Machine Learning for Heart Disease Diagnosis

Mandava Nidhish
Department of Computer Science
Vellore Institute of Technology
Chennai, India

Dogiparthi Aasrith
Department of Computer Science
Vellore Institute of Technology
Chennai, India

Y. Kalyan
Department of Computer Science
Vellore Institute of Technology
Chennai, India

***Abstract*—This has various machine learning algorithms to predict heart disease using a comprehensive dataset. The algorithm visualization includes Logistic Regression, Gaussian Naive Bayes, K-Nearest Neighbours (KNN), and an ensemble method known as the Voting Classifier. The dataset comprises diverse patient attributes, allowing for a holistic analysis to develop accurate predictive models. The process begins with data pre-processing and exploratory data analysis to understand the dataset's characteristics. Missing values are handled, and statistical measures are computed to gain insights into the data distribution. The target variable, indicating the presence or absence of heart disease, is examined for its distribution across the dataset. Following data preparation, the machine learning models are trained and evaluated. Logistic Regression, a linear classification model, is trained to predict heart disease based on patient attributes. Evaluation metrics such as accuracy, precision, recall, F1-score, and confusion matrix analysis are utilized to assess the model's performance on both training and test datasets. Additionally, Gaussian Naive Bayes, a probabilistic classification algorithm, is trained and evaluated. K-Nearest Neighbours (KNN) is implemented as a non-parametric classification method, leveraging the similarity between data points to make predictions. Furthermore, an ensemble approach is adopted using the Voting Classifier, which combines predictions from multiple base estimators, including Logistic Regression, Naive Bayes, and Random Forest. This ensemble method enhances predictive accuracy and robustness by aggregating individual model predictions. The performance of each model is evaluated using various metrics, providing a comprehensive understanding of their effectiveness in predicting heart disease. Furthermore, visualization techniques such as confusion matrices and receiver operating characteristic (ROC) curves are employed to illustrate model performance visually.**

Keywords—Machine Learning, Logistic Regression, KNN, Naïve Bayes, Voting Classification, Ensembled Learning.

I. INTRODUCTION

Cardiovascular diseases (CVDs) are relentless adversaries to global public health, exacting an immense toll in terms of mortality, morbidity, and healthcare expenditure. These multifaceted ailments encompass a spectrum of conditions affecting the heart and blood vessels, including coronary artery disease, stroke, and heart failure. Despite advances in medical science and preventive measures, CVDs persist as a leading cause of death worldwide. The imperative to mitigate this burden underscores the importance of proactive strategies for risk assessment and early intervention. In this study, we embark on a journey through the realm of predictive analytics, employing machine learning algorithms to discern patterns within complex datasets and forecast the likelihood of heart disease occurrence.

Machine learning has emerged as a potent ally in the battle against cardiovascular diseases, offering computational tools to unravel intricate relationships between clinical variables and disease outcomes. Logistic regression, Naive Bayes, K-Nearest Neighbors (KNN), and ensemble methods stand out as formidable contenders in the domain of classification algorithms, each with its unique strengths and applications. Logistic regression, a stalwart of statistical modeling, excels in modeling binary outcomes and estimating the probability of disease occurrence based on predictor variables. Naive Bayes, grounded in Bayes' theorem and the assumption of feature independence, boasts simplicity and computational efficiency, making it a popular choice for classification tasks, especially in domains with limited data availability. KNN, a non-parametric algorithm, leverages proximity-based reasoning to assign class labels to data points, rendering it adept at handling complex data structures and nonlinear relationships.

In this project, we undertake a comprehensive evaluation of these machine learning algorithms, scrutinizing their performance across a diverse array of performance metrics. By harnessing the power of real-world datasets containing a rich tapestry of clinical and demographic information, we aim to extract the essential insights for heart disease risk assessment and management. Through rigorous experimentation and model refinement, we endeavour to unveil the latent predictive potential embedded within these algorithms, illuminating the path towards precision medicine and personalized care. By elucidating the strengths and limitations of each approach, we seek to empower healthcare professionals with robust decision support tools, facilitating early detection, risk stratification, and targeted interventions to curtail the scourge of cardiovascular diseases.

In summary, this study embodies a concerted effort to harness the transformative potential of machine learning in the realm of cardiovascular health. By leveraging the predictive process of logistic regression, Naive Bayes, KNN, and ensemble methods, we aspire to transcend traditional paradigms of risk assessment and catalyse a paradigm shift towards proactive, data-driven healthcare. Through collaborative endeavours between clinicians, data scientists, and healthcare stakeholders, we envisage a future where predictive analytics serve as analytics tools are essential partners in our efforts to combat cardiovascular diseases, paving the way for a healthier and stronger global community.

