorFlow/PyTorch supported the transformer-based feature representation. The SHAP library was employed to provide explainable AI analysis for interpreting model predictions.

Table 1 presents the experimental environment used for implementing and evaluating the proposed cardiovascular disease prediction framework. The implementation was carried out using Python with machine learning, deep learning, and data processing libraries.

Table 1 Experimental Environment

Component	Description
Programming Language	Python
Machine Learning Framework	Scikit-learn
Deep Learning Framework	TensorFlow / PyTorch
Explainability Framework	SHAP
Data Processing	Pandas, NumPy

4.2 Dataset Preparation

The cardiovascular disease dataset consists of clinical attributes such as age, cholesterol level, blood pressure, chest pain type, electrocardiographic results, and maximum heart rate. The target variable indicates the presence or absence of cardiovascular disease.

Before training, the dataset was preprocessed by handling missing values, normalizing numerical attributes, and encoding categorical features. The data were divided into training and testing sets using an 80–20 split, where 80% of the data were used for training and 20% for model evaluation. To improve reliability, k-fold cross-validation was applied during model training.

4.3 Model Training

The proposed model consists of two main stages. First, a transformer encoder extracts contextual feature representations from the clinical dataset using an attention mechanism. These encoded features are then used to train multiple machine learning classifiers.

An ensemble strategy combines the predictions from several base learners, including logistic regression, support vector machines, random forests, and gradient boosting models. The final prediction is obtained using majority voting or probability averaging.

4.4 Performance Evaluation Metrics

To evaluate the effectiveness of the proposed cardiovascular disease prediction framework, several widely used classification performance metrics were employed. These metrics provide a comprehensive assessment of the predictive capability of the model in terms of correctness, reliability, and sensitivity. The evaluation metrics used in this study include accuracy, precision, recall, and F1-score.

Accuracy measures the overall proportion of correctly classified instances among all predictions made by the model. It is defined as

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (4)$$

where TP represents true positive cases, TN represents true negative cases, FP represents false positive cases, and FN represents false negative cases.

Precision evaluates the proportion of correctly predicted positive instances among all instances predicted as positive. It is expressed as

$$Precision = \frac{TP}{TP+FP} \qquad (5)$$

This metric indicates how reliable the model's positive predictions are.

Recall, also known as sensitivity, measures the ability of the model to correctly identify actual positive cases. It is defined as

$$Recall = \frac{TP}{TP+FN} \qquad (6)$$

High recall indicates that the model can effectively detect patients with cardiovascular disease.

The F1-score combines precision and recall into a single metric by computing their harmonic mean. It provides a balanced evaluation of classification performance, particularly when dealing with imbalanced datasets.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \qquad (7)$$

These evaluation metrics provide a comprehensive measure of the classification performance of the proposed model and enable fair comparison with baseline machine learning approaches.

Table 2 presents the performance evaluation metrics used to assess the effectiveness of the proposed cardiovascular disease prediction model. These metrics provide a comprehensive assessment of the model’s predictive accuracy, reliability, and classification capability.

Table 2 Performance Evaluation Metrics

Metric	Description
Accuracy	Overall prediction correctness
Precision	Correct positive predictions among predicted positives
Recall	Ability to detect actual positive cases
F1-score	Balance between precision and recall
AUC	Model capability to distinguish between classes

5. Results and Discussion

The proposed Transformer-Enhanced Ensemble Learning model was compared with conventional machine learning algorithms used for cardiovascular disease prediction. As shown in Table 3, traditional models such as Logistic Regression, Decision Tree, and SVM achieved accuracies of 83.4%, 85.2%, and 86.7%, respectively. Ensemble models like Random Forest and Gradient Boosting performed better with 89.3% and 90.1% accuracy.

However, the proposed model achieved the highest accuracy of 93.6%, outperforming all baseline methods. This improvement is due to the combination of transformer-based feature extraction and ensemble learning, which enhances feature representation and prediction robustness. Additionally, the model achieved the highest AUC value, indicating better capability in distinguishing between diseased and non-diseased cases.

