

Explainable Ensemble Learning Models for Early Detection of Heart Disease

Raed Hassan Laftah ^{1*}, Karim Hashim Kraidi Al-Saedi ²

^{1,2} Department of Computer Science, Al-Mustansiriyah University College of Science, Baghdad, Iraq

Email: ¹ raed.dirassat@gmail.com, ² dr.karim@uomustansiriyah.edu.iq

*Corresponding Author

Abstract—Coronary diseases (CVD) are a major global health concern, and timely, accurate diagnosis is crucial for effective treatment and management. As machine learning is, has steadily been on the improvement way, and it's there where we find the transformative potential for enhancing the diagnostic accuracy for their predictive accuracy using the Local Interpretable Model-agnostic Explanations technique to ensure the explainability of our models. With the advancement of machine learning, we aim to enhance diagnostic accuracy by developing a high-precision prediction tool for heart disease using various ML models. We utilized a Kaggle dataset to implement several ML models, including Random Forest, Gradient Boosting, CatBoost, K-Nearest Neighbor, Naive Bayes, Support Vector Machine, and AdaBoost, with appropriate data preprocessing. The soft voting ensemble method, combining various models, achieved a notable 98.54% accuracy and 99% precision, recall, and f1-score, with Random Forest, CatBoost, and the Voting Classifier outperforming others. These results indicate that our model is highly reliable and sets a new standard for CVD prediction. Future research should focus on validating this model with larger datasets and exploring deep learning approaches.

Keywords—Cardiovascular Disease Prediction; Machine Learning; Ensemble Learning; Soft Voting Classifier.

I. INTRODUCTION

Cardiovascular diseases (CVDs) are the leading cause of death globally, accounting for 17.9 million deaths annually [1]. This number is projected to rise to 23.6 million by 2030 [2], with a significant impact on low- to middle-income countries [3][4]. Contributory risk factors to HD are diverse and include, for example, unhealthy diets, tobacco use, and the high prevalence of obesity in developed countries [5][6]. Risk factors for heart disease include unhealthy diets, tobacco use, and obesity. In low- to middle-income countries, the adoption of similar lifestyles and rapid urbanization have led to a rising prevalence of chronic diseases [4]. Furthermore, CVDs impose a substantial economic burden, estimated at USD 3.7 trillion globally from 2010 to 2015 [7][8]. This reveals that the diseases are characterized by massive financial effects on an economy [24] [25]. They help in the fighting of HD through the availability and access to diagnostic technologies, for instance, electrocardiograms and CT scans, especially in resource-limited settings. This ultimately leads to late or failed diagnoses that further increase the health and economic burdens [4]. Early detection and accurate diagnosis are crucial for effective treatment and prevention of CVDs, especially in resource-limited settings. Machine learning (ML) and data mining techniques offer promising solutions by identifying patterns in large datasets,

which can lead to more accurate and early diagnosis of heart diseases [26] [27], which may be lifesaving, including for the affected individual and the health system burdening their finances. Data mining is the analytic process that helps in identifying necessary information from large datasets in different domains, among which medicine has its need in this area of activity. This is exactly what healthcare needs most: to extract the hidden pattern from vast repositories of patient data for making better diagnostic decisions [9]. Artificial intelligence, being relatively new and fast growing, comes hand in hand with ML to play the most crucial role in analyzing these extensive medical datasets [28] [29]. Its algorithm customizes the hard and non-linear interactions of all the other diverse health factors in such a way that the prediction accuracy is maximized and minimized in the difference between expected and real outcomes [10]. These have become very essential in the diagnosis, instrumental for the prediction of different diseases, including HD [11] [30] [31].

Previous years have seen an application of these techniques in predicting the probability of HD development [32] [33]. Factors predicting it include diabetes, high blood pressure, levels of cholesterol, and abnormal pulse rates [12]. However, often the medical data available are incomplete, which can impact the accuracy of HD prediction [34] [35]. It has been demonstrated in previous work that data mining has applications in the prediction of diseases [36][37][13], but accurate predictions are yet to be decided, especially with specific regard to the progression of diseases like HD [14].

Numerous studies have utilized machine learning models for the classification of heart diseases to this point [60, 61], yet many of these models have lacked sufficient explanatory power. Recent advancements in explainable machine learning [62, 63] have introduced a fresh research perspective focused on explicating the models deployed in the classification of heart diseases. In this area, a limited yet significant body of work has emerged. The most prevalent frameworks for elucidating machine learning models include LIME [62] and SHapley Additive exPlanations (SHAP) [65]. The researchers in [64] applied SHAP to elucidate the impact of coronary heart disease on mortality due to heart failure.

This study aims to enhance the prediction accuracy of heart disease (HD) by utilizing various machine learning (ML) models applied to a dataset from the UCI repository, which is publicly available on Kaggle. We will evaluate individual ML algorithms [11] and integrate ensemble learning techniques to identify the most effective predictive



model. Additionally, to ensure the explainability of our results, we incorporate the use of explainable artificial intelligence (XAI) with the Local Interpretable Model-Agnostic Explanations (LIME) technique, allowing for a clearer understanding of how our models make their predictions. This approach enhances the reliability of the predictions and makes the outcomes transparent and interpretable for healthcare professionals.

The research contributions of our article can be articulated as follows:

- This research significantly advances the predictive accuracy of heart disease diagnostics by employing a comprehensive ensemble of machine learning models. By integrating various algorithms, the study provides a robust framework that enhances the ability to predict heart disease more accurately than using individual models alone.
- The research innovates in the optimization of ensemble learning strategies to combine the strengths of individual ML models effectively. This leads to superior performance in terms of predictive accuracy, reliability, and applicability in real-world clinical settings.
- A key contribution of this study is the application of explainable artificial intelligence techniques, specifically the Local Interpretable Model-Agnostic Explanations (LIME). This approach allows for the elucidation of the decision-making processes of ML models, fostering trust and understanding among healthcare professionals by providing clear explanations of predictive outcomes.
- By improving the accuracy of HD predictions, this research will contribute significantly to medical diagnostics and preventive healthcare, providing valuable insights into the likelihood of CVDs based on patient attributes such as age, gender, and medical history.

II. RELATED WORK

Recent advances in data mining and machine learning (ML) have significantly impacted healthcare, particularly cardiology, by efficiently analyzing large volumes of medical data. This has opened new avenues for algorithm development and testing. Heart disease (HD) remains a major cause of mortality in developing countries, making it a key research area. Utilizing data mining and ML to identify risk factors and early symptoms can greatly enhance the early detection and prevention of HD, ultimately minimizing its public health impact.

