I.J. Intelligent Systems and Applications, **2025, 5, 49-67**
Published Online on October 8, 2025 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijisa.2025.05.04

Advanced Heart Attack Prediction Using a Stacked Ensemble Machine Learning Model and Diverse Data Integration

Sultanul Arifeen Hamim
Department of Computer Science, American International University-Bangladesh, Dhaka, 1219, Bangladesh
E-mail: arifeenhamim@gmail.com
ORCID iD: https://orcid.org/0009-0008-4402-0739

Rakin S. Aftab*
Department of Computer Science, American International University-Bangladesh, Dhaka, 1219, Bangladesh
E-mail: rakinsadaftab@gmail.com
ORCID iD: https://orcid.org/0000-0001-7164-5517
*Corresponding author

M. Ahmed
Department of Computer Science, American International University-Bangladesh, Dhaka, 1219, Bangladesh
E-mail: marufmir420@gmail.com
ORCID iD: https://orcid.org/0009-0007-7181-4393

Farzana Faiza
Department of Computer Science, American International University-Bangladesh, Dhaka, 1219, Bangladesh
E-mail: farzanafaiza601@gmail.com
ORCID iD: https://orcid.org/0009-0008-5831-1000

M. F. Mridha
Department of Computer Science, American International University-Bangladesh, Dhaka, 1219, Bangladesh
E-mail: firoz.mridha@aiub.edu
ORCID iD: https://orcid.org/0000-0001-5738-1631

Received: 03 June 2025; Revised: 15 July 2025; Accepted: 21 August 2025; Published: 08 October 2025

Abstract: Heart attacks continue to be one of the primary causes of death globally, highlighting the critical need for advanced predictive models to improve early diagnosis and timely intervention. This study presents a comprehensive machine learning (ML) approach to heart attack prediction, integrating multiple datasets from diverse sources to construct a robust and accurate predictive model. The research employs a stacking ensemble model, which combines the strengths of individual ML algorithms to improve overall performance. Extensive data preprocessing steps were carefully undertaken to preserve the dataset's integrity and maintain its quality. The results demonstrate a superior accuracy of 97.48%, significantly outperforming state-of-the-art approaches. The high level of accuracy indicates the model’s potential effectiveness in the clinical setting for early detection of heart attack and prevention. However, the proposed model is influenced by the quality and diversity of the integrated datasets, which could affect its generalizability across broader populations. Challenges encountered during the model's development include optimizing hyperparameters for multiple classifiers, ensuring data preprocessing consistency, and balancing computational efficiency with model interpretability. The results underscore the pivotal contribution of advanced ML approaches in revolutionizing the management of cardiovascular attack. By addressing the complexities and variabilities inherent in heart attack prediction, the work provides a pathway towards more effective and personalized cardiovascular disease management strategies, demonstrating the transformative potential of ML in healthcare.

Index Terms: Myocardial Healthcare, Heart Attack Prediction, Machine Learning, Data Integration, Hybrid Stacked Ensemble Model.

1. Introduction

Recent advances in technology had a transformative effect on the provision of healthcare, specifically cardiovascular health. Cardiac arrests, a prominent contributor to mortality worldwide [1], require inventive and proactive approaches to prevention and treatment. The heart, an essential organ, facilitates blood circulation, delivers oxygen, and eliminates metabolic wastes [2]. A heart attack, medically termed myocardial infarction (MI), occurs when blood supply to a section of the heart muscle becomes blocked, typically caused by a blood clot. This obstruction leads to injury to the cardiac tissue [3].

Hypertension, high cholesterol levels, and tobacco use are significant risk factors that contribute to the formation of blockages [4]. Given that a substantial proportion of the world's population displays these risk factors, the frequency of deaths caused by heart attacks remains elevated [5]. Timely detection of probable MI is crucial, requiring monitoring of hypertension, lipid profiles, body mass index, and glycemic levels [6]. Nevertheless, the substantial surveillance and documentation necessary for every patient pose significant obstacles [7,8].

Early prediction of heart attacks is important because heart attacks, is still the leading cause of mortality worldwide. Early prediction and timely intervention reduce the chances of mortality and morbidity related to heart attack [1]. Prediction of heart attack before the event is beneficial not only for patients, who are given preventive care, advice for behavioral changes, and medications to mitigate risk factors and maintain good health, but also for healthcare providers, as accurate predictive models enhance service delivery in diagnosis, patient management, and resource utilization [2]. Further, the overall healthcare system reaps benefits from lower emergency admissions and, importantly, treatment costs that will be used for the management of heart attack at a later, more severe stage. Predictive models also empower patients by providing actionable health insights, leading to proactive patient engagement [4]. Advanced ML techniques used for the prediction of heart attacks can identify hidden patterns and risk factors present in huge, complex datasets, which eventually lead to the generation of more precise and reliable predictions [9]. This not only brings success in individual patient outcomes but also leads to success in public health and prevention. In general, the potential impact of successfully predicting the onset of a heart attack is massive, bringing about better survival rates for patients, improvements in the quality of life of people, and the effectiveness of healthcare systems.

In recent times, advancements in ML and artificial intelligence (AI) have brought significant changes to numerous domains, particularly healthcare. As a branch of AI, ML uses computational methods to process data, enabling systems to learn patterns and generate predictions. This innovation has demonstrated remarkable capabilities in handling vast and intricate datasets within healthcare, uncovering trends and providing accurate outcome predictions. Specifically, in the domain of heart attack prediction, ML models have demonstrated superior performance over traditional statistical methods by leveraging the ability to handle vast and multifaceted datasets [10].

The previous research advanced numerous ML techniques in the prediction of heart attacks, and some methods applied are, for example, logistic regression (LR), support vector machines (SVM), random forests (RF), or even neural methods. In a review, a process was revealed that revealed the capacity of the models to increase the predicted accuracy, hence improving the diagnosis for earlier intervention. For instance, ensemble techniques—that is, techniques to combine models to boost overall performance—have become new directions in the area of previous research. Such methods as boosting and stacking are employed to improve the accuracy of the predictions by reducing the drawbacks of single models.

Various research investigates a variety of advanced algorithms to improve the forecasting and control of MI. It deals with these issues by looking at various data science (DS) methods, such as well-known models like LR, SVM [11,12] and naive bayes (NB), as well as more advanced ensemble methods such as RF, extreme gradient boosting (XGBoost), and gradient boosting classifier. In addition, some assess the prediction capabilities of K-nearest neighbors (KNN), multi-layer perceptron (MLP) classifier, light gradient-boosting machine (LGBM) classifier, and gaussian mixture [13, 14]. These algorithms were selected for optimal characteristics in handling complex and diverse heart health data. The following classifiers are available: adaptive boosting (AdaBoost) classifier, perceptron, nearest centroid, bernoulli naive bayes, stochastic gradient descent (SGD) classifier, calibrated classifier cross validation, linear support vector classification (LinearSVC), linear discriminant analysis, logistic regression cross validation (LR-CV), and quadratic discriminant analysis.

This paper presents a hybrid ML approach to predicting heart attacks. The aim is to combine several datasets from different sources in order to build one large, comprehensive predictive model using a stacking ensemble approach. There is extensive research toward the prediction of heart attacks using a variety of techniques under the ML umbrella, namely: LR, KNN, RF, SVM, and XGBoost. Greatly improved prediction accuracy has been achieved, and the study of many other factors has now been made possible with the help of these models to provide early diagnosis and intervention [11-14]. It is important for the data to undergo strict preprocessing in terms of imputations, attribute normalization, and data selection to improve the quality of the dataset. Ensemble techniques, with a particular emphasis on stacking, are beneficial because combining the strengths of two or more models results in an optimal combined model. The extraordinary accuracy of this model outperforms previous methods and proves to be very useful for the early detection of a heart attack. It underlines the potential to make a breakthrough in applying advanced ML strategies to tackle cardiovascular attacks, henceforth enabling feasible and tailored interventions in the clinical field. The contributions introduced by this research are as follows:

- Integration of Diverse Datasets: This research combines multiple datasets from various sources—Kaggle, Mendeley Data, and the UCI ML Repository—to improve the model's reliability and applicability. This approach overcomes the constraints associated with relying solely on single-source datasets often seen in similar studies.
- Stacking Ensemble Approach: The research implements a sophisticated stacking ensemble model that combines the strengths of diverse individual ML algorithms (LR, KNN, RF, SVM, and XGBoost) and employs a LR meta-classifier to achieve superior performance.
- Superior Predictive Metrics: The proposed model attains a performance level with an accuracy of 97.48%, a precision of 97.64%, a recall of 98.51%, and an F1-score of 97.78%. These evaluation metrics underscore its superiority over standalone classifiers and contemporary techniques, emphasizing its potential for early detection of heart attacks.
- Clinically Validated Feature Importance: Tree-based models are utilized to identify and validate clinically significant features such as ST slope, chest pain type, and maximum heart rate. These findings are corroborated with clinical knowledge, ensuring the interpretability and reliability of the model.
- Comprehensive Model Evaluation: The effectiveness of the proposed model is thoroughly validated through an evaluation process that employs both standard and advanced performance metrics such as accuracy, precision, recall, and F1-score. Additionally, visual representations like confusion matrices and ROC curves are utilized to assess and confirm the model's reliability.
- Practical Clinical Applicability: The study underscores the transformative potential of advanced ML techniques in cardiovascular disease management. The proposed model demonstrates its capacity for real-world clinical application, supporting personalized risk assessment and early intervention strategies.

The structure of this paper is outlined as follows: Section 2 reviews the relevant literature in detail, and Section 3 explains the methodology utilized in this research. Section 4 describes the proposed model, while Section 5 highlights the results and analysis. A comprehensive discussion of the findings is provided in Section 6. Lastly, Section 7 concludes the paper with a summary of key insights and suggestions for future research directions.

2. Literature Review

Healthcare professionals have relied on a variety of traditional methods for heart attack prediction, utilizing clinical risk factors such as age, gender, family history, and medical history. Diagnostic tools like echocardiography, which evaluates the heart's structure and function, and electrocardiograms (ECGs), which detect heart abnormalities, play crucial roles in assessing heart health [15,16]. Specifically, ECG assists in managing treatment plans for patients with congestive heart failure, while cardiac catheterization is used to diagnose coronary artery disease (CAD) and other cardiovascular conditions [17].