Table 3 Performance Comparison

Model	Accuracy
Logistic Regression	83.4
Decision Tree	85.2
SVM	86.7
Random Forest	89.3
Gradient Boosting	90.1
Proposed Model	93.6

Figure 5 shows the Receiver Operating Characteristic (ROC) curve comparison of different prediction models. The ROC curve plots the True Positive Rate (Sensitivity) against the False Positive Rate (1 – Specificity) to evaluate the classification performance.

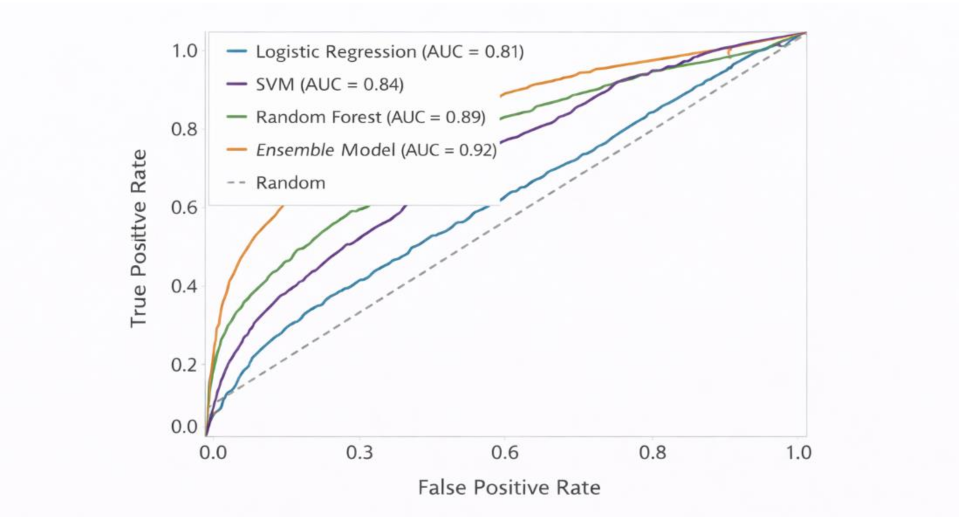

Fig.5 ROC Curve Comparison

The curves represent models such as Logistic Regression, Support Vector Machine (SVM), Random Forest, and the proposed Ensemble Model. The Area Under the Curve (AUC) indicates the model's ability to distinguish between diseased and non-diseased cases. The proposed ensemble model achieves the highest AUC, demonstrating better prediction performance compared to the individual baseline models.

The results demonstrate that transformer-based feature representation combined with ensemble learning significantly improves classification performance.

Explainable AI analysis identified key clinical factors influencing cardiovascular risk, including age, cholesterol level, chest pain type, and maximum heart rate.

Figure 6 shows the SHAP Feature Importance Plot, which ranks clinical features based on their contribution to the cardiovascular disease prediction model.