Studies have demonstrated various ML techniques' efficacy in predicting cardiovascular diseases (CVDs). For instance, in [12], a hybrid model combining Random Forest (RF) and Linear Model (HRFLM) achieved an accuracy of 88.7% in CVD prediction. Another study [15] employed RF, Decision Tree (DT), Logistic Regression (LR), Naive Bayes (NB), and Support Vector Machine (SVM), with DT achieving the highest precision of 93.19%. In [16], a K-Nearest Neighbor (KNN) model achieved an accuracy of 90.8% using the Cleveland HD dataset. Study [15] showed that RF was used, along with other ML algorithms like the Decision Tree (DT), Logistic Regression (LR), NB, and

SVM, in predicting heart failure from the Cleveland Clinic Foundation data. These models have been cross-validated by 10 folds, which further establish that the DT algorithm leads in precision with 93.19%, followed closely by the SVM of 92.30%. This is indicative that the use of ML in the prediction of HD is very effective and hence forth will be a great tool in future research focusing on the prediction of HD. The researchers tried to develop a prediction model for CVD using ML approaches on the Cleveland HD dataset from the UCI repository, which contains 303 cases with 17 features [16]. The authors had applied to their testing various supervised classification methods with KNN had a maximum accuracy of 90.8%. Summarizing, the study has indicated that the KNN model seems to be the best, with the highest percentage accuracy of 90.8%. Research [4] highlighted the effectiveness of ensemble methods, with a stacked model combining KNN, RF, and SVM achieving 75.1% accuracy. Another study [17] evaluated eight data mining classifiers with different cross-validation methods, finding the Neural Network most precise for large datasets (71.82%) and RF for smaller datasets (89.01%).

The investigation in [17] delved into the role of data mining techniques and four cross-validation methods in HD prediction, employing eight data mining classifiers on datasets from Kaggle and the UCI machine-learning repository. The study aimed to discover the most accurate combination for HD prediction, employing accuracy, precision, recall, and F-measure as metrics. For large datasets (70,000 records), the Neural Network with the holdout method was found to be the most precise (71.82%), whereas for smaller datasets (303 records), the Repeated Random method with RF was most effective (89.01%). This research seeks to assist physicians in early HD detection to potentially improve productivity within business organizations.

In [18], ML techniques identified critical CVD risk factors in patients with metabolic-associated fatty liver disease, achieving accuracies of 85.11% and 79.17% for high-risk and low-risk groups, respectively. In [19], SVM led in precision (91.67%) for CVD diagnosis using electronic health records. Other studies [20][21][22] demonstrated the potential of ML models like AdaBoost and ensemble classifiers in achieving high accuracies for HD prediction.

A study [19] demonstrated the developed intelligent systems for accurate CVD diagnosis using electronic health records. The research evaluated four classification algorithms, following a preprocessing stage with data and feature extraction. The result of this model evaluated on metrics of accuracy, precision, recall, and F1-score with SVM led in precision with 91.67%.

A new method of increasing the prediction accuracy of the HD was established through algorithms on the basis of six algorithmic models. According to [20], the research established a new approach for This research takes the Cleveland dataset from the Cleveland database, and the data is from IEEE Dataport, wherein the model has been effectively tuned using GridSearchCV and using five-fold cross-validation. In the Cleveland dataset, AdaBoost has been most effective to the tune of 90%, whereas on the IEEE Dataport dataset, LR displayed 90.16% accuracy. The soft

voting ensemble classifier of novel nature that integrates these six algorithms presents improved accuracies of 93.44% and 95% for the Cleveland and IEEE Dataport datasets, respectively. This means better performance over the individual algorithms. A study of [21] therefore focuses on predictive algorithms for early detection of CVD and accurate prediction that form central points for improving patient diagnosis and treatment. The study took into account the data of the patients collected from Khyber Teaching Hospital and Lady Reading Hospital, based on which DT, RF, LR, NB, and SVM techniques were put to take up the study classification and prediction of CVD. The algorithm with the highest accuracy (85.01%) was the RF, which hints at a possible universal use of this model in healthcare for C CVD prediction. The investigation in [22] introduced a ML-based prediction model (MLbPM) for HD, incorporating data scaling, optimal data splitting ratios, best parameter selection, and algorithm application. Evaluated using the University of California Irvine HD dataset, the MLbPM, particularly with the LR algorithm, achieved remarkable accuracy of 96.7%. This model outperformed others in accuracy, presenting a significant advancement in HD prediction methodologies.

III. METHODOLOGY

This section outlines the methodology for developing a heart disease prediction model, including data preprocessing, model training, evaluation, and the use of ensemble learning and Explainable AI (XAI) techniques.

A. Proposed Heart Disease Model

The proposed heart disease prediction model integrates multiple machine learning algorithms to enhance prediction accuracy. This process involves data preprocessing, model training, and evaluation using ensemble learning techniques. A schematic representation of an HD Prediction model Data from Kaggle originating is shown on Fig. 1. The process starts with EDA (Exploratory Data Analysis) to get an insight of the dataset's most prominent patterns and to spot some possible outliers. Data Engineering includes various stages such as EDA, in which the data is normalized in order to prepare it for modeling.

The training and test sets are created from the predecoded dataset to help validate the models after. Multiple individual models are learnt to be trained, including RF, GB, CatBoost, KNN, NB, SVM, and AdaBoost. Every model leverages the data for training, after which accuracy is tested.

The following step is training of individual models, then an ensemble learning model using the soft voting method will be used. Each of the models mentioned above produces a prediction based on the specific model's weighting scheme. The final probability of prediction is then the average of the predicted probabilities. The case in point is that this type of methodology usually ensures a higher quality of performance, because it makes use of the strong sides of the given models to counteract their weak points. The performance of both the individual models and ensemble model is measured using a set of metrics.

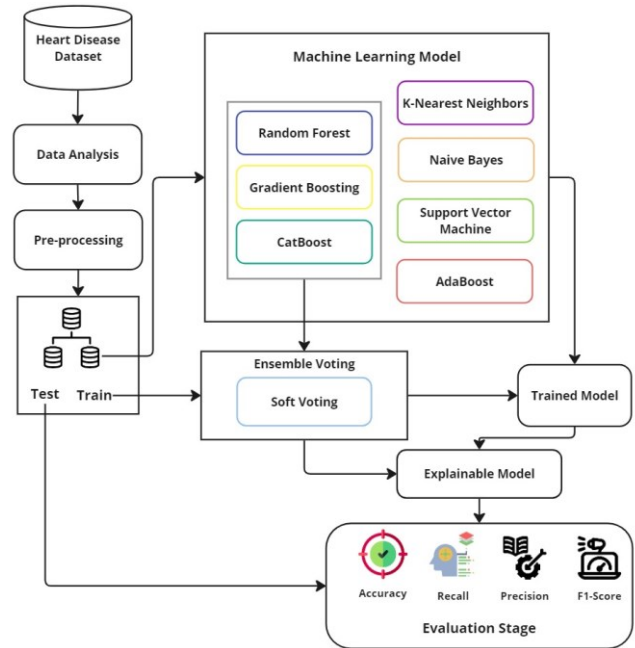


Fig. 1. Flowchart of proposed heart disease model

B. Dataset Description

The dataset used in this study is derived from the Cleveland database, available on the UCI repository and Kaggle. It includes 1025 instances with 14 attributes relevant to heart disease diagnosis, such as age, gender, chest pain type, resting blood pressure, and cholesterol levels. It also includes results from resting electrocardiographic tests, maximum heart rate achieved, exercise-induced angina, ST depression induced by exercise relative to rest, the slope of the peak exercise ST segment, and the number of major vessels colored by fluoroscopy, as well as the presence of thalassemia—a blood disorder affecting hemoglobin levels. The final attribute, 'target,' indicates the presence or absence of HD.