The demand for more sophisticated predictive models has grown, with machine learning (ML) emerging as a transformative approach to improving accuracy and efficiency in heart attack prediction [18]. ML algorithms enable systems to learn from data, uncover patterns, and make informed decisions, offering the ability to process and analyze vast datasets [19]. This capability positions ML as a pivotal tool in advancing precise and reliable heart attack prediction, paving the way for innovative solutions in cardiovascular care.

Recent research on heart attack prediction has demonstrated different approaches to the process. For example, [20] conducts a systematic literature review and bibliometric analysis of studies on heart attack prediction using ML. The review identifies key trends and methodologies in the field, emphasizing the impact of various ML algorithms on prediction accuracy. The study highlights the need for comprehensive research to address existing gaps and improve early warning systems for heart attack. Heart attack remains a leading cause of mortality worldwide, necessitating advanced diagnostic tools for early detection and intervention. Traditional diagnostic methods are often invasive, time-consuming, and prone to errors, highlighting the need for innovative solutions. [21] provides a comprehensive review of the role of ML in heart attack prediction, emphasizing the potential of ML to enhance early attack detection and reduce healthcare costs. Critical aspects discussed include feature selection, evaluation metrics, and recent trends in ML applications for heart attack prediction. [22] focuses on boosting ensemble techniques for heart attack prediction, highlighting the effectiveness of methods such as AdaBoost, categorical boosting (CatBoost), gradient boosting, LGBM, and RF. The study utilizes a substantial dataset from the UCI ML library, demonstrating that AdaBoost outperforms other algorithms with an accuracy of 95%. The optimally configured and improved long short-term memory (OCI-LSTM) model, introduced in [23], is an enhanced long-short-term memory model designed for heart attack prediction. It utilizes the salp swarm algorithm (SSA) for feature selection and the genetic algorithm (GA) for network configuration optimization. Achieving an impressive accuracy of 97.11%, the OCI-LSTM model outperforms other models, such as deep neural networks and deep belief networks.

Another study, presented in [24], introduces the ensemble heuristic-metaheuristic feature fusion learning (EHMFFL) algorithm to improve heart attack diagnosis using tabular data. This approach incorporates heuristic knowledge, such as the pearson correlation coefficient (PCC), with the metaheuristic-driven grey wolf optimizer (GWO) to select optimal feature subsets for various ML models. The ensemble model includes multiple base learners, such as

SVM, KNN, LR, RF, NB, decision trees (DT), and XGBoost techniques. The algorithm demonstrates high accuracy rates of 91.8% and 88.9% on the Cleveland and Statlog datasets, respectively, surpassing state-of-the-art techniques in heart attack diagnosis.

AI and ML have shown significant progress in cardiology predictive analytics. Different research papers employed various AI techniques, such as LR, KNN, RF, and linear support vector machines (LSVC), to create an AI-driven tool for predicting the likelihood of heart attacks. The study in [25] found LR to be the most precise for initial screening, identifying chest pain as the most crucial predictor and RF as a dependable model. Another research paper, [9], analyzed the impact of anthropometric features on the risk of heart attacks using ML techniques like NB, SVM, DT, RF, KNN, and LR. A comparison of seven ML models for classifying cardiac disease using datasets from Cleveland, Statlog, and Hungary showed that RF outperformed the others, achieving 94.96% accuracy after adjusting hyperparameters and performing cross-validation [26]. Additionally, despite obstacles such as data imbalance and the need for efficient preprocessing techniques, studies on predicting diabetes using ML algorithms incorporating physiological and behavioral aspects have shown the effectiveness of ensemble methods like RF [27]. These studies highlight how ML and AI have the potential to revolutionize preventative medicine by enabling early intervention based on non-invasive data.

Similarly, [28] compared several ML techniques for heart attack prediction, including SVM, RF, KNN, DT, NB, and LR. The study utilized the Cleveland dataset and focused on performance metrics such as accuracy, precision, recall, and F1 score. The SVM model yielded the highest accuracy of 89.34%, demonstrating its robustness in classification tasks. Researchers in [29] also employed these ML algorithms, such as LR, RF, and KNN, on data from the Harvard Dataverse Repository, incorporating various data types, including clinical and behavioral factors.

Numerous ML and optimization techniques have been studied to predict cardiac disease, with notable improvements in interpretability and accuracy. A study that employed the Cleveland dataset and presented a transformer model based on self-attention achieved an accuracy of 96.51% and offered interpretability via feature contribution analysis [30]. Using an enhanced CatBoost model with lasso regression, the CVD-OCSCatBoost framework increased prediction accuracy and convergence speed to 73.67% [31]. Across five datasets, a heart attack risk prediction model that selected features using the whale optimization algorithm (WOA) showed notable increases in accuracy, precision, recall, and F1 score [32]. A unique transfer learning-based feature engineering technique, validated through 10-fold cross-validation, attained an outstanding 97.5% accuracy in predicting heart failure survival using an RF model and the SMOTE technique [33]. particle swarm optimization (PSO) combined with artificial neural networks (ANN) displayed exemplary performance in accuracy, sensitivity, and specificity for heart attack prediction [34].

Utilizing the UCI Heart Disease dataset, the NB classifier produced results for individuals with and without heart disease that were 83.21% and 83.81% accurate, respectively [35]. RF produced the highest accuracy, at 95.63%. Moreover, a hybrid model combining SVM and quantum-behaved particle swarm optimization (QPSO) demonstrated better sensitivity, specificity, and precision than conventional approaches, achieving an accuracy of 96.31% [10]. This research illustrates how ML and optimization methods can enhance the prediction of cardiac disease and aid in clinical decision-making.

3. Methodology

Despite advancements in cardiovascular healthcare, accurately predicting heart attacks remains a significant global health challenge. Current methods for predicting MI may benefit from integrating a wider array of data, including medical history, genetic traits, and epidemiological factors such as lifestyle and demographics. Traditional approaches often overlook the diverse and complex nature of heart attacks, which vary significantly in symptoms and causes. This limitation is increasingly problematic given the rising incidence of heart attacks across different populations, highlighting the urgent need for more sophisticated prediction systems. AI offers a promising avenue for addressing these challenges by leveraging dense data to uncover complex patterns that traditional methods might miss. However, there remains a substantial gap in evaluating and comparing the effectiveness of various AI models in this specific context. Figure 1 illustrates the pathway of the research methodology.

3.1. Environmental Setup

The computational analysis and model development for this research were conducted using Google Colab, a cloud-based platform that provides a robust environment for DS and ML tasks. Google Colab offers several advantages, including access to powerful graphics processing units (GPU) and tensor processing units (TPU), which significantly accelerate the training of complex models. The platform also supports the execution of Python code, enabling seamless integration with popular DS libraries such as NumPy, Pandas, Scikit-learn, and Keras. The datasets were imported into Google Colab, where data preprocessing and model implementation were carried out. This setup ensured a collaborative and efficient workflow, with the added benefit of cloud storage for easy access and sharing of datasets and results. Utilizing Google Colab also facilitated the reproducibility of experiments, as the entire codebase and computational environment could be shared through notebooks, ensuring that all steps in the research process were transparently documented and easily replicable.

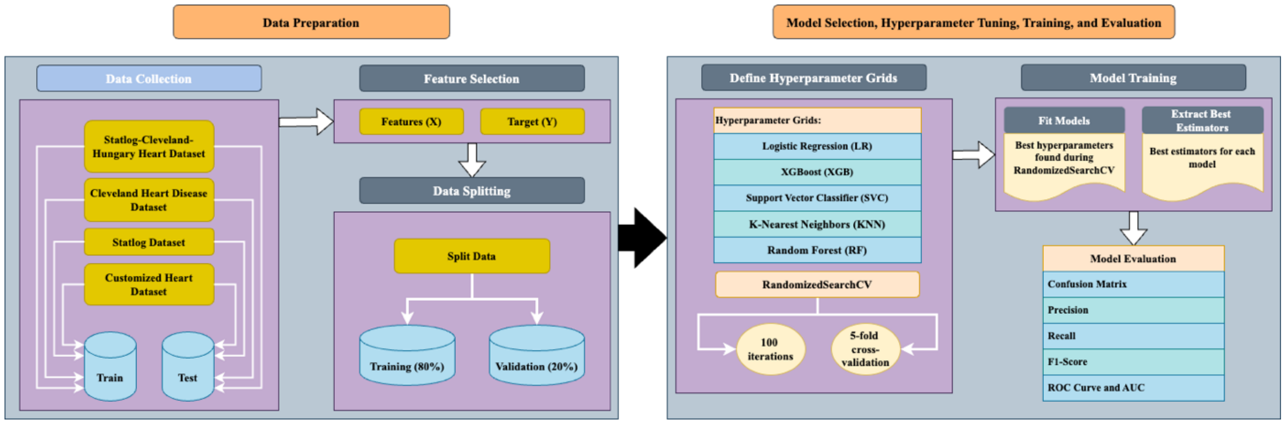


Fig.1. Proposed research methodology

3.2. Data Collection

The datasets selected for this study are widely recognized and frequently used in heart attack prediction research. To ensure a thorough analysis and the development of a robust model, three publicly available datasets were utilized. The *heart_statlog_cleveland_hungary_final* dataset was sourced from IEEE DataPort, which provides a curated collection of patient records, offering critical data for creating a model with broad clinical relevance. The *Heart_disease_cleveland_new* dataset was obtained from the UCI ML Repository, a well-established source of high-quality datasets, particularly in medical domains. The *Statlog* dataset, acquired from Kaggle, was also incorporated. These datasets are commonly referenced in the literature and have played a key role in advancing heart attack prediction methodologies.

3.3. Data Preprocessing

In this research, several data preprocessing steps were meticulously applied to ensure the quality and integrity of the dataset. Missing values were addressed using multiple imputation methods, including mean, median, and mode imputation, as well as forward and backward filling techniques to replace missing values with adjacent data points. Categorical values were transformed into numerical values to facilitate compatibility with ML algorithms. Normalization was performed using Min-Max scaling to scale features into the range [0, 1], ensuring that all variables contributed equally to the model and improving convergence rates during training. Outliers were handled by identifying and either removing or adjusting anomalous data points to prevent skewed results. Finally, feature selection was conducted to identify and retain the most relevant attributes, thereby reducing dimensionality and enhancing model performance. These preprocessing steps were crucial in preparing the dataset for robust and accurate heart attack prediction.