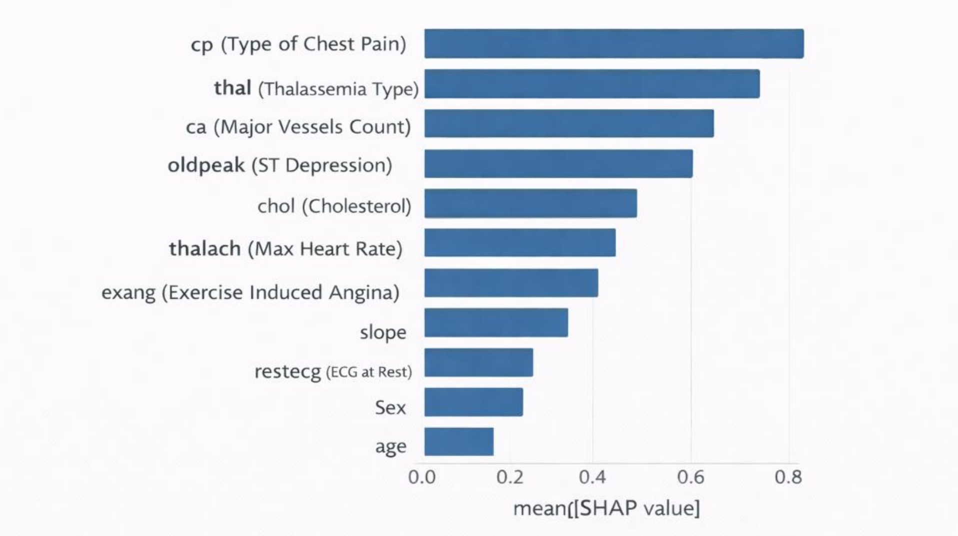

Fig.6 SHAP Feature Importance Plot

The horizontal bars represent the mean absolute SHAP values, indicating the overall importance of each feature. Features such as chest pain type (cp), thalassemia (thal), number of major vessels (ca), and ST depression (oldpeak) have the highest impact on the model's predictions. This analysis helps identify the most significant clinical risk factors influencing cardiovascular disease detection and improves the interpretability of the proposed model.

6. CONCLUSION

Although the proposed Transformer-Enhanced Ensemble Learning framework demonstrates promising performance for cardiovascular disease prediction, several research directions remain for further improvement and practical deployment.

First, future studies can evaluate the proposed framework using larger and more diverse clinical datasets collected from multiple healthcare institutions. Incorporating heterogeneous patient populations would improve the generalization capability of the model and enable more reliable real-world clinical applications.

Second, the current study focuses primarily on structured clinical attributes. Future work may integrate multimodal medical data, including electrocardiogram (ECG) signals, medical imaging, wearable sensor data, and electronic health records. Combining multiple data modalities can enhance predictive accuracy and provide a more comprehensive representation of patient health conditions.

Another promising direction involves the use of advanced transformer architectures and lightweight attention mechanisms to improve computational efficiency. Optimizing transformer models for structured healthcare datasets may reduce training complexity while maintaining strong predictive performance.

Additionally, federated learning approaches could be explored to enable collaborative model training across multiple hospitals without sharing sensitive patient data. Such privacy-preserving learning frameworks would address data security concerns and facilitate large-scale healthcare analytics.

Finally, the proposed framework could be extended into real-time clinical decision-support systems integrated with hospital information systems. This would allow healthcare professionals to utilize AI-assisted prediction tools during routine clinical assessments, enabling earlier detection of cardiovascular risk and supporting preventive healthcare strategies.

7. ACKNOWLEDGMENT

The authors would like to acknowledge the support and resources provided by their respective institutions that facilitated the completion of this research work. The authors also express their gratitude to the researchers and organizations that made publicly available cardiovascular disease datasets and machine learning libraries, which were instrumental in conducting the experimental evaluation presented in this study.

REFERENCES

[1] S. Santhosh, K. Chadaga, R. V. Arjunan, and S. D'Souza, "Cardiac Clarity: Harnessing Machine Learning to Accurately Predict Heart Disease," IEEE Access, vol. 13, pp. 97529–97544, 2025, doi: 10.1109/ACCESS.2025.3573760.

[2] N. L. Fitriyani, M. Syafrudin, G. Alfian, and J. Rhee, "HDPM: An Effective Heart Disease Prediction Model for Clinical Decision Support Systems," IEEE Access, vol. 8, pp. 133034–133050, 2020, doi: 10.1109/ACCESS.2020.3010511.

[3] F. Mahmud et al., "HybridTabNet-QC: A Transformer-Based Clinical Feature Fusion Framework to Predict Heart Disease Risk," IEEE Open Journal of the Computer Society, vol. 7, pp. 1–13, 2026, doi: 10.1109/OJCS.2025.3637308.

[4] V. R. Parihar, R. S. Nage, and A. S. Dahane, "Image analysis and image mining techniques: A review," Journal of Image Processing and Artificial Intelligence, vol. 3, no. 2, pp. 1–18, 2017.