TABLE I. DATASET FEATURES DESCRIPTION

Attribute	Explanation
age	Individual's age.
sex	Individual's gender (1 = male, 0 = female).
cp	Chest pain type (0 = typical angina, 1 = atypical angina, 2 = non-anginal pain, 3 = asymptomatic).
trestbps	Resting blood pressure (mm Hg).
chol	Serum cholesterol (mg/dl).
fbs	Fasting blood sugar > 120 mg/dl (1 = true, 0 = false).
restecg	Resting electrocardiographic findings (0 = normal, 1 = ST-T wave abnormality, 2 = left ventricular hypertrophy).
thalach	Maximum heart rate achieved.
exang	Angina induced by exercise (1 = yes, 0 = no).
oldpeak	ST depression from exercise relative to rest.
slope	Slope of peak exercise ST segment (0 = upsloping, 1 = flat, 2 = downsloping).
ca	Major vessels (0-3) shown by fluoroscopy.
thal	Thalassemia type (0 = normal, 1 = fixed defect, 2 = reversible defect).
target	Heart disease diagnosis (0 = no, 1 = yes).

C. Data Analysis

Our EDA process was meticulously designed to uncover patterns and detect outliers, ensuring a robust foundation for

the subsequent modeling steps [39]. We employed statistical techniques to examine the distributions of continuous variables such as age, resting blood pressure, and serum cholesterol. These variables displayed approximately normal distributions but with potential outliers, which were identified using the Interquartile Range (IQR) method. Outliers were then treated by capping at the 1st and 99th percentiles to mitigate their impact on model accuracy.

For categorical variables [38], analysis was conducted using frequency bar charts, which highlighted demographic disparities, such as a higher prevalence of heart disease among males and varying types of chest pain, with typical angina appearing less frequently. These insights led to the development of feature engineering strategies tailored to enhance model interpretability and performance. We converted categorical variables into dummy variables to facilitate their use in machine learning models and created interaction terms between age and serum cholesterol to explore potential synergistic effects on heart disease risk.

Additionally, missing values were handled through multiple imputation techniques, informed by the variable type and distribution, to preserve data integrity and prevent bias in model training. For instance, missing continuous variables were imputed with the median value, and categorical variables with the mode, ensuring consistency and maintaining the statistical properties of the dataset.

These detailed EDA steps have provided a clear, data-driven basis for our preprocessing and modeling decisions, enhancing the transparency and reproducibility of our study, as illustrated in Fig. 2.

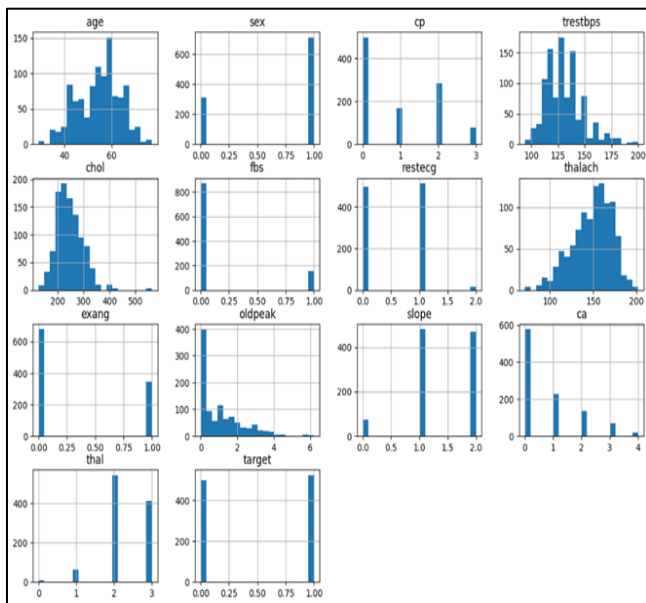


Fig. 2. Distribution of clinical attributes in heart disease

D. Data Pre-processing

In the preprocessing phase of our HD prediction study, we employed the Standard Scaler method to normalize the feature set, ensuring that each variable contributes equally to the model's predictive power by bringing them to a common scale without distorting differences in the ranges of values. This step is critical, as the varying scales of raw data can bias

the algorithms, affecting the accuracy of predictions [40] [41]. The dataset was normalized using the Standard Scaler method to ensure equal contribution of each variable to the model. The data was then split into training (80%) and testing (20%) sets to validate the models' performance [42][43]. This split is designed to validate the performance of the models [44][45], allowing us to train the algorithms on a large portion of the data and then test their predictive capabilities on unseen data, thereby ensuring the robustness and generalizability of our models [46].

E. Individual Machine Learning Approach

A variety of algorithms are tailored to enhance the accuracy and efficiency of predictive modeling [82]. Notable among these are Random Forest (RF), Gradient Boosting (GB), CatBoost, K-Nearest Neighbors (KNN), Naive Bayes (NB), Support Vector Machines (SVM), and AdaBoost. Each of these models brings a unique set of strengths and methodologies to tackle classification problems of heart disease prediction.

- The Random Forest (RF) classifier [66], a key tool for heart disease diagnosis, constructs multiple decision trees during training and selects the most common outcome from these trees to make predictions. This ensemble method helps minimize overfitting better than a single decision tree by using diverse subsets of the data for each tree, enhancing the model's generalization capabilities. By setting a fixed `random_state`, we ensure reproducibility, crucial for model validation and comparison. RF effectively handles the bias-variance tradeoff, making it adept at using a variety of features—from demographic to medical—to predict heart disease, while also highlighting the most influential factors in its predictions [74] [75].
- Support Vector Machines (SVM) excel in classification by identifying the optimal hyperplane that maximizes the margin between two classes [67-69]. In heart disease diagnosis, SVM classifies conditions based on a clear gap between positive and negative cases. It operates by projecting data points into a high-dimensional space, ensuring a maximum margin between the closest examples of each class. This configuration leads to confident classification of new instances. Using the scikit-learn library's SVC function with a fixed `random_state` ensures reproducibility. Depending on data complexity, various kernels like linear, polynomial, or radial basis function (RBF) can be used to handle linear or more intricate relationships among features.
- The K-Nearest Neighbors (KNN) algorithm classifies data based on the proximity of similar cases within the feature space, making it ideal for tasks like heart disease diagnosis [76]. KNN determines the classification by examining the 'K' nearest labeled data points to a new patient's data, assigning a binary output where '0' indicates no heart disease and '1' indicates its presence, based on the majority class among these neighbors. Implemented via the `KNeighborsClassifier()` from the scikit-learn library, it calculates distances (typically Euclidean) between points, selecting the 'K' closest to decide the class by majority. Selecting an optimal 'K'

value is critical, as smaller 'K' can be noise-sensitive, while larger 'K' increases computation and may reduce precision.