A. Dataset Selection

In this research, five datasets were utilized: Cleveland, UCI, Statlog, Harvard Cleveland, and a Custom Dataset, selected for the extensive use in previous studies and relevance to heart attack prediction. These datasets provide diverse clinical and demographic information, forming a robust foundation for model evaluation. The Cleveland dataset, with 13 input variables and 303 instances, is a widely recognized benchmark, while the UCI dataset, featuring 15 variables and 920 instances, adds complexity and richness. The Statlog dataset, with 13 variables and 270 instances, offers structured validation data, and the Harvard Cleveland dataset, comprising 18 variables and 299 instances, enhances feature diversity. To further generalize the model, a Custom Dataset was created by integrating the instances of these datasets and retaining 14 common features, resulting in a unified dataset with 1792 samples. During preprocessing, 602 instances were removed from the Custom Dataset to ensure data quality and consistency. This selection and preprocessing ensure the model's robustness across varied datasets and facilitate comparability with existing studies. Potential biases, such as sampling bias due to population specificity, feature omission from commonality constraints, and class imbalances, were mitigated through preprocessing steps, including normalization, feature alignment, and balanced training. These measures enabled the model to handle heterogeneous data effectively, and the results demonstrate strong performance across all datasets, validating the model's robustness and reliability.

B. Dataset Analysis

This analysis utilized three established datasets focused on cardiac conditions, specifically Cleveland, Statlog, Hungary, along with a customized dataset. Table 1 outlines the shared attributes across these datasets, alongside additional features, with the final attribute indicating heart attack status. To further explore feature distribution, Figures 2 and 3 provide visualizations showcasing the relationships between maximum heart rate, age, and sex, as well as the distribution of other features. These visual insights are essential for recognizing patterns and correlations within the data, forming a foundation for creating accurate heart attack prediction models.

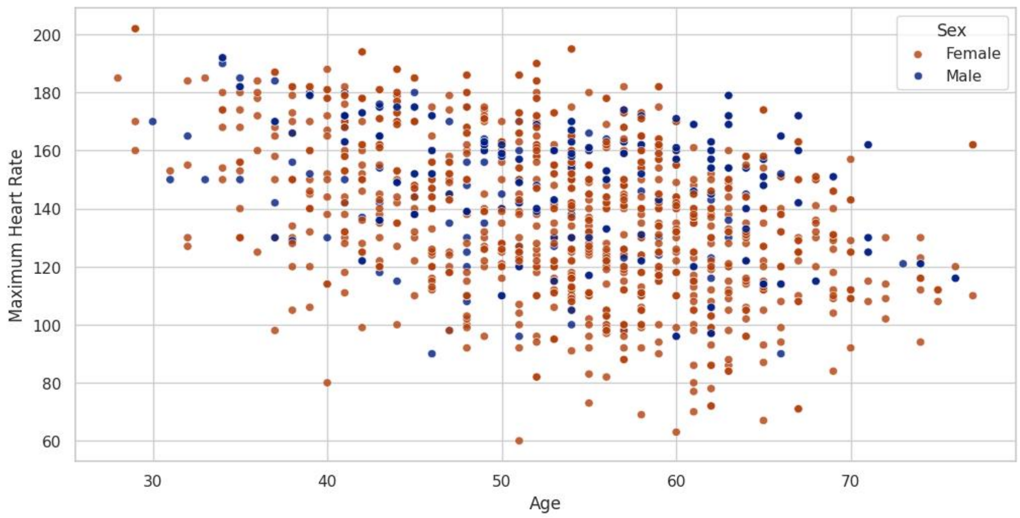
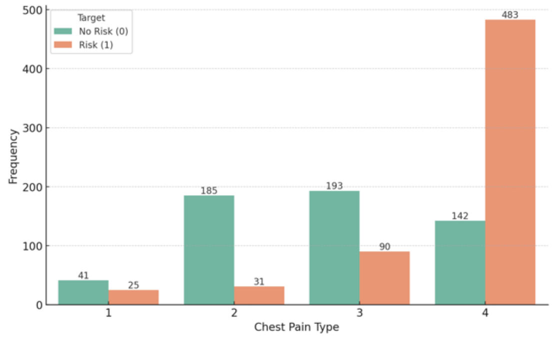
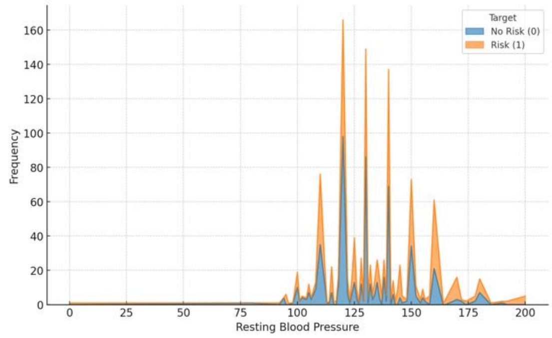


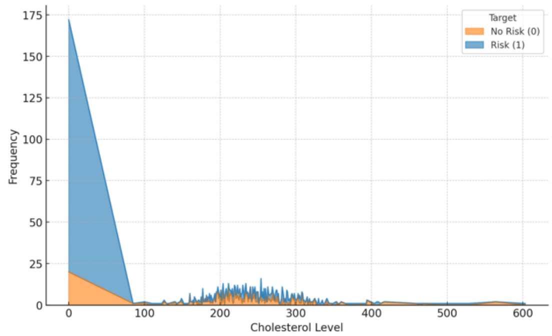
Fig.2. Relationship between maximum heart rate, age, and sex



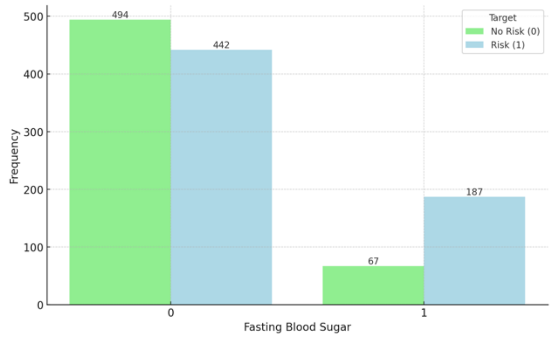
(a) Chest pain type in relation to target



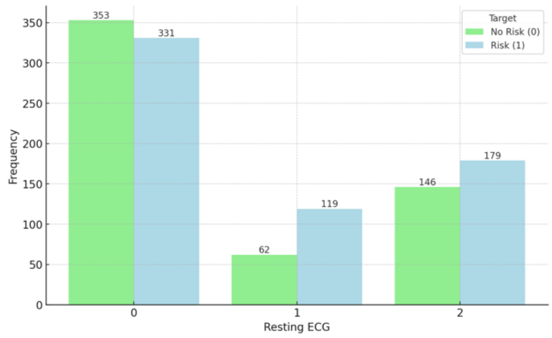
(b) Resting blood pressure in relation to target



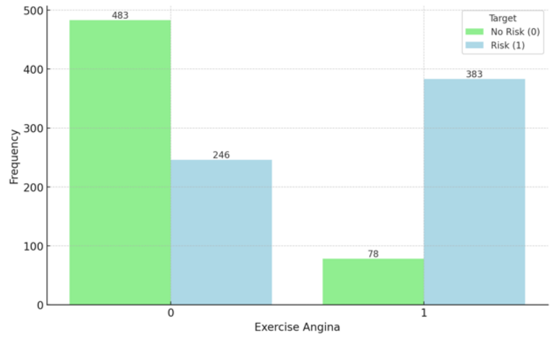
(c) Cholesterol level in relation to target



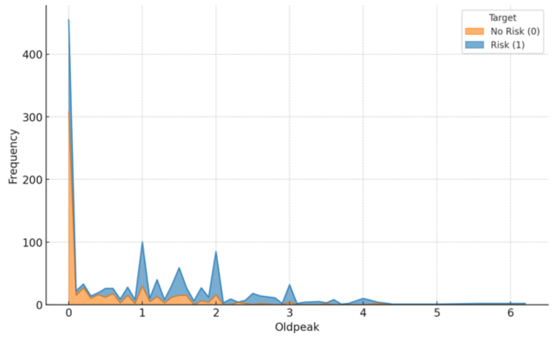
(d) Fasting blood sugar in relation to target



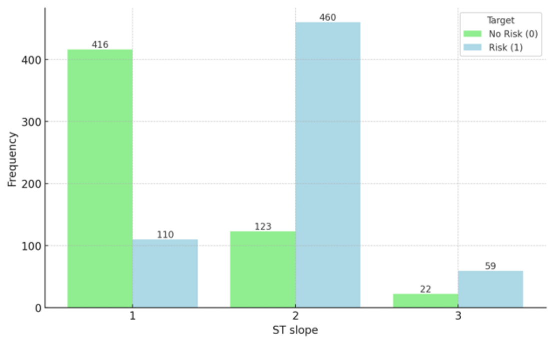
(e) Resting ECG in relation to target



(f) Exercise angina in relation to target



(g) ST depression induced by exercise relative to rest in relation to target



(h) The slope of the peak exercise ST segment in relation to target

Fig.3. Various distribution of the features

C. *Overview of the Datasets*

The datasets utilized in this study comprise a comprehensive set of cardiovascular health indicators essential for the analysis and prediction of heart attack. The dataset includes the following features: age, sex, chest pain type, resting blood pressure, cholesterol levels, fasting blood sugar, resting electrocardiographic results, maximum heart rate achieved, exercise-induced angina, ST depression induced by exercise relative to rest, the slope of the peak exercise ST segment, the number of major vessels, and thalassemia. Each feature is represented as an integer or floating-point number, with specific encoded values for categorical variables. For instance, sex is encoded as 0 for females and 1 for males, while chest pain type is categorized into non-angina (0), typical angina (1), atypical angina (2), and asymptomatic (3). Fasting blood sugar levels are encoded as 1 for values greater than 120 mg/dl and 0 otherwise. Resting electrocardiographic results are classified as normal (0), ST elevation (1), and left ventricular hypertrophy (2). Exercise-induced angina is indicated by 1 for presence and 0 for absence. The slope of the peak exercise ST segment is categorized as up-sloping (1), flat (2), and down-sloping (3). The dataset also includes a target variable indicating the presence (1) or absence (0) of heart attack. This rich collection of features provides a robust basis for developing predictive models to assess heart attack risk.