II. Literature review

The literature survey encompasses several studies focusing on the prediction and detection of cardiac disease using machine learning (ML) algorithms [2]. Each study emphasizes the importance of early detection and accurate prediction of heart disease. Various techniques such as feature selection, algorithm comparison, and model performance evaluation are explored [3]. The research highlights the significance of leveraging ML methods to forecast cardiac disorders, with a focus on improving model performance through the selection of informative features and the evaluation of different algorithms [4]. Key metrics such as accuracy, precision, recall, and F1-score are utilized to assess the performance of the predictive models, with notable achievements in achieving high accuracy scores like 99.02% accuracy [5].This encompasses several studies focused on utilizing machine learning (ML) algorithms for predicting and detecting heart disease, underlining the importance of early detection for effective treatment and risk reduction. Python emerges as a favoured programming language due to its accessibility and robust ML libraries. Across the studies, various ML algorithms such as Logistic Regression, k Nearest Neighbors (KNN), Decision Trees, Random Forest, Naïve Bayes, Support Vector Machine (SVM), and k-nearest neighbour algorithm are investigated [7,8]. Findings reveal a range of accuracies: Random Forest demonstrates strong performance with accuracies ranging from 83.52% to 88.16% [6,10], Logistic Regression and Naïve Bayes also showcase notable accuracies, while specific figures vary among the studies. These collective research efforts underscore the potential of ML in advancing heart disease prediction and detection, offering valuable insights into algorithm performance and its implications for healthcare practice.

Among the Logistic regression, K-Nearest neighbour, decision tree, random forest and XG Boost Logistic regression is the best model for the prediction of cardio vascular disease with accuracy of 90.16 [11]. And also, the hybrid model of random forest and decision forest has shown the more accuracy than the individual model with the accuracy of 88.7% [13]. The study highlights the ML methods used for classification, providing the comparative analysis between them [14]. The implementation of the models takes use of vast range of feature integrations in addition to numerous techniques of categorization that are already widely known to general public [15].

Some of them have applied the genetic algorithm on the data set to optimize the features which results in increase of accuracy to its maximum [17]. The three major's danger signs for heart disease are smoking, high blood pressure and cholesterol, and 47% of all US citizens have at least one of these risk factors. This machine learning algorithm helps in prediction of the risk and help the people. [18,20]. As we seen the random forest, logistic regression are the best models for the prediction of heart disease but support vector machine is also the best algorithm for predicting the disease [19].

Not only the above-mentioned algorithms they have also talked about the gradient boosting classifier and MLP classifier are the models that are investigated for the better accuracy [21]. In some analysis they have also concluded that ANN algorithm, hybrid algorithm of gradient boosting and decision tree is best for the heath related services [22,24]. 1 in 14 people live with heart and circulatory diseases and also 200 million people are also estimated to live with coronary heart disease. By creating a machine learning model, we can help them by being aware of these diseases before getting effected [25]. In this survey, we have analysed the performance of various heart disease prediction techniques, namely ABC-SVM, ANFIS, SVM-ANN, SVM-SSVM, Genetic Algorithm, Neural Network Ensemble, FNN, Majority Vote Based Ensemble Classifier etc [26,27]. Many of them also had dome using the deep learning algorithms like RNN, CNN which concluded that CNN algorithm is the beat algorithm for measuring the disease or predicting [28]. As mentioned above the algorithms like logistic regression, naïve bay's and etc of them have given the accuracy near to 85% -92% [29]. At last the cleaved database of UCI has got the best accuracy using the decision trees that is 83.87% [30].

A "black box" is a model that is so complicated that it is difficult, if not impossible, for a person to understand. Especially in the field of medicine, where many choices have profound consequences for patients' lives, the inability to understand the reasoning behind prediction models might erode their credibility [40,42]. For the FHS dataset the SVM classifier has the got the highest accuracy of 56.53% [43]. All the above algorithms have been used before for the sake of heart failure [47,48]. So that they can be prevented from getting the heart failure and follow the precautions for the diseases they have been identified using the, machine learning and deep learning algorithms [49,50].

Table-1: *Accuracies and performance metrics from the research papers*

Research Paper	Models	Performance Metrics
[1]	Naive Bayes	Accuracy:74%
[3]	Logistic Regression	Accuracy:66.01%,Precision:66.15%,Recall:66.13%,F1-Score:65.96%
	KNN	Accuracy:76.45%,Precision:80.48%,Recall:78.37%,F1-Score:78.15%
	Random Forest	Accuracy:70.01%,Precision:74.06%,Recall:70.01%,F1-Score:68.24%
[4]	Logistic Regression	Accuracy:83.1%,Precision:84.4%%,Recall:86%,F1-Score:85.1%
[5]	KNN	Accuracy:82.9%,Precision:77.5%,Recall:90.9%,F1-Score:83.6%
	SVM	Accuracy:86.3%,Precision:80.8%,Recall:93.9%,F1-Score:86.8%
[6]	Logistic Regression	Accuracy:84.21%
	KNN	Accuracy:73.68%
	Random Forest	Accuracy:88.16%
[7]	Logistic Regression	Accuracy:81.3%,Precision:81.4%,Recall:81.1%,F1-Score:80.9%
	Naive Bayes	Accuracy:81.4%,Precision:81.3%,Recall:81.3%,F1-Score:81.1%
	KNN	Accuracy:73.5%,Precision:73.5%,Recall:72.9%,F1-Score:72.8%
[9]	Logistic Regression	Accuracy:80.2%,Precision:82