- The Naive Bayes (NB) classifier, utilized for heart disease prediction [77] [78], operates on Bayes' theorem with an assumption of feature independence, making it suitable for high-dimensional datasets. Specifically, the GaussianNB() function is used, which presumes that continuous features like age and cholesterol follow a Gaussian distribution. This classifier estimates the probability of heart disease presence ('1') or absence ('0') by treating each feature's effect independently, simplifying calculations and enabling efficient estimation of one-dimensional distributions. Gaussian Naive Bayes proves effective even when features deviate from normal distribution, by computing and comparing the posterior probabilities for each class to make predictions.
- Gradient Boosting (GB) is a powerful ensemble technique that builds models sequentially, with each new model focusing on correcting errors made by previous ones. It combines multiple weak prediction models, typically decision trees, to create a strong predictive model. In heart disease prediction, GB progressively refines its predictions by increasing the weight of incorrectly classified instances, ensuring subsequent models focus more on difficult cases. This approach helps in effectively handling the complex patterns and interactions within medical data, resulting in a robust model capable of accurately predicting the presence or absence of heart disease with high reliability [79] [80]. The final prediction is made by aggregating the outcomes from all the individual trees, providing a balanced consideration of various risk factors.
- CatBoost is an advanced gradient boosting algorithm that excels in handling categorical data directly with minimal preprocessing required [81]. It is designed to effectively prevent overfitting, making it particularly suitable for complex datasets like those used in heart disease prediction. CatBoost builds on decision trees by systematically selecting the best splits during the learning process, optimizing predictions through ordered boosting. This approach ensures high accuracy by addressing potential biases in feature selection. For heart disease classification, CatBoost leverages its capability to process diverse medical datasets, integrating numerous health indicators to predict the likelihood of heart disease with enhanced precision and stability.
- AdaBoost, short for Adaptive Boosting, is a robust ensemble technique that combines multiple weak classifiers into a strong one by iteratively adjusting the weights of incorrectly classified instances [82]. In the context of heart disease prediction, AdaBoost starts with a base classifier, often a decision tree, and focuses subsequent classifiers on the errors of the previous ones. Each classifier's influence in the final decision is weighted based on its accuracy, with more accurate classifiers having more impact. This adaptive process enhances the model's ability to distinguish between patients with and without heart disease, effectively

improving prediction accuracy by emphasizing harder-to-classify instances in successive training rounds.

F. Ensemble Learning Approach

Ensemble learning leverages the strengths of multiple predictive models to enhance the overall accuracy and robustness of predictions [47] [70-72](Fig. 3). In our approach, we employed a soft voting ensemble method [51], where the decision is derived from the average of the predicted probabilities of selected models, specifically Random Forest (RF), Gradient Boosting (GB), and CatBoost. Each of these models brings distinct advantages to the ensemble that RF excels in reducing overfitting while maintaining high accuracy by averaging multiple deep decision trees that individually consider a random subset of features and data points. GB incrementally builds an ensemble by correcting the residuals of earlier trees, optimized for performance and less prone to overfitting, making it highly effective for complex classification problems. CatBoost is particularly effective with categorical data and reduces the typical gradient boosting issues of overfitting through systematic ordered boosting and advanced handling of categorical variables.

The soft voting mechanism used in our ensemble learning method [51] does not simply pick the most frequent prediction. Instead, it averages the probabilities assigned to each class by the selected models, thereby integrating the strengths of each. This probabilistic approach allows the ensemble to benefit from the diverse strengths of the included models, such as RF's handling of random features, GB's sequential improvement on mistakes, and CatBoost's efficiency with categorical inputs. Such an ensemble not only balances the individual weaknesses of the models but also enhances prediction reliability and generalization to new data. This ensemble framework, therefore, acts as a sophisticated, balanced, and robust tool for predicting heart disease, potentially outperforming any single classifier operating independently [52] [53].



Fig. 3. Ensemble learning architecture

G. Explainable AI (XAI)

Explainable AI (XAI) plays a pivotal role in demystifying the predictions made by AI models, particularly in sensitive fields like healthcare. Our methodology utilizes Local Interpretable Model-Agnostic Explanations (LIME) to enhance the interpretability of our predictive models [54] [55][56] [59]. This technique opposes the "black box" approach by elucidating how input features affect predictions, thus providing transparency and building trust with healthcare professionals.

In this study, LIME was specifically applied to instances where the model predicted the presence or absence of heart disease. By breaking down the contribution of each feature—such as age, cholesterol levels, and blood pressure—LIME highlighted which factors were most influential in predicting

patient outcomes [23]. For example, in a case where the model predicted a high risk of heart disease, LIME could reveal that elevated cholesterol levels and age were the dominant factors influencing this prediction, allowing clinicians to understand the basis of the AI's decision.

Furthermore, we provided case studies in which LIME's explanations assisted doctors in making informed treatment decisions. These case studies demonstrated scenarios where the model's output, validated by LIME, aligned with clinical expectations or provided new insights into risk factors that were not initially obvious to the treating physicians.

By integrating LIME, we aim not only to improve model transparency but also to empower healthcare professionals to leverage AI-assisted diagnostics confidently. This approach ensures that the foundation of AI recommendations is clear, enhancing the decision-making process in clinical settings and fostering deeper trust in AI-driven medical diagnostics [57] [58].

This enhanced focus on XAI through practical applications of LIME establishes its value in heart disease prediction, offering clear illustrations of how model predictions are derived and how they can be effectively communicated and utilized in medical practice.

IV. RESULT ANALYSIS AND DISCUSSION

A. Performance Matrices

In our study, the performance of the heart disease (HD) prediction models is evaluated using a comprehensive suite of metrics, each offering unique insights into the model's effectiveness for a binary classification problem [73]. Accuracy measures the overall proportion of correct predictions (both true positives and true negatives) relative to the total dataset, indicating the model's general effectiveness. Precision quantifies the accuracy of the positive predictions, reflecting the ratio of true positives to the total predicted as positive; this is crucial in medical diagnostics to avoid false alarms. Recall, or sensitivity, measures the proportion of actual positives correctly identified, highlighting the model's capability to detect all relevant cases, which is vital for ensuring no condition goes unnoticed. The F1-score provides a balance between precision and recall by calculating their harmonic mean, offering a single metric that considers both the reliability of positive predictions and the model's ability to identify all positives. This is particularly important in clinical settings where both avoiding false positives and capturing all true positives are critical.

However, these metrics have limitations. For instance, accuracy can be misleading in imbalanced datasets where one class significantly outnumbers the other. Precision and recall may not fully capture the model's performance nuances, as they do not consider the true negatives and false negatives, respectively. Hence, while these metrics provide a solid foundation for evaluating model performance, they must be interpreted in conjunction with a thorough understanding of their mathematical implications and the clinical context of HD prediction. By elucidating these metrics, we aim to ensure that the results are interpreted correctly and that the model's applicability to clinical diagnosis is clearly

understood, fostering trust and reliability in AI-driven diagnostics.

$$ACC = \frac{TN + TP}{TP + TN + FP + FN} \quad (1)$$

$$PRE = \frac{TP}{FP + TP} \quad (2)$$

$$REC = \frac{TP}{FN + TP} \quad (3)$$

$$F1 - S = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

B. Evaluation of Individual Machine Learning methods

In our study, we harnessed the computational power of Google Colab, executing our models on a standard computing system equipped with a 4800-H processor and 16 GB of RAM. The HD dataset under scrutiny contained 1025 instances, each described by 14 attributes, providing a rich matrix of data for analysis. A comprehensive evaluation of a spectrum of classifiers was conducted to discern the presence of HD, with the results detailed in Table II of our study.