Table 1. Overview of features in the dataset

Feature Name	Description	Data Type
age	Age of the person	Numerical
sex	Gender of the person	Numerical
cp	Chest pain type	Numerical
trtbps	Resting blood pressure	Numerical
chol	Cholesterol	Numerical
fbs	Fasting blood sugar	Numerical
rest_ecg	Resting electrocardiographic results	Numerical
thalach	Maximum heart rate achieved	Numerical
exang	Exercise induced angina	Numerical
oldpeak	ST depression induced by exercise relative to rest	Floating Point
slope	The slope of the peak exercise ST segment	Numerical
target	Heart attack	Numerical

D. *Feature Importance*

For evaluating the importance of each feature in predicting heart attacks, tree-based models, specifically random forests (RF), were employed. These models are particularly favored in clinical data analysis due to the ability to handle various data types (e.g., int64 and float64) and the inherent feature selection capabilities. RF determines feature importance by assessing how features reduce impurity when creating decision trees. The model is generally robust to feature correlations because it randomly selects subsets of features for each split, thereby reducing the impact of multicollinearity. However, correlated features or interactions can still affect performance indirectly. Highly correlated features may dilute individual importance scores, making it harder to interpret contributions. Furthermore, if critical interactions between features are overlooked during the random sampling process, the model's performance may degrade. Addressing these issues through feature selection or engineering to highlight key interactions can significantly enhance model performance. This approach is beneficial as it considers both the direct and interaction effects of features on the prediction outcome, providing a holistic view of how the data influences the target variable. Table 2 presents the importance values of each feature as determined by the model.

Table 2. Feature importance values

Feature Name	Feature Importance
slope	0.19
thalach	0.13
cp	0.12
oldpeak	0.11
chol	0.11
age	0.09
trtbps	0.08
exang	0.07
sex	0.04
rest_ecg	0.03
fbs	0.02

The table 2 shows the importance of the features used in the relative prediction of heart attack. The most important predictor of the occurrence of the attack as per the predictive model is the ST_sloping with a score of 0.19. This shows that the slope of the ST segment in an ECG plays an important role in predicting heart attacks. This feature is an indicator of abnormalities in the ST_Segment that closely associate with the occurrence of cardiac events. The second most important predictor is chest_pain_type with a score of 0.13. This means that the different types of chest pain are more important in the prediction of heart attacks; therefore, specific patterns of chest pain are significant leading indicators of cardiac issues. The third most significant predictor, which represents the ST depression induced by exercise relative to rest, has an importance score of 0.12. According to the definition, this shows the importance of exercise-induced changes in the ST_Segmentation as one of the roles to assess the risk of heart attacks.

Max heart rate and cholesterol were both assigned an importance value of 0.11, indicating moderate significance but critical contributions to the model. Max heart rate under physical activity is an important measure of the overall condition of the heart, while cholesterol level is a well-known cardiovascular disease risk factor. Some features with moderate and inclusion importance are age 0.09, resting bps 0.08, and exercise-induced angina 0.07. The risk of heart attack usually increases with age, demonstrating a cumulative risk over time. On the other hand, blood pressure and angina are reflections of the current cardiovascular load. Variables such as sex, resting ECG, and fasting blood sugar appear to be less important features in the model after getting feature importance scores of 0.04, 0.03, and 0.02, respectively. Although these are significant issues, the values presented tend to be lower, indicating a nominal degree of contribution to this model among the list of features.

3.4. Model Selection

In the context of developing a robust heart attack prediction model, the selection of base classifiers for the stacking ensemble is guided by complementary strengths and diverse methodologies, which collectively enhance predictive performance. LR is chosen for its simplicity and effectiveness in binary classification problems, providing a strong probabilistic interpretation of the relationship between features and the target. KNN is included due to its non-parametric nature, which allows it to capture complex patterns in the data without assuming an underlying distribution. RF is a powerful ensemble method itself, offering high accuracy and resilience to overfitting by averaging multiple DT. support vector classifier (SVC) is selected for its robustness in high-dimensional spaces and its effectiveness in handling cases where the classes are not linearly separable. Lastly, XGBoost is a highly efficient and flexible gradient boosting algorithm known for its superior performance in many ML competitions and real-world applications.

By integrating these diverse classifiers, the stacking ensemble aims to leverage the strengths of each model while mitigating individual weaknesses. The LR model as the final estimator serves to effectively combine the predictions from the base classifiers, ensuring a balanced and well-calibrated final prediction. This approach is expected to improve overall model accuracy, precision, recall, and F1 score, ultimately leading to more reliable heart attack predictions.

A. Logistic Regression

LR is one of the top classifiers in supervised ML, which is applied to classification tasks. It is typically applied to a huge dataset in order to train the model more accurately. It is employed to address categorization issues where the answers are binary, meaning they can be either 0 or 1 [36].

$$z_{LR} = w_{LR}^{T} X + b_{LR} \tag{1}$$

This equation 1 represents the linear combination of the input features X with weights w_{LR} and bias b_{LR}. It calculates the linear predictor z_{LR} which is used to determine the probability of the positive class.

$$P_{LR}(y = 1 \mid X) = \sigma(z_{LR}) = \frac{1}{1+e^{-z_{LR}}} \tag{2}$$

In equation 2, the sigmoid function $\sigma(z_{LR})$ transforms the linear predictor z_{LR} into a probability value between 0 and 1, representing the likelihood of the positive class.

B. K-Nearest Neighbors

KNN is another often-used classifier for solving problems related to regression and classification. Its widespread use can be attributed to its ease of implementation when compared to other ML regressors. Numerous industries, such as finance, healthcare, forestry, image identification, etc., have used KNN for pattern detection. Because the KNN classifier generates classification rules from training samples, it is referred to as a packaging strategy. It fits into the supervised learning group. It has significant drawbacks and is simple to comprehend and apply, but its efficiency decelerates as data volume increases. Using the Euclidean distance, it determines the distance needed for categorization [37,38].

$$d(X, X_i) = \sqrt{\sum_{j=1}^{n} \left(x_j - x_{ij}\right)^2} \tag{3}$$

This equation 3 calculates the Euclidean distance between the input instance X and each training instance X_i. The distances are used to identify the k-nearest neighbors.

$$\hat{y}_{KNN} = mode(y_1, y_2, \dots, y_k) \tag{4}$$

In equation 4, for classification, the KNN model predicts the class $\hat{y}_{KNN}$based on the majority class (mode) of the k-nearest neighbors.

C. Random Forest

Another widely known classifier is the RF. For the given problem, during the training time, it builds several DT and outputs the mode of the classes for classification. RF uses the ensemble learning method largely for classification. At each split, a random subset of characteristics is considered, adding diversity to the trees. Each tree is trained on a bootstrap sample of the dataset. The results of all the trees are combined to get the final classification in order to not overfit and increase generalization performance [36,39].

$$f_t(X) = w_{q_t}(X) \tag{5}$$

Each decision tree f_t in the RF makes a prediction $w_{qt}(X)$ for the input instance X. Here, $q_t(X)$ maps the input to a specific leaf node in the tree, and $w_{q_t}(X)$ is the prediction associated with that leaf, as defined in equation 5.

$$\hat{y}_{RF} = mode(f_1(X), f_2(X), \dots, f_T(X)) \tag{6}$$

In equation 6, the final prediction $\hat{y}_{RF}$ of the RF is the majority vote (mode) of the predictions from all individual trees $f_1, f_2 \dots, f_t$.

D. Support Vector Machine

Being a ML regressor, it most likely solves the regression and classification problems [40]. Its primary use is in handling classification-related problems. Because it can yield extremely precise results with fewer processing resources, the DM community uses it primarily. Finding the best hyperplane to divide the data into two classes is its main objective. This method, however, rejects two significant obstacles: changing the parameters and choosing an appropriate primary function [41].

$$f_{SVM}(X) = w_{SVM}^{T} X + b_{SVM} \tag{7}$$

This equation 7 defines the decision function $f_{SVM}(X)$ for the SVM, which computes a linear combination of the input features X with weights w_{SVM} and bias b_{SVM}

$$\hat{y}_{SVM} = sign(f_{SVM}(X)) \tag{8}$$

In equation 8, the SVM predicts the class $\hat{y}_{SVM}$ based on the sign of the decision function $f_{SVM}(X)$. If the value is positive, the class is 1; if negative, the class is -1.

E. Extreme Gradient Boosting

XGBoost is an improved version of the gradient boosting algorithms that are designed to be fast and effective. In classification problems, it constructs a series of DT one after another, where each new tree tries to fix the mistakes of the previous ones by minimizing some specified loss function. The algorithm uses various regularization methods to avoid overfitting and blends gradient descent with second-order Taylor approximation for efficient optimization. The final classification result is the weighted sum of the predictions of all trees, whereby the model can get more accurate and robust [39,40].

$$Obj(\theta) = \sum_{i=1}^{n} l(\mathrm{y_i}, \hat{y}_i) + \sum_{k=1}^{k} \Omega(\mathrm{f_k}) \tag{9}$$

The objective function $Obj(\theta)$ includes the loss function l to measure the difference between the actual $\mathrm{y_i}$ and predicted $\hat{y}_i$ values, and the regularization term $\Omega(\mathrm{f_k})$ to control the complexity of the trees and prevent overfitting, as defined in equation 9.

$$\hat{y}_{XGB}^{(t)} = \hat{y}_{XGB}^{(t-1)} + f_t(X) \tag{10}$$

In equation 10, the XGBoost makes predictions by adding the output of each new tree $f_t(X)$ to the previous prediction $\hat{y}_{XGB}^{(t-1)}$. This iterative process continues until the final prediction $\hat{y}_{XGB}$ is obtained.

3.5. Criterias of Model Evaluation

The confusion matrix has been used for the proposed research in order to precisely analyze the model. It is composed of a 2 x 2 matrix, where the trained model's projected value is displayed on one side and the dataset's actual values are displayed on the other. The performance measures accuracy, precision, recall, and F1 score are elucidated here. The following formula provides accuracy.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{11}$$

The measure of the total number of correct predictions made from all of the data's input values is called accuracy in equation 11. Equation 12 provides the precision value, which is used to determine the model's accuracy.

$$Precision = \frac{TP}{TP+FP} \tag{12}$$

Similar to this, recall gauges the model's validity by assessing how well a class is predicted using all of the input values that the model has received. Equation 13 provides the recall formula.

$$Recall = \frac{TP}{TP+FN} \tag{13}$$

The F1-Score is the harmonic mean of Precision and Recall, providing a single metric that balances both concerns. It is particularly useful when the dataset is imbalanced, as it accounts for both false positives and false negatives. Equation 14 provides the formula for the f1-score.

$$F1 - Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{14}$$

The Area Under the Receiver Operating Characteristic Curve (AUC-ROC) measures the model's ability to distinguish between positive and negative instances across different threshold values. AUC-ROC is effective for evaluating the overall performance of the model, as it considers all possible classification thresholds. Equation 15 provides the AUC formula.

$$AUC = \int_0^1 TPR(t)\, dFPR(t) \tag{15}$$

Here, the True Positive Rate at threshold t is represented as $TPR(t)$ and $FPR(t)$ is the False Positive Rate at threshold t, defined as in equation 16.

$$FPR = \frac{FP}{FP+TN} \tag{16}$$

By using these metrics, a comprehensive assessment of the heart attack prediction model can be ensured, balancing accuracy, sensitivity, and the model's ability to generalize across different thresholds.