[5] T. Patil, N. Honmane, A. Mahajan, U. Khune, and A. Bangar, "Working of a blood donation application using GPS," International Journal of Innovative Research in Computer and Communication Engineering, vol. 11, no. 3, pp. 1–7, Mar. 2023.

[6] A. Pisal, N. Honmane, T. Pawar, V. D. Pawar, and V. V. Pawar, "A survey paper on using facial recognition to propose music," International Journal of Innovative Research in Computer and Communication Engineering, vol. 11, no. 3, pp. 1–8, Mar. 2023.

[7] T. Patil, N. Honmane, A. Mahajan, U. Khune, and A. Bangar, "BloodFinder application using GPS and Kotlin," International Journal of Advanced Research in Science, Communication and Technology (IJARSCT), vol. 3, no. 15, pp. 1–6, May 2023.

[8] N. Honmane, P. Kadbane, S. Hande, S. Goyal, and A. Jadhav, "Smart attendance: Automated facial recognition system," International Journal for Research Trends and Innovation (IJRTI), vol. 10, no. 3, Mar. 2025.

[9] V. R. Parihar, "Image segmentation based on graph theory and threshold," in Image Segmentation: A Guide to Image Mining, ICSES Book Series, vol. 4, no. 4, pp. 61–82, Dec. 2018.

[10] N. Mohan, V. Jain, and G. Agrawal, "Heart Disease Prediction with Supervised Machine Learning Algorithms," in Proc. 5th Int. Conf. Information Systems and Computer Networks (ISCON), Mathura, India, 2021, pp. 1–3, doi: 10.1109/ISCON52037.2021.9702314.

[11] H. F. El-Sofany, "Predicting Heart Diseases Using Machine Learning and Various Data Classification Techniques," IEEE Access, vol. 12, pp. 106146–106160, 2024, doi: 10.1109/ACCESS.2024.3437181.

[12] J. Talaviya, D. Trivedi, R. Ramani, A. Diwan, and R. Mahadeva, "A Data-Driven Approach to Cardiac Health: Machine Learning to Predict Heart Diseases," in Proc. IEEE Technology and Engineering Management Conf. Asia Pacific (TEMSCON-ASPAC), Bengaluru, India, 2023, pp. 1–7.

[13] R. S. Bhaduaria, I. Javid, and A. Khara, "Advanced Heart Attack Risk Prediction Using Stacked Hybrid Machine Learning," Journal of Mobile Multimedia, vol. 21, no. 3–4, pp. 393–406, 2025, doi: 10.13052/jmm1550-4646.21343.

[14] H. Khan, N. Javaid, T. Bashir, M. Akbar, N. Alrajeh, and S. Aslam, "Heart Disease Prediction Using Novel Ensemble and Blending-Based Cardiovascular Disease Detection Networks: EnsCVDD-Net and BlCVDD-Net," IEEE Access, vol. 12, pp. 109230–109254, 2024, doi: 10.1109/ACCESS.2024.342124.

[15] M. Naz et al., "Meta-Ensemble Learning to Predict Heart Diseases: A Stacking-Based Approach with Explainable AI," IEEE Access, vol. 13, pp. 137271–137290, 2025, doi: 10.1109/ACCESS.2025.3588683.

[16] B. Gayathri, V. N. Kumar, R. Tiwari, and V. Indumathi, "An Innovative Prediction Model for Heart Disease Detection Using Artificial Intelligence and Machine Learning Methods," in Proc. 1st Int. Conf. Sustainable Computing and Integrated Communication in Changing Landscape of AI (ICSCAI), Greater Noida, India, 2024.

[17] A. Gadde and S. Chintala, "Comprehensive Study on Heart Disease Prediction and Risk Stratification Using Explainable Artificial Intelligence," in Proc. Int. Conf. Intelligent Systems and Computational Networks (ICISCN), Bidar, India, 2025.

[18] V. R. Parihar, "Automatic wireless health monitoring system," Reinvention International: An International Journal of Thesis Projects and Dissertation, vol. 1, no. 1, pp. 84–95, 2018.