The performance of each classifier was measured by its accuracy, with tree-based algorithms emerging as the most effective. The RF and CatBoost models both achieved an exemplary accuracy score of 0.985366, indicating their superior predictive capabilities in this context. GB also performed impressively, with an accuracy of 0.931707, underscoring its potential utility in medical diagnostics.

Further down the accuracy spectrum, the SVM registered a substantial accuracy of 0.887805, while AdaBoost followed closely with 0.878049, both reflecting strong predictive performances. The KNN algorithm, with an accuracy of 0.834146. Lastly, the NB classifier, while offering a baseline accuracy of 0.800, confirmed the challenging nature of the task and the sophistication required in the predictive modeling of HD. These results not only reflect the effectiveness of tree-based models in handling the complexity of medical data but also highlight the potential of ensemble techniques in enhancing prediction accuracy in the field of HD diagnosis.

TABLE II. ACCURACY OF DIFFERENT MACHINE LEARNING MODELS FOR HEALTH DISEASE PREDICTION

Model	Accuracy
RF	0.985366
CatBoost	0.985366
GB	0.931707
SVM	0.887805
AdaBoost	0.878049
K-NN	0.834146
NB	0.800

The results from our model, as elucidated through LIME. The interpretability of our AI model is significantly enhanced through the use of Local Interpretable Model-Agnostic Explanations (LIME), as illustrated in Fig. 3 and Fig. 4. In Fig. 4, LIME analyzes a prediction where there is a high likelihood (87% probability) of a heart attack. This high-risk prediction is primarily influenced by critical cardiac

indicators including a lower slope of the peak exercise ST segment (≤ -0.62) and a reduced maximum heart rate (thalach ≤ -0.74). Additionally, factors such as age (≥ 0.72) and exercise-induced angina (exang ≤ 1.40) contribute to the risk assessment, emphasizing their significance in the model's predictive logic.

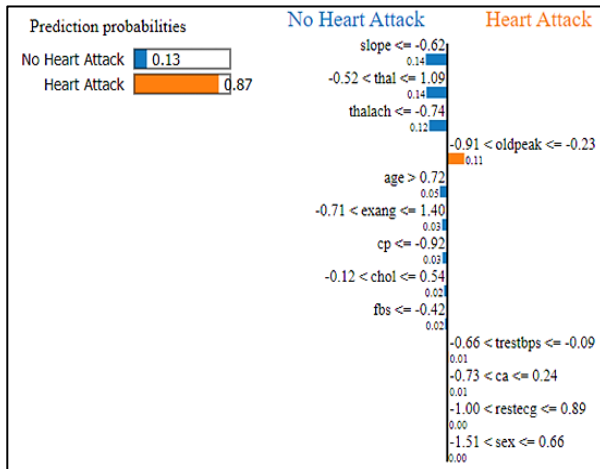


Fig. 4. Interpretable machine learning predictions for heart attack risk using LIME

Fig. 5 contrasts this by depicting a scenario with a low probability (1% probability) of a heart attack, where the model's confidence in predicting the absence of heart disease is particularly driven by an oldpeak value (> 0.49) and low cholesterol levels (chol ≤ -0.68). The decision tree paths shown elucidate the reasoning behind the model's predictions, with the branch widths visually representing the impact of each feature on the decision-making process, thereby offering clear insights into the factors mitigating against the risk of heart disease.

These visualizations validate the model's accuracy and furnish healthcare professionals with understandable, actionable insights from AI-driven diagnostics, essential for informed clinical decision-making. By employing LIME, we offer a detailed interpretation of the model's predictions at an individual level, ensuring transparency and comprehensibility in how decisions are derived. Such clear explanations are vital for clinicians who depend on these predictions to guide patient care, enabling them to make well-informed decisions with confidence.

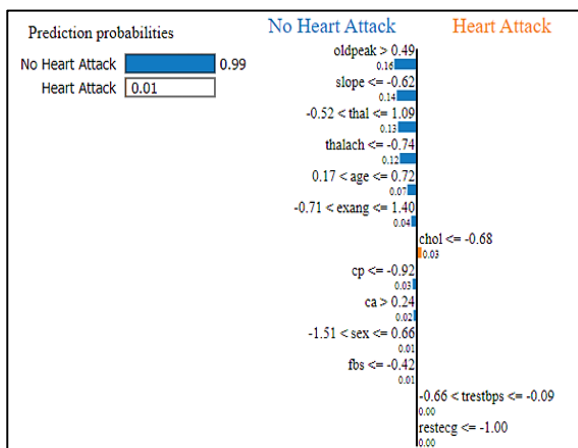


Fig. 5. Predictive analysis of low heart attack risk using LIME

C. Evaluation of Individual Machine Learning methods

The performance of the ensemble learning model, specifically the soft Voting Classifier, has exhibited exceptional results in the task of HD prediction. With an accuracy of 0.9854, the model demonstrates a near-perfect classification rate. The sensitivity, or true positive rate, stands at 0.9709, indicating that the model correctly identifies 97.09% of actual HD cases. Moreover, the model achieved a specificity of 1.00, which means it perfectly identifies individuals without the disease, as there were no false positives.

The confusion matrix, as shown in Fig. 6, further solidifies the model's robustness, where out of 205 cases, 102 were correctly identified as no heart attack (true negatives), and 100 were correctly identified as heart attack cases (true positives), with only 3 instances of false negatives, where the model incorrectly predicted no heart attack.

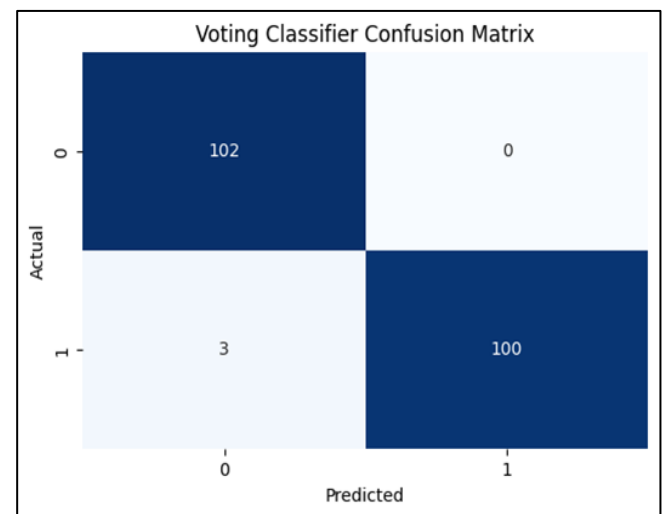


Fig. 6. Confusion matrix of ensemble learning model

The classification report in Fig. 7 provides detailed insights into the model's precision, recall, and F1-score for both classes. For the 'No Heart Attack' class, the model achieved a precision of 0.97 and a recall of 1.00, culminating in an F1-score of 0.99. The 'Heart Attack' class saw a perfect precision of 1.00, a recall of 0.97, and an F1-score of 0.99. The support column indicates the number of actual occurrences of each class in the dataset, which stands at 102 for 'No Heart Attack' and 103 for 'Heart Attack.' These metrics, combined with the support figures, attest to the model's high reliability and validity in classifying HD.

	precision	recall	f1-score	support
No Heart Attack	0.97	1.00	0.99	102
Heart Attack	1.00	0.97	0.99	103
accuracy			0.99	205
macro avg	0.99	0.99	0.99	205
weighted avg	0.99	0.99	0.99	205

Fig. 7. Classification report of ensemble learning model

D. Comparison Results

I In the context of contemporary research on heart disease (HD) prediction, our proposed ensemble learning model,

employing a soft Voting Classifier, has demonstrated outstanding accuracy, achieving 98.54%. This performance not only surpasses notable benchmarks set by recent studies but also highlights the model's superior predictive capabilities. For instance, Boukhatem et al. [19] reported an accuracy of 91.67% using an SVM model, Khan et al. [21] achieved 85.01% with the RF algorithm, and Bizimana et al. [22] reached 96.7% using LR. These comparisons are detailed in Table III, underscoring the enhanced accuracy of our approach.