4. Proposed Model

The proposed model aims to enhance the accuracy and efficiency of heart attack prediction by leveraging a robust stacking ensemble approach. This model integrates the predictive strengths of several base classifiers, including LR, KNN, RF, SVM, and XGBoost, through a meticulous hyperparameter tuning and model training process. By combining the individual predictions of these diverse algorithms into a meta-feature matrix and training a final LR estimator on this matrix, the proposed model aspires to achieve superior predictive performance. This section details the design, implementation, and evaluation of this comprehensive model, highlighting its potential to provide reliable and rapid heart attack predictions.

Model Architecture

The proposed model employs a stacking ensemble framework, integrating multiple machine-learning algorithms to improve heart attack prediction performance. The dataset utilized comprises a range of clinical and demographic attributes associated with heart attack presence or absence. The data was initially separated into independent variables and the target outcome. It was then divided into training and validation sets in an 80:20 ratio, with 80% of the data reserved for training and 20% for validation. This approach maintained the integrity of the validation set, ensuring an unbiased evaluation of the model's generalization ability.

The crux of the architecture is the hyperparameter tuning applied to each base classifier via RandomizedSearchCV. This approach performs a comprehensive search over specified hyperparameter distributions to determine optimal configurations. The model incorporates five types of base classifiers: LR, XGBoost, SVC, KNN, and RF, each with a tailored hyperparameter space suited to its characteristics. For LR, the C parameter was sampled from a uniform distribution between 0.1 and 10. For XGBoost, the n_estimators parameter was sampled from a discrete range of 50 to 150, while the max_depth parameter was drawn from a range of 3 to 10. For SVC, the C parameter was sampled from a uniform distribution between 0.1 and 10, with the kernel parameter allowed to vary between linear and rbf. For KNN, the n_neighbors parameter was selected from a range of 3 to 10. For RF, the n_estimators parameter was sampled from 50 to 150, and the max_depth parameter included values of None, 10, 20, and 30.

RandomizedSearchCV was configured with 100 iterations and a five-fold cross-validation strategy to ensure robust and unbiased evaluation of hyperparameter combinations. This configuration balanced computational efficiency with thorough exploration, ensuring the base models operated under optimal conditions while contributing effectively to the ensemble. After determining the optimal hyperparameters, the best models yielded by RandomizedSearchCV were extracted and integrated into the stacking technique as base classifiers.

Stacking, an ensemble learning method, combines the outputs of multiple base learners, each trained to solve the same problem but from different perspectives. The meta-classifier in this architecture, a LR model, aggregates the predictions of the base classifiers to produce the final output, leveraging the diverse strengths of the ensemble while mitigating individual weaknesses. This meticulous hyperparameter tuning and stacking approach ensures the robustness and superior performance of the proposed model.

In the proposed architecture, the meta-classifier was a Logistic Regression model trained on the predictions of the base learners. Such a two-level structure allows the stacking classifier to learn and correct the shortcomings of the base classifiers to improve predictive performance. The final determination by the meta-classifier considers the outputs from all base models, effectively leveraging the strengths while mitigating the weaknesses of each. Mathematically, the stacking classifier can be represented as follows: Let $P_{LR}(y = 1 \mid X)$, $P_{KNN}(y = 1 \mid X)$, $P_{RF}(y = 1 \mid X)$, $P_{SVM}(y = 1 \mid X)$, and $P_{XGB}(y = 1 \mid X)$ denote the probabilities of the positive class predicted by the LR, KNN, RF, SVM, and XGBoost classifiers, respectively.

$$Z = \begin{pmatrix} P_{LR}(y = 1 \mid X) \\ P_{KNN}(y = 1 \mid X) \\ P_{RF}(y = 1 \mid X) \\ P_{SVM}(y = 1 \mid X) \\ P_{XGB}(y = 1 \mid X) \end{pmatrix} \tag{17}$$

In equation 17, the outputs of the base models (predicted probabilities for the positive class) are combined into a new feature vector Z. Each element in Z represents the prediction from one of the base models. The final estimator (LR) takes the combined predictions Z and computes a linear combination with weights w_{stack} and biases b_{stack} to produce the final prediction score z_{stack}, as defined in equation 18.

$$z_{stack} = w_{stack}^{T} Z + b_{stack} \tag{18}$$

In equation 19, the sigmoid function $\sigma(z_{stack})$ transforms the final prediction score z_{stack} into a probability value between 0 and 1, representing the likelihood of the positive class.

$$P_{stack}(y = 1 \mid Z) = \sigma(z_{stack}) = \frac{1}{1+e^{-z_{stack}}} \tag{19}$$

The integration of outputs is performed by aggregating the predicted probabilities of the positive class from each base classifier into the meta-feature matrix Z. This structured representation enables the meta-classifier to analyze the contributions of each base classifier to the final prediction. During training, the meta-classifier learns optimal weights w_{stack} for each base classifier's output, which are determined based on the individual predictive performance. Base classifiers with higher accuracy are assigned greater weights, ensuring that predictions influence the final output more significantly. This approach inherently balances the contributions of diverse base classifiers, leveraging strengths while mitigating weaknesses.

The training of the stacking classifier enables predictions on the validation set, allowing performance assessment through metrics such as accuracy, precision, recall, and F1 score. These metrics collectively evaluate the model's ability to classify unknown samples accurately. Additionally, a confusion matrix is generated, illustrating true positives, true negatives, false positives, and false negatives, which helps visualize the types of errors made by the model. The Receiver Operating Characteristic (ROC) curve is also plotted, and the area under the curve (AUC) is computed to assess the model's discriminative capability. The ROC curve provides insights into the trade-off between sensitivity and specificity, while the AUC quantifies the overall effectiveness of distinguishing between classes. This comprehensive evaluation ensures that the stacking ensemble leverages the predictions of individual classifiers to deliver a reliable and robust model for heart attack prediction.

The notion of having a sort of strong combination of multiple base estimators and a strong metamodel in the concept of stacking is what actually makes the stacking ensemble model robust in its performance. The employed architecture would correctly capture the real manifold information since it considers the capabilities of the algorithms used through varied base classifiers. This actually makes it a robust and powerful tool for heart attack prediction. The optimization process will fine-tune the ensemble hyperparameters to ensure maximum optimal operation of each component, delivering high accuracy and reliable final set predictions. Comprehensive evaluation metrics have proved the efficiency of the stacking ensemble, which lends it an applicable tool for the development of clinical decision support systems. Figure 4 gives a complete visual representation of the comprehensive model architecture for the heart attack prediction system.

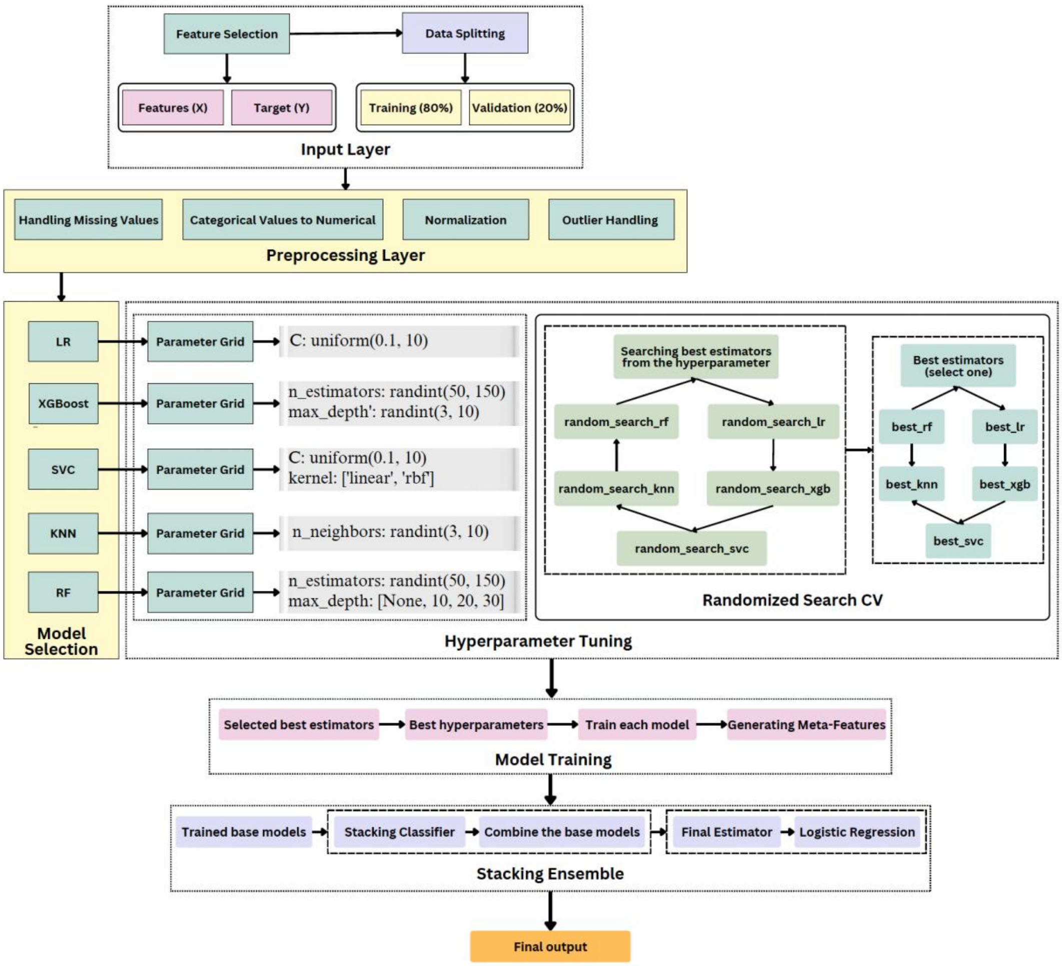


Fig.4. Architecture of the proposed model

5. Result Analysis

The correlation heatmap visually represents the relationships between various features in the dataset related to heart attack prediction. The heatmap in Figure 5 uses color coding to indicate the strength and direction of correlations, with values ranging from -1 to 1. A value closer to 1 implies a strong positive correlation, while a value closer to -1 indicates a strong negative correlation. Key observations include a moderately positive correlation between 'chest pain type' and 'target' (0.46), suggesting that certain types of chest pain are associated with a higher likelihood of a heart attack. 'Exercise angina' also shows a notable positive correlation with 'target' (0.48), indicating its potential as a predictive feature. Conversely, 'max heart rate' demonstrates a negative correlation with 'target' (-0.41), implying that higher maximum heart rates are associated with a lower risk of heart attack. The correlation between 'ST slope' and 'target' (0.51) is the strongest observed positive correlation, highlighting its significance in heart attack prediction. The overall pattern of correlations provides insights into the interdependencies among the features, which can guide feature selection and model development in predicting heart attacks.