[19] R. Bhuvaneswari, P. Kumar, and S. Kaviya, "Explainable AI-Driven Heart Disease Prediction," in Proc. Int. Conf. Visual Analytics and Data Visualization (ICVADV), Tirunelveli, India, 2025, pp. 959–964.

[20] J. P. S., T. Rao, and A. Kodipalli, "Heart Disease Prediction with AHP and Fuzzy TOPSIS Methods," in Proc. Int. Conf. Cognitive Computing in Engineering, Communications, Sciences and Biomedical Health Informatics (IC3ECSBHI), Greater Noida, India, 2025.

[21] A. Srivastav, A. Pandey, and S. Chaturvedi, "Heart Disease Prediction Using Artificial Intelligence," in Proc. 12th Int. Conf. Computing for Sustainable Global Development (INDIACom), Delhi, India, 2025.

[22] Y. Narayan, N. Kumar, V. Kumar, Manya, and D. Karan, "Heart Disease Identification Using Machine Learning Algorithm," in Proc. IEEE Int. Conf. Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI), Gwalior, India, 2025.

[23] V. R. Parihar, "Image segmentation based on graph theory and threshold," in Image Segmentation: A Guide to Image Mining, vol. 4, no. 4, International Computer Science and Engineering Society, Dec. 2018.

[24] S. P. Alampally et al., "Optimizing Cardiovascular Disease Prediction Using Machine Learning Models on Heart Disease Dataset," in Proc. 7th Int. Conf. Information Systems and Computer Networks (ISCON), Mathura, India, 2025.

[25] A. Kanmani, G. S. Gopika, and M. Sreekrishna, "Streamlit-Based Multi-Disease Prediction System for Early Detection of Kidney Disease, Breast Cancer and Heart Disease," in Proc. IEEE Int. Conf. Contemporary Computing and Communications (InC4), Bangalore, India, 2025.

[26] K. Pati, S. Banik, M. S. Sk, A. Srimani, S. Ghosh, and S. Dasgupta, "An AI-Driven GPT Model for Automated Heart Disease Detection in ECG Dataset," in Proc. AI-Driven Smart Healthcare for Society 5.0, Kolkata, India, 2025.

[27] N. S. Magar and P. Kaur, "Survey of Heart Disease Prediction Models Using Machine Learning and Deep Learning Techniques," in Proc. Int. Conf. Engineering, Technology and Management (ICETM), Oakdale, NY, USA, 2025

[28] M. F. Ahmed et al., "Heart Disease Prediction in App-Based Systems: Optimized Machine Learning Model with Reduced Features and Explainable AI," in Proc. Int. Conf. Quantum Photonics, Artificial Intelligence, and Networking (QPAIN), Rangpur, Bangladesh, 2025.

[29] T. P. and A. R., "DL-Based Heart Disease Prediction System Using CT Scan Images," in Proc. IEEE Int. Conf. Information Technology, Electronics and Intelligent Communication Systems (ICITEICS), Bangalore, India, 2024.

[30] V. V. Bhakare, V. R. Parihar, A. D. Wankhade, A. R. Shah, and S. R. Vaidya, "Kidney Stone Prediction using Neural Network," in Proc. One Day National Level Inter-Disciplinary E-Conference on Contemporary Dynamics in Humanities, Science & Technology, Apr. 2024.

[31] R. S. Nage, "A Review and Comparative Analysis on Image Mining Technique," Image Segmentation, vol. 4, no. 4, pp. 51–60, 2018.

[32] V. R. Parihar, A. Y. Tonge, and P. D. Ganorkar, "Heartbeat and temperature monitoring system for remote patients using Arduino," International Journal of Advanced Engineering Research and Science (IJAERS), vol. 4, no. 5, pp. 55–58, May 2017.

[33] A. P. Agarkar and V. R. Parihar, “Plant leaf disease detection using machine learning,” WSEAS Transactions on Computers, vol. 24, pp. 136–141, 2025.