While our model exhibits high accuracy, it is essential to discuss its generalizability and external validity, particularly in diverse populations and healthcare settings. The high accuracy achieved in a controlled study setting might not directly translate to varied clinical environments where demographic, genetic, and lifestyle factors can significantly influence model performance. To address these concerns, further validation studies are necessary, particularly in geographically and demographically diverse populations to ensure that the model maintains its reliability across different groups. Additionally, we acknowledge potential sources of bias, such as sampling bias and measurement inconsistencies, which could affect the model's performance in real-world settings.

Addressing these factors and conducting extensive external validation will enhance the robustness of our findings and ensure that our model can be effectively scaled and applied in various clinical scenarios, thus supporting its broader adoption for HD prediction.

TABLE III. COMPARISON OF HEART DISEASE PREDICTION ACCURACIES

Reference	Method	Results (Accuracy)
Boukhatem et al. [19]	SVM	91.67%
Khan et al. [21]	RF	85.01%
Bizimana et al. [22]	LR	96.7%
Our Proposed Model	Soft Voting Classifier	98.54%

V. CONCLUSION

Heart disease (HD) remains a significant public health challenge, requiring innovative approaches for early detection and treatment. It is influenced by a complex interplay of genetic, lifestyle, and environmental factors. Accurate prediction and diagnosis are crucial for improving patient outcomes, necessitating advanced analytical tools. In this context, we have developed a state-of-the-art ensemble learning model that integrates multiple well-proven algorithms like Random Forest (RF), Gradient Boosting (GB), and CatBoost through a strategic soft voting mechanism. This approach enhances the model's predictive accuracy, achieving an impressive 98.54% accuracy and leverages the strengths of each constituent classifier to improve overall reliability.

Importantly, we recognize the critical need for model interpretability and transparency in healthcare applications. Decisions based on our model can have profound implications for patient outcomes; therefore, ensuring that the decision-making processes are transparent and interpretable

to both clinicians and patients is essential. We have incorporated Explainable AI (XAI) techniques, particularly Local Interpretable Model-Agnostic Explanations (LIME), to make our model's predictions understandable and actionable. This enhances trust and facilitates the adoption of the model in clinical settings.

However, the conclusion of our study also acknowledges the necessity of balancing technical performance with the broader implications for healthcare delivery. Future improvements will focus on enhancing the model's generalization capabilities through the use of larger and more diverse datasets, potentially integrating deep learning models to capture more complex patterns and subtle correlations that simpler models might miss. Moreover, ethical and regulatory considerations are paramount as predictive models become more integrated into clinical decision-making. We are committed to addressing ethical concerns such as patient privacy, data security, algorithmic bias, and informed consent. Compliance with regulatory frameworks like HIPAA and GDPR is also a priority to ensure patient safety and data protection.

REFERENCES

- [1] M. D. Seckeler and T. R. Hoke, "The worldwide epidemiology of acute rheumatic fever and rheumatic heart disease," *Clinical epidemiology*, pp. 67–84, 2011.
- [2] K. Saxena et al., "Efficient heart disease prediction system," *Procedia Computer Science*, vol. 85, pp. 962–969, 2016.
- [3] C. S. Dangare and S. S. Apte, "Improved study of heart disease prediction system using data mining classification techniques," *International Journal of Computer Applications*, vol. 47, no. 10, pp. 44–48, 2012.
- [4] V. Shorewala, "Early detection of coronary heart disease using ensemble techniques," *Informatics in Medicine Unlocked*, vol. 26, p. 100655, 2021.
- [5] H. N. Murthy and M. Meenakshi, "Dimensionality reduction using neuro-genetic approach for early prediction of coronary heart disease," in *International conference on circuits, communication, control and computing*, pp. 329–332, 2014.
- [6] E. J. Benjamin et al., "Heart disease and stroke statistics—2019 update: a report from the american heart association," *Circulation*, vol. 139, no. 10, pp. e56–e528, 2019.
- [7] D. Mozaffarian et al., "Heart disease and stroke statistics—2016 update: a report from the american heart association," *circulation*, vol. 133, no. 4, pp. e38–e360, 2016.
- [8] J. Maiga et al., "Comparison of machine learning models in prediction of cardiovascular disease using health record data," in *2019 international conference on informatics, multimedia, cyber and information system (ICIMCIS)*, pp. 45–48, 2019.
- [9] J. Soni et al., "Predictive data mining for medical diagnosis: An overview of heart disease prediction," *International Journal of Computer Applications*, vol. 17, no. 8, pp. 43–48, 2011.
- [10] S. F. Weng, J. Reys, J. Kai, J. M. Garibaldi, and N. Qureshi, "Can machine-learning improve cardiovascular risk prediction using routine clinical data?," *PloS one*, vol. 12, no. 4, p. e0174944, 2017.
- [11] V. Ramalingam, A. Dandapath, and M. K. Raja, "Heart disease prediction using machine learning techniques: a survey," *International Journal of Engineering & Technology*, vol. 7, no. 2.8, pp. 684–687, 2018.
- [12] S. Mohan, C. Thirumalai, and G. Srivastava, "Effective heart disease prediction using hybrid machine learning techniques," *IEEE access*, vol. 7, pp. 81 542–81 554, 2019.
- [13] M. Nabeel, S. Majeed, M. J. Awan, H. Muslih-ud Din, M. Wasique, and R. Nasir, "Review on effective disease prediction through data mining techniques," *International Journal on Electrical Engineering & Informatics*, vol. 13, no. 3, 2021.