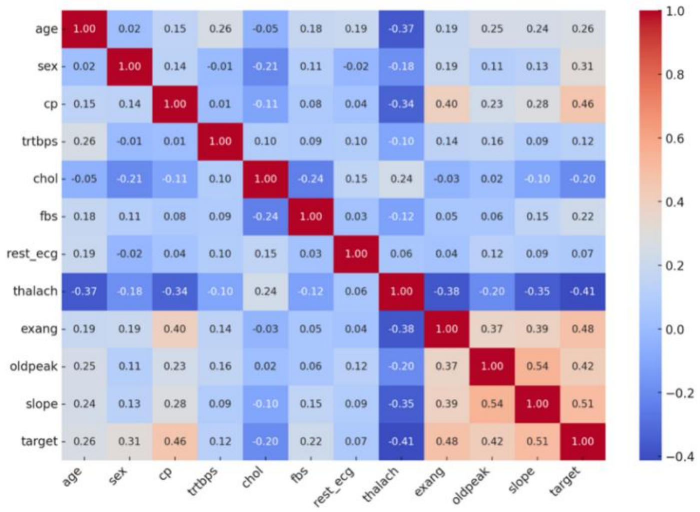


Fig.5. Correlation matrix

The confusion matrix in Figure 6 illustrates the performance of the heart attack prediction model by displaying the actual versus predicted classifications. The matrix reveals that the model accurately classified 100 negative cases and 128 positive cases, demonstrating high precision and recall in both categories. However, it also indicates 7 false positives, where the model incorrectly predicted heart attacks in patients who did not have one, and 3 false negatives, where it failed to predict heart attacks in patients who actually had one. These misclassifications have important implications in a clinical context. False positives, while causing unnecessary stress, anxiety, and additional diagnostic costs for patients, are less critical compared to false negatives. False negatives represent a missed diagnosis of heart attack risk, which could delay timely interventions, potentially leading to severe adverse outcomes, including mortality. To address these risks, the model's classification threshold can be adjusted to prioritize reducing false negatives, or complementary diagnostic tools and clinical evaluations can be integrated to improve overall reliability. Despite these misclassifications, the model’s low false positive and false negative rates indicate its robustness and reliability, supporting its application in clinical settings for early detection and prevention strategies.

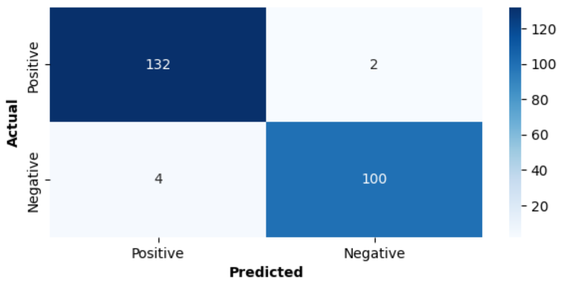


Fig.6. Confusion matrix of the proposed model

The ROC curve shown in Figure 7 assesses the heart attack prediction model's performance by depicting the relationship between the true positive rate and the false positive rate across different threshold values. With an area under the curve (AUC) of 0.97, the model demonstrates exceptional performance, as an AUC of 1 signifies a perfect model. The steep ascent in the curve highlights the model's ability to achieve a high true positive rate while maintaining a low false positive rate, further underscoring its predictive effectiveness. This suggests that the model is highly effective in distinguishing between patients who have and do not have a heart attack. The curve's shape further indicates that the model maintains a high sensitivity while minimizing the false positive rate, highlighting its robustness and reliability for clinical applications. The high AUC value reinforces the model's potential utility in predictive diagnostics, ensuring early and accurate detection of heart attacks.

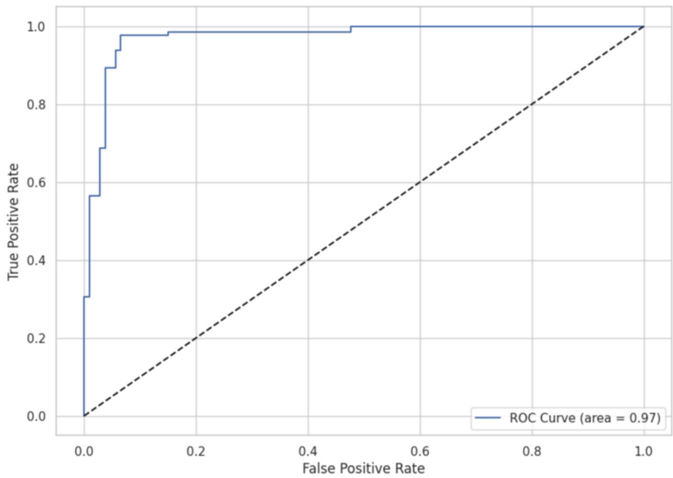


Fig.7. ROC curve

Table 3. Accuracy comparison in heart attack prediction models

Reference	Model Description	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
-	LR	84.03	84	87	86
-	LR_CV	85.71	87	87	87
-	RF	94.54	94	96	95
-	ET	93.70	95	94	94
-	XGB	2.02	91	95	93
-	LGBM	94.54	95	95	95
-	SVM	72.69	78	70	74
-	Linear SVC	76.47	71	96	82
-	KNN	73.53	77	74	75
-	NB	86.13	87	88	87
-	Nearest Centroid	53.78	62	42	50
-	Bernoulli NB	77.31	78	82	80
-	Linear Discriminant	85.29	87	86	87
-	Label Propagation	66.39	95	41	57
-	Calibrated Classifier CV	81.51	83	84	83
-	Quadratic Discriminant	84.03	85	86	86
-	MLP Classifier	84.03	83	89	86
-	Bagging Classifier	91.18	94	89	92
-	AdaBoost Classifier	88.66	88	92	90
-	Gradient Boosting	91.18	91	93	92
-	Decision Tree	89.08	93	87	90
[10]	QPSO-SVM	96.31	-	-	-
[23]	OCI-LSTM	97.11	-	-	97.32
[24]	EHMFFL	91.8	91.4	91.4	92.8
[30]	Self-Attention-Based Transformer Model	96.51	-	-	-
[31]	OCSCatBoost	73.67	-	72.17	73.29
Proposed	**Hybrid Ensemble**	**97.48**	**97.06**	**98.51**	**97.78**

Table 3 compares the performance metrics of various models used for heart attack prediction, including accuracy, precision, recall, and F1-score. The models reviewed are OCI-LSTM [9], EHMFFL [35], a self-attention-based transformer model [17], OCSCatBoost [18], QPSO-SVM [23], and the Proposed Hybrid Ensemble. Among these, the proposed hybrid ensemble model demonstrates the highest overall performance, with an accuracy of 97.48%, a precision of 97.64%, a recall of 98.51%, and an f1-score of 97.78%. These metrics indicate that the proposed hybrid ensemble model not only achieves the highest accuracy but also excels in precision, recall, and f1-score, making it the most effective model for heart attack prediction among the ones compared. The superior performance across all evaluated metrics suggests that the proposed hybrid ensemble model is highly reliable and robust, providing an optimal balance between identifying true positive cases and minimizing false positives. Table 4 shows the ablation study with different classified datasets.

Table 4. Ablation study with different classified dataset

Datasets	Features Count	Total Instance	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Cleveland [42]	14	303	90.13	87.8	93.5	90.6
UCI [43]	16	920	95.00	93.90	95.24	94.56
Harvard Cleveland [44]	19	299	96.99	97.06	97.63	97.35
Statlog [45]	14	270	93.70	93.38	94.07	93.73
Custom Dataset [46]	14	1190	96.22	96.14	96.30	96.22

6. Discussion

The proposed model utilizes the Stacking Ensemble approach, which aggregates the output of different ML algorithms such as LR, KNN, RF, SVM, and XGBoost to enhance heart attack prediction accuracy. This approach achieves a prediction accuracy of 97.48%, outperforming other methods. Additionally, individual classifiers such as RF with 94.54%, LR with 84.03%, and SVM with 72.69% demonstrate significant improvements, delivering better results compared to existing state-of-the-art methods like the OCI-LSTM model with 97.11% and the EHMFFL algorithm with 91.8%.

The RF model demonstrated robustness but tended to overfit on small datasets, which reduced its generalized performance. SVM performed well in high-dimensional data; however, the nonlinear relationships within the dataset posed limitations to its effectiveness. In contrast, XGBoost exhibited strong individual performance, although it fell short when compared to the combined predictive power of the ensemble model.

The architecture of the stacking ensemble model capitalizes on the strengths of each base model while effectively mitigating respective weaknesses. For instance, RF enhances the simplicity and interpretability of LR with its capability to handle complexity, while also incorporating the precision offered by SVM. The ensemble captures multifaceted information from the data in a more nuanced manner, resulting in improved predictive performance and reliability.

The interpretability of the proposed model is crucial for its integration into real-world clinical workflows. To ensure clinicians can understand and trust the predictions, feature importance analysis using RF was employed to identify key contributing features such as ST slope, chest pain type, and maximum heart rate, which align with established clinical indicators of heart attack risk. The stacking ensemble further enhances interpretability by using a LR meta-classifier, which associates feature contributions with specific weights. Moreover, the model outputs probabilities alongside classifications, allowing clinicians to assess the confidence of each prediction. For example, a high probability of the positive class signals a stronger likelihood of heart attack risk, prompting closer examination or early intervention. These features ensure that the model functions as a transparent tool to support clinical judgment rather than as an opaque black-box system.