- [14] T. Ramesh, U. K. Lilhore, M. Poongodi, S. Simaiya, A. Kaur, and M. Hamdi, "Predictive analysis of heart diseases with machine learning approaches," *Malaysian Journal of Computer Science*, pp. 132–148, 2022.
- [15] F. S. Alotaibi, "Implementation of machine learning model to predict heart failure disease," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 6, 2019.
- [16] D. Shah, S. Patel, and S. K. Bharti, "Heart disease prediction using machine learning techniques," *SN Computer Science*, vol. 1, pp. 1–6, 2020.
- [17] A. I. Shima Ouf, "A proposed paradigm for intelligent heart disease prediction system using data mining techniques," *Journal of Southwest Jiaotong University*, vol. 56, no. 4, 2021.
- [18] K. Drozd'z *et al.*, "Risk factors for cardiovascular disease in patients with metabolic-associated fatty liver disease: a machine learning approach," *Cardiovascular Diabetology*, vol. 21, no. 1, p. 240, 2022.
- [19] C. Boukhatem, H. Y. Youssef, and A. B. Nassif, "Heart disease prediction using machine learning," in *2022 Advances in Science and Engineering Technology International Conferences (ASET)*, pp. 1–6, 2022.
- [20] N. Chandrasekhar and S. Peddakrishna, "Enhancing heart disease prediction accuracy through machine learning techniques and optimization," *Processes*, vol. 11, no. 4, p. 1210, 2023.
- [21] A. Khan *et al.*, "A novel study on machine learning algorithm-based cardiovascular disease prediction," *Health & Social Care in the Community*, vol. 2023, 2023.
- [22] P. C. Bizimana *et al.*, "An effective machine learning-based model for an early heart disease prediction," *BioMed Research International*, vol. 2023, 2023.
- [23] M. T. Ribeiro, S. Singh, and C. Guestrin, "Anchors: High-precision model-agnostic explanations," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 32, no. 1, 2018.
- [24] M. S. Eichenbaum, S. Rebelo, and M. Trabandt, "The macroeconomics of epidemics," *The Review of Financial Studies*, vol. 34, no. 11, pp. 5149–5187, 2021.
- [25] Y. Ismahene, "Infectious diseases, trade, and economic growth: a panel analysis of developed and developing countries," *Journal of the Knowledge Economy*, vol. 13, no. 3, pp. 2547–2583, 2022.
- [26] T. Chatzinikolaou, E. Vogiatzi, A. Kousis, and C. Tjortjis, "Smart healthcare support using data mining and machine learning," in *IoT and WSN based Smart Cities: A Machine Learning Perspective*, pp. 27–48, 2022.
- [27] A. S. Albahri *et al.*, "Role of biological data mining and machine learning techniques in detecting and diagnosing the novel coronavirus (COVID-19): a systematic review," *Journal of medical systems*, vol. 44, pp. 1–11, 2020.
- [28] M. Rana and M. Bhushan, "Machine learning and deep learning approach for medical image analysis: diagnosis to detection," *Multimedia Tools and Applications*, vol. 82, no. 17, pp. 26731–26769, 2023.
- [29] M. Rahman, "Artificial Intelligence (AI) and Machine Learning (ML)," in *Medical Imaging Informatics towards Diagnostic Decision Making*, p. 238, 2023.
- [30] P. C. Bizimana, Z. Zhang, M. Asim, A. A. A. El-Latif, and M. Hammad, "Learning-based techniques for heart disease prediction: a survey of models and performance metrics," *Multimedia Tools and Applications*, pp. 1–55, 2023.
- [31] M. Badawy, N. Ramadan, and H. A. Hefny, "Healthcare predictive analytics using machine learning and deep learning techniques: a survey," *Journal of Electrical Systems and Information Technology*, vol. 10, no. 1, p. 40, 2023.
- [32] J. Ouwerkerk *et al.*, "Machine learning in Huntington's disease: exploring the Enroll-HD dataset for prognosis and driving capability prediction," *Orphanet Journal of Rare Diseases*, vol. 18, no. 1, p. 218, 2023.
- [33] S. D. Prakoso, A. E. Permanasari, and A. R. Pratama, "Heart Disease Prediction Using Machine Learning: A Systematic Literature Review," in *2023 10th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, pp. 155–159, 2023.
- [34] M. Fan, X. Peng, X. Niu, T. Cui, and Q. He, "Missing data imputation, prediction, and feature selection in diagnosis of vaginal prolapse," *BMC Medical Research Methodology*, vol. 23, no. 1, p. 259, 2023.
- [35] L. Hu, X. Cheng, C. Wen, and Y. Ren, "Medical prediction from missing data with max-minus negative regularized dropout," *Frontiers in Neuroscience*, vol. 17, p. 1221970, 2023.
- [36] D. L. Olson and Ö. M. Araz, "Applications of Predictive Data Mining in Healthcare," in *Data Mining and Analytics in Healthcare Management: Applications and Tools*, pp. 105–116, 2023.
- [37] S. G. Kanakaraddi, K. C. Gull, J. Bali, A. K. Chikaraddi, and S. Giraddi, "Disease prediction using data mining and machine learning techniques," *Advanced Prognostic Predictive Modelling in Healthcare Data Analytics*, pp. 71–92, 2021.
- [38] A. Agresti, *Categorical data analysis*. John Wiley & Sons, vol. 792, 2012.
- [39] B. Shahbaba and B. Shahbaba, "Statistical Inference for the Relationship Between Two Variables," *Biostatistics with R: An Introduction to Statistics Through Biological Data*, pp. 193–219, 2012.
- [40] L. H. Nazer *et al.*, "Bias in artificial intelligence algorithms and recommendations for mitigation," *PLOS digital health*, vol. 2, no. 6, p. e0000278, 2023.
- [41] L. Belenguer, "AI bias: exploring discriminatory algorithmic decision-making models and the application of possible machine-centric solutions adapted from the pharmaceutical industry," *AI and Ethics*, vol. 2, no. 4, pp. 771–787, 2022.
- [42] H. Liu and M. Cocca, "Semi-random partitioning of data into training and test sets in granular computing context," *Granular Computing*, vol. 2, pp. 357–386, 2017.
- [43] K. Korjus, M. N. Hebart, and R. Vicente, "An efficient data partitioning to improve classification performance while keeping parameters interpretable," *PloS one*, vol. 11, no. 8, p. e0161788, 2016.
- [44] T. P. Debray *et al.*, "A guide to systematic review and meta-analysis of prediction model performance," *BMJ*, vol. 356, 2017.
- [45] R. D. Riley *et al.*, "Evaluation of clinical prediction models (part 2): how to undertake an external validation study," *BMJ*, vol. 384, 2024.
- [46] L. Kumar, C. Anitha, V. N. Ghodke, N. Nithya, V. A. Drave, and F. Azmath, "Deep Learning Based Healthcare Method for Effective Heart Disease Prediction," *EAI Endorsed Transactions on Pervasive Health and Technology*, vol. 9, pp. 1–6, 2023.
- [47] A. Ahmed Mohammed and Z. Aung, "Ensemble learning approach for probabilistic forecasting of solar power generation," *Energies*, vol. 9, no. 12, p. 1017, 2016.
- [48] R. R. Sarra, A. M. Dinar, M. A. Mohammed, M. K. A. Ghani, and M. A. Albahar, "A robust framework for data generative and heart disease prediction based on efficient deep learning models," *Diagnostics*, vol. 12, no. 12, p. 2899, 2022.
- [49] D. Hassan, H. I. Hussein, and M. M. Hassan, "Heart disease prediction based on pre-trained deep neural networks combined with principal component analysis," *Biomedical signal processing and control*, vol. 79, p. 104019, 2023.
- [50] A. A. Mohammed, W. Yaqub, and Z. Aung, "Probabilistic forecasting of solar power: An ensemble learning approach," in *Intelligent Decision Technologies: Proceedings of the 7th KES International Conference on Intelligent Decision Technologies (KES-IDT 2015)*, pp. 449–458, 2015.
- [51] M. U. Salur and I. Aydın, "A soft voting ensemble learning-based approach for multimodal sentiment analysis," *Neural Computing and Applications*, vol. 34, no. 21, pp. 18391–18406, 2022.
- [52] G. Wang, X. Yang, and X. Zhu, "Single Classifier Selection for Ensemble Learning," in *Advanced Data Mining and Applications: 12th International Conference*, pp. 312–328, 2016.
- [53] S. Y. Hera, M. Amjad, and M. K. Saba, "Improving heart disease prediction using multi-tier ensemble model," *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 11, no. 1, p. 41, 2022.
- [54] S. Praveen and K. Joshi, "Explainable Artificial Intelligence in Health Care: How XAI Improves User Trust in High-Risk Decisions," in *Explainable Edge AI: A Futuristic Computing Perspective*, pp. 89–99, 2022.
- [55] C. Manresa-Yee, M. F. Roig-Maimó, S. Ramis, and R. Mas-Sansó, "Advances in XAI: Explanation interfaces in healthcare," in *Handbook*