The practical implications of the model are extensive, particularly in the context of clinical care. Early and accurate predictions can facilitate timely interventions, potentially saving countless lives and significantly reducing the burden on healthcare systems and providers. An integrated approach to diverse datasets enables a more comprehensive analysis, strengthening the reliability and applicability of predictions across varied patient populations. Also, its high degree of accuracy makes it possible for better stratification of risk and individualized treatment. Healthcare providers can use such predictions, for instance, to monitor high-risk patients more closely, implement preventive measures, and streamline resource allocation. All these quantum improvements ought to, at least, translate into major improvements in patient outcomes and public health at large.

This renders this model highly beneficial to cardiologists, general practitioners, as well as to the healthcare institutions specializing in cardiovascular disease management. It can very easily be integrated into the EHR systems and enable real-time risk assessments that may be used by the clinical experts for supporting decisions. The benefits would also extend to the patients by providing advice and appropriate interventions based on predicted risk, enabling proactive health management.

The model can also be used as a source of more research by researchers, besides health care providers, into predictive analytics and personalized medicine. It can also be applied to the work of insurance companies to accurately quantify risk profiles, enabling the design of personalized insurance packages and pricing, thereby optimizing services and enhancing customer satisfaction. Overall, its high predictive accuracy and robustness confirm the actual value of this model for real-world applications, which is expected to make it a very powerful tool in the ongoing fight against cardiovascular attack.

7. Conclusions

This study chiefly aimed at developing a highly accurate and robust early heart attack detection prediction model based on advanced ML techniques. It is proposed that multiple integrated datasets would be combined into a stacking ensemble method for improving the predictive power of ordinary models. Here as well, while integrating several ML algorithms, very meticulous data preprocessing was performed to develop LR, KNN, RF, SVM, and XGBoost. Thus, the model achieved an accuracy of 97.48%, which significantly outperforms individual classifiers and other current state-of-the-art approaches.

The results suggest that the stacking ensemble model effectively leverages the strengths of each base classifier, amplifying predictive capabilities while mitigating weaknesses. This integration enhances the model's robustness and reliability, making it well-suited for clinical applications. Furthermore, the combination of different datasets ensures stability across varied patient demographics, supporting its potential applicability in real-world settings. Despite these achievements, there remains a gap between the conceptual accuracy of the prediction model and the full-scale realism required for precise real-world applications. Thus, there is a continuing need to refine and improve this technology, with increasing consideration of varied data types to accommodate advanced deep learning techniques that further enhance accuracy and reliability.

Statistical tests, such as confidence intervals and p-values, were not included in the current analysis due to certain constraints, but these will be incorporated in future work to enhance the reliability of the findings. Furthermore, the results presented are currently limited to specific datasets and lack external validation on unseen data. Evaluating model generalizability across diverse populations and scenarios will be a priority in future studies. While comparisons with other models, such as OCI-LSTM and EHMFFL, are included, there is a need for more in-depth evaluation, which will also be addressed in subsequent work to provide a comprehensive analysis of the model’s capabilities.

This study concludes with a novel approach to heart attack prediction that synergizes multiple ML techniques, implemented with strict preprocessing and evaluation methods on the training data. This results in a powerful tool for timely and personalized cardiovascular disease management, with immense potential to assist both patients and the healthcare system. This work forms a strong foundation for further research and development in the field, where even more accurate and efficient predictive models can be created.

Reference

[1] P. Ghosh et al., "Efficient Prediction of Cardiovascular Disease Using Machine Learning Algorithms With Relief and LASSO Feature Selection Techniques," in IEEE Access, vol. 9, pp. 19304-19326, 2021, doi: 10.1109/ACCESS.2021.3053759.

[2] H. Liu, R. G. Crespo, and O. S. Martínez, “Enhancing Privacy and Data Security across Healthcare Applications Using Blockchain and Distributed Ledger Concepts,” Healthcare, vol. 8, no. 3, p. 243, Jul. 2020, doi: https://doi.org/10.3390/healthcare8030243.

[3] S. Song, T. Chen, and G. Antoniou, “ANFIS Models for Heart Disease Prediction,” 2021 the 5th International Conference on Innovation in Artificial Intelligence, Mar. 2021, doi: https://doi.org/10.1145/3461353.3461354.

[4] S. Shehzadi, M. A. Hassan, M. Rizwan, N. Kryvinska, and K. Vincent, “Diagnosis of Chronic Ischemic Heart Disease Using Machine Learning Techniques,” Computational Intelligence and Neuroscience, vol. 2022, pp. 1–9, Jun. 2022, doi: https://doi.org/10.1155/2022/3823350.

[5] M. A. Alim, S. Habib, Y. Farooq and A. Rafay, "Robust Heart Disease Prediction: A Novel Approach based on Significant Feature and Ensemble learning Model," 2020 3rd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, Pakistan, 2020, pp. 1-5, doi: 10.1109/iCoMET48670.2020.9074135.

[6] M. E. Ozer, P. O. Sarica, and K. Y. Arga, “New Machine Learning Applications to Accelerate Personalized Medicine in Breast Cancer: Rise of the Support Vector Machines,” OMICS: A Journal of Integrative Biology, vol. 24, no. 5, pp. 241–246, May 2020, doi: https://doi.org/10.1089/omi.2020.0001.

[7] L. Zahid et al., "A Spectrogram-Based Deep Feature Assisted Computer-Aided Diagnostic System for Parkinson’s Disease," in IEEE Access, vol. 8, pp. 35482-35495, 2020, doi: 10.1109/ACCESS.2020.2974008.

[8] J. Zhang, Y. Zhou, and Q. Luo, “An improved sine cosine water wave optimization algorithm for global optimization,” Journal of Intelligent & Fuzzy Systems, vol. 34, no. 4, pp. 2129–2141, Apr. 2018, doi: https://doi.org/10.3233/jifs-171001.

[9] R. Jousef, “Heart Attack Prediction using Machine Learning Algorithms,” Journal of Electrical Systems, vol. 20, no. 5s, pp. 1428–1436, Apr. 2024, doi: https://doi.org/10.52783/jes.2474.

[10] E. I. Elsedimy, S. M. M. AboHashish, and F. Algarni, “New cardiovascular disease prediction approach using support vector machine and quantum-behaved particle swarm optimization,” Multimedia Tools and Applications, Aug. 2023, doi: https://doi.org/10.1007/s11042-023-16194-z.

[11] D. Zhao, “Epidemiological Features of Cardiovascular Disease in Asia,” JACC: Asia, vol. 1, no. 1, pp. 1–13, Jun. 2021, doi: https://doi.org/10.1016/j.jacasi.2021.04.007.

[12] S. Steven et al., “Vascular Inflammation and Oxidative Stress: Major Triggers for Cardiovascular Disease,” Oxidative Medicine and Cellular Longevity, vol. 2019, Jun. 2019, doi: https://doi.org/10.1155/2019/7092151.

[13] A. Mehmood et al., “Prediction of Heart Disease Using Deep Convolutional Neural Networks,” Arabian Journal for Science and Engineering, vol. 46, no. 4, pp. 3409–3422, Jan. 2021, doi: https://doi.org/10.1007/s13369-020-05105-1.

[14] S. K. J. and G. S., "Prediction of Heart Disease Using Machine Learning Algorithms," 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), Chennai, India, 2019, pp. 1-5, doi: 10.1109/ICIICT1.2019.8741465.

[15] G. Heusch, “Coronary blood flow in heart failure: cause, consequence and bystander,” Basic Research in Cardiology, vol. 117, no. 1, Jan. 2022, doi: https://doi.org/10.1007/s00395-022-00909-8.

[16] J. Choi, S. Lee, M. Chang, Y. Lee, G. C. Oh, and H.-Y. Lee, “Deep learning of ECG waveforms for diagnosis of heart failure with a reduced left ventricular ejection fraction,” Scientific Reports, vol. 12, no. 1, Aug. 2022, doi: https://doi.org/10.1038/s41598-022-18640-8.

[17] Z. L. Steinberg and H. S. Singh, “How to plan and perform a diagnostic catheterisation in adult patients with congenital heart disease,” Heart, p. heartjnl-2022-321005, Oct. 2022, doi: https://doi.org/10.1136/heartjnl-2022-321005.

[18] G. Ye et al., “Short-Term Risk Estimation and Treatment Planning for Cardiovascular Disease Patients after First Diagnostic Catheterizations with Machine Learning Models,” Applied sciences, vol. 13, no. 8, pp. 5191–5191, Apr. 2023, doi: https://doi.org/10.3390/app13085191.

[19] N. E. Menaceur, S. Kouah, and Makhlouf Derdour, “IoT-Based Intelligent Medical Decision Support System for Cardiovascular Diseases,” Lecture notes in networks and systems, pp. 122–126, Jan. 2023, doi: https://doi.org/10.1007/978-3-031-44146-2_12.

[20] G.-C. Javier, R. Crisostomo-Castro, J. Vilabuleje, J. López-Goycochea, and J. N. Valenzuela, “HEART ATTACK PREDICTION USING MACHINE LEARNING: A COMPREHENSIVE SYSTEMATIC REVIEW AND BIBLIOMETRIC ANALYSIS,” 102, 2023, Available: https://hdl.handle.net/20.500.13067/3185.

[21] M. A. Naser, A. A. Majeed, M. Alsabah, T. R. Al-Shaikhli, and K. M. Kaky, “A Review of Machine Learning’s Role in Cardiovascular Disease Prediction: Recent Advances and Future Challenges,” Algorithms, vol. 17, no. 2, p. 78, Feb. 2024, doi: https://doi.org/10.3390/a17020078.

[22] N, Nissa, S. Jamwal, and M. Neshat, “A Technical Comparative Heart Disease Prediction Framework Using Boosting Ensemble Techniques,” Computation, vol. 12, no. 1, pp. 15–15, Jan. 2024, doi: https://doi.org/10.3390/computation12010015.

[23] T.K. Revathi, S. Balasubramaniam, V. Sureshkumar, and S. Dhanasekaran, “An Improved Long Short-Term Memory Algorithm for Cardiovascular Disease Prediction,” Diagnostics, vol. 14, no. 3, pp. 239–239, Jan. 2024, doi: https://doi.org/10.3390/diagnostics14030239.