- of *Artificial Intelligence in Healthcare: Vol 2: Practicalities and Prospects*, pp. 357-369, 2021.
- [56] J. Gupta and K. R. Seeja, "A Comparative Study and Systematic Analysis of XAI Models and their Applications in Healthcare," *Archives of Computational Methods in Engineering*, pp. 1-26, 2024.
- [57] C. Metta *et al.*, "Improving trust and confidence in medical skin lesion diagnosis through explainable deep learning," *International Journal of Data Science and Analytics*, pp. 1-13, 2023.
- [58] R. Lukyanenko, W. Maass, and V. C. Storey, "Trust in artificial intelligence: From a Foundational Trust Framework to emerging research opportunities," *Electronic Markets*, vol. 32, no. 4, pp. 1993-2020, 2022.
- [59] Henriques, A., Parola, H., Gonçalves, R., & Rodrigues, M. (2024, March). Integrating Explainable AI: Breakthroughs in Medical Diagnosis and Surgery. In World Conference on Information Systems and Technologies (pp. 254-272). Cham: Springer Nature Switzerland.
- [60] G. N. Ahmad, H. Fatima, S. Ullah, and A. S. Saidi, "Efficient medical diagnosis of human heart diseases using machine learning techniques with and without GridSearchCV," *IEEE Access*, vol. 10, pp. 80151-80173, 2022.
- [61] J. P. Li, A. U. Haq, S. U. Din, J. Khan, A. Khan, and A. Saboor, "Heart disease identification method using machine learning classification in e-healthcare," *IEEE Access*, vol. 8, pp. 107562-107582.
- [62] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should i trust you?' Explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 1135-1144, 2016.
- [63] P. Besse, C. Castets-Renard, A. Garivier, and J. M. Loubes, "Can everyday ai be ethical? machine learning algorithm fairness," *Machine Learning Algorithm Fairness (May 20, 2018)*, vol. 6, no. 3, 2019.
- [64] K. Wang *et al.*, "Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP," *Computers in biology and medicine*, vol. 137, p. 104813, 2021.
- [65] T. A. Assegie, "Evaluation of local interpretable model-agnostic explanation and shapley additive explanation for chronic heart disease detection," *Proc Eng Technol Innov*, vol. 23, pp. 48-59, 2023.
- [66] R. Goel, "Heart disease prediction using various algorithms of machine learning," in *Proceedings of the International Conference on Innovative Computing & Communication (ICICC)*, 2021.
- [67] A. Kilic, "Artificial intelligence and machine learning in cardiovascular health care," *The Annals of thoracic surgery*, vol. 109, no. 5, pp. 1323-1329, 2020.
- [68] S. Uddin, A. Khan, M. E. Hossain, and M. A. Moni, "Comparing different supervised machine learning algorithms for disease prediction," *BMC medical informatics and decision making*, vol. 19, no. 1, pp. 1-16, 2019.
- [69] P. Aparna and K. M. Sharma, "Detection of A Fib and its Classification using SVM," in *2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, pp. 116-120, 2020.
- [70] M. A. Naji, S. El Filali, M. Bouhlal, E. H. Benlahmar, R. A. Abdelouahid, and O. Debauche, "Breast cancer prediction and diagnosis through a new approach based on majority voting ensemble classifier," *Procedia Computer Science*, vol. 191, pp. 481-486, 2021.
- [71] S. Kumari, D. Kumar, and M. Mittal, "An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier," *International Journal of Cognitive Computing in Engineering*, vol. 2, pp. 40-46, 2021.
- [72] N. Kumar and D. Kumar, "Machine learning based heart disease diagnosis using non-invasive methods: A review," in *Journal of Physics: Conference Series*, vol. 1950, no. 1, p. 012081, 2021.
- [73] Y. Muhammad, M. Tahir, M. Hayat, and K. T. Chong, "Early and accurate detection and diagnosis of heart disease using intelligent computational model," *Scientific reports*, vol. 10, no. 1, p. 19747, 2020.
- [74] P. K. Misra, N. Kumar, A. Misra, and A. Khang, "Heart disease prediction using logistic regression and random forest classifier," in *Data-Centric AI Solutions and Emerging Technologies in the Healthcare Ecosystem*, pp. 83-112, 2023.
- [75] A. K. Chaudhuri, S. Das, and A. Ray, "An Improved Random Forest Model for Detecting Heart Disease," in *Data-Centric AI Solutions and Emerging Technologies in the Healthcare Ecosystem*, pp. 143-164, 2024.
- [76] N. Samal, M. Kaur, R. K. Singhal, and J. S. P. Singh, "Heart Disease Analysis Research Using K-Nearest Neighbor: a Review," *EasyChair Preprint*, 2023.
- [77] M. Maydanchi, "Comparative Study of decision tree, adaboost, random forest, Naïve Bayes, KNN, and perceptron for heart disease prediction," in *SoutheastCon 2023*, pp. 204-208, 2023.
- [78] T. S. Patel, D. P. Patel, M. Sanyal, and P. S. Shrivastav, "Prediction of heart disease and survivability using support vector machine and Naive Bayes algorithm," *bioRxiv*, 2023.
- [79] T. O. Omotehinwa, D. O. Oyewola, and E. G. Moun, "Optimizing the light gradient-boosting machine algorithm for an efficient early detection of coronary heart disease," *Informatics and Health*, vol. 1, no. 2, pp. 70-81, 2024.
- [80] R. Suhendra *et al.*, "Cardiovascular Disease Prediction Using Gradient Boosting Classifier," *Infolitika Journal of Data Science*, vol. 1, no. 2, pp. 56-62, 2023.
- [81] F. Ahmed, M. Saleem, Z. Rajpoot, and A. Noor, "Intelligent Heart Disease Prediction Using CatBoost Empowered with XAI," *International Journal of Computational and Innovative Sciences*, vol. 2, no. 4, pp. 8-13, 2023.
- [82] H. Singh, T. Gupta, and J. Sidhu, "Prediction of heart disease using machine learning techniques," in *2021 Sixth International Conference on Image Information Processing (ICIIP)*, vol. 6, pp. 164-169, 2021.