[24] M. Shokouhifar, M. Hasanvand, E. Moharamkhani, and F. Werner, “Ensemble Heuristic–Metaheuristic Feature Fusion Learning for Heart Disease Diagnosis Using Tabular Data,” Algorithms, vol. 17, no. 1, pp. 34–34, Jan. 2024, doi: https://doi.org/10.3390/a17010034.

[25] I. Rojek, P. Kotlarz, M. Kozielski, M. Jagodziński, and Z. Królikowski, “Development of AI-Based Prediction of Heart Attack Risk as an Element of Preventive Medicine,” Electronics, vol. 13, no. 2, p. 272, Jan. 2024, doi: https://doi.org/10.3390/electronics13020272.

[26] A. Hammoud, A. Karaki, R. Tafreshi, S. Abdulla, and M. Wahid, “Coronary Heart Disease Prediction: A Comparative Study of Machine Learning Algorithms,” Journal of advances in information technology, vol. 15, no. 1, pp. 27–32, Jan. 2024, doi: https://doi.org/10.12720/jait.15.1.27-32.

[27] T. Ullah et al., "Machine Learning-Based Cardiovascular Disease Detection Using Optimal Feature Selection," in IEEE Access, vol. 12, pp. 16431-16446, 2024, doi: 10.1109/ACCESS.2024.3359910.

[28] S. Ware, S. K. Rakesh, and B. Choudhary, “Heart Attack Prediction by using Machine Learning Techniques,” International Journal of Recent Technology and Engineering, vol. 8, no. 5, pp. 1577–1580, Jan. 2020, doi: https://doi.org/10.35940/ijrte.d9439.018520.

[29] R. Mia et al., “Exploring Machine Learning for Predicting Cerebral Stroke: A Study in Discovery,” Electronics (Basel), vol. 13, no. 4, pp. 686–686, Feb. 2024, doi: https://doi.org/10.3390/electronics13040686.

[30] A. Ur Rahman, Y. Alsenani, A. Zafar, K. Ullah, K. M. Rabie, and T. Shongwe, “Enhancing heart disease prediction using a self-attention-based transformer model,” Scientific Reports, vol. 14, no. 1, Jan. 2024, doi: https://doi.org/10.1038/s41598-024-51184-7.

[31] Z. Qiu, Y. Qiao, W. Shi, and X. Liu, “A robust framework for enhancing cardiovascular disease risk prediction using an optimized category boosting model,” Mathematical Biosciences and Engineering, vol. 21, no. 2, pp. 2943–2969, Jan. 2024, doi: https://doi.org/10.3934/mbe.2024131.

[32] S. A. Atimbire, J. K. Appati, and E. Owusu, “Empirical exploration of whale optimisation algorithm for heart disease prediction,” Scientific Reports, vol. 14, no. 1, p. 4530, Feb. 2024, doi: https://doi.org/10.1038/s41598-024-54990-1.

[33] A. M. Qadri, M. S. A. Hashmi, A. Raza, S. A. J. Zaidi, and A. ur Rehman, “Heart failure survival prediction using novel transfer learning based probabilistic features,” PeerJ Computer Science, vol. 10, p. e1894, Mar. 2024, doi: https://doi.org/10.7717/peerj-cs.1894.

[34] Pavithra M.P and P. Maneesha, “HEART DISEASE PREDICTION USING BIO INSPIRED ALGORITHMS,” International Journal of Engineering Technology and Management Sciences, vol. 8, no. 1, pp. 125–129, Jan. 2024, doi: https://doi.org/10.46647/ijetms.2024.v08i01.015.

[35] S. H. Hasanah, “Application of Machine Learning for Heart Disease Classification Using Naive Bayes,” Jurnal Matematika MANTIK, vol. 8, no. 1, pp. 68–77, Jun. 2022, doi: https://doi.org/10.15642/mantik.2022.8.1.68-77.

[36] P. K. Misra, N. Kumar, A. Misra, and A. Khang, “Heart Disease Prediction Using Logistic Regression and Random Forest Classifier,” CRC Press eBooks, pp. 83–112, Aug. 2023, doi: https://doi.org/10.1201/9781003356189-6.

[37] N. Ukey, Z. Yang, B. Li, G. Zhang, Y. Hu, and W. Zhang, “Survey on Exact kNN Queries over High-Dimensional Data Space,” Sensors, vol. 23, no. 2, pp. 629–629, Jan. 2023, doi: https://doi.org/10.3390/s23020629.

[38] S. Adhikary and S. Banerjee, “Introduction to Distributed Nearest Hash: On Further Optimizing Cloud Based Distributed kNN Variant,” Procedia Computer Science, vol. 218, pp. 1571–1580, 2023, doi: https://doi.org/10.1016/j.procs.2023.01.135.

[39] X. Guo et al., “Critical role of climate factors for groundwater potential mapping in arid regions: Insights from random forest, XGBoost, and LightGBM algorithms,” Journal of Hydrology, vol. 621, pp. 129599–129599, Jun. 2023, doi:

https://doi.org/10.1016/j.jhydrol.2023.129599.
[40] S. Ghosal and A. Jain, "Depression and Suicide Risk Detection on Social Media using fastText Embedding and XGBoost Classifier," Procedia Computer Science, vol. 218, pp. 1631–1639, 2023, doi: https://doi.org/10.1016/j.procs.2023.01.141.
[41] A. Roy and S. Chakraborty, "Support vector machine in structural reliability analysis: A review," Reliability Engineering & System Safety, vol. 233, p. 109126, May 2023, doi: https://doi.org/10.1016/j.ress.2023.109126.
[42] A. Janosi, W. Steinbrunn, M. Pfisterer, and R. Detrano. "Heart Disease," UCI Machine Learning Repository, 1989. [Online]. Available: https://doi.org/10.24432/C52P4X.
[43] R. Detrano et al., "International application of a new probability algorithm for the diagnosis of coronary artery disease," The American Journal of Cardiology, vol. 64, no. 5, pp. 304–310, Aug. 1989, doi: https://doi.org/10.1016/0002-9149(89)90524-9.
[44] C. Bartley, Replication Data for: Cleveland Heart Disease. Harvard Dataverse, 2016. doi: 10.7910/DVN/QWXVNT.
[45] "Statlog (Heart)," UCI Machine Learning Repository, [online]. Available: https://doi.org/10.24432/C57303.
[46] R. S. Aftab (2024), "Prediction of Heart Attack", Mendeley Data, V2, doi: 10.17632/yrwd336rkz.2.

Authors' Profiles

Sultanul A. Hamim is an instructor at American International University-Bangladesh (AIUB) in the Department of Computer Science under the Faculty of Science and Technology. He is interested in research areas including ML, Deep learning, Data Mining, and ML Data Analytics, and also has experience with projects involving ML, data analysis, database architecture and cybersecurity who hopes to make a difference by using his abilities to facilitate well-informed decision-making and strengthen the safety of his employer's operations.

Rakin S. Aftab completed his bachelor's degree in computer science and engineering from the American International University-Bangladesh (AIUB). He is focused on ML, AI, and DS, with a particular interest in exploring neural networks and their applications in deep learning (DL), including convolutional neural networks (CNNs), artificial neural networks (ANNs), and deep hypercomplex neural networks. His work also encompasses networking within AI frameworks and the software development life cycle (SDLC), with the goal of optimizing AI-driven software development and enhancing technological solutions in AI and analytics. He is eager to apply his diverse skill set and innovative thinking to drive success and adapt to challenges. Committed to problem-solving, continuous learning, and making positive contributions, Rakin seeks opportunities for growth and collaboration. He can be contacted at rakinsadaftab@gmail.com.

Mir M. Ahmed has been passionate about computers from a young age, which naturally led him to pursue studies in computer science and technologies. He graduated from the American International University-Bangladesh (AIUB) with a Bachelor of Science degree in Computer Science and Engineering. His academic journey has been driven by a desire to learn new skills and technologies, marking him as a dedicated researcher focused on ML, AI, and DS. He is particularly interested in exploring neural networks and their applications in deep learning, including convolutional neural networks (CNNs) and artificial neural networks (ANNs).

Farzana Faiza attends the Faculty of Science and Technology (FST) at American International University-Bangladesh (AIUB), where she studies computer science in the Department of Computer Science. She enjoys exploring new frontiers in AI, computer vision, and deep learning (DL). The potential of deep learning to improve society is of great interest to her. She also strongly supports efforts to increase diversity and inclusion in the tech sector.

M. F. Mridha (Senior Member, IEEE) is currently working as an Associate Professor in the Department of Computer Science, at American International University-Bangladesh (AIUB). Before that, he worked as an Associate Professor and Chairman in the Department of Computer Science and Engineering, at Bangladesh University of Business and Technology (BUBT). He also worked as a CSE department faculty member at the University of Asia Pacific and as a graduate head from 2012 to 2019. He received his Ph.D. in AI/ML from Jahangirnagar University in 2017. His research experience in academia and industry has resulted in over 120 journal and conference publications. His research work contributed to the reputed Journal of Scientific Reports Nature, Knowledge-Based Systems, Artificial Intelligence Review, IEEE Access, Sensors, Cancers, Biology and Applied Sciences, etc. His research interests include artificial intelligence (AI), machine learning, deep learning, and natural language processing (NLP). For more than 10 (Ten) years, he has been with the master's and undergraduate students as a supervisor of their thesis work. His research interests include artificial intelligence (AI), machine learning, natural language processing (NLP), big data analysis, etc. He has served as a program committee member in several international conferences/workshops. He served as an Academic editor of several journals including PLOS ONE Journal. He has served as a reviewer of reputed journals like IEEE Transactions on Neural Networks, IEEE Access, Knowledge Base System, Expert System, Bioinformatics, Springer Nature, MDPI, etc., and conferences like ICCIT, HONET, ICIEV, IJCCI, ICAEE, ICCAIE, ICSIPA, SCORED, ISIEA, APACE, ICOS, ISCAIE, BEIAC, ISWTA, IC3e, ISWTA, CoAST, icIVPR, ICSCT, 3ICT, DATA21, etc.

How to cite this paper: Sultanul Arifeen Hamim, Rakin S. Aftab, M. Ahmed, Farzana Faiza, M. F. Mridha, "Advanced Heart Attack Prediction Using a Stacked Ensemble Machine Learning Model and Diverse Data Integration", International Journal of Intelligent Systems and Applications(IJISA), Vol.17, No.5, pp.49-67, 2025. DOI:10.5815/ijisa.2025.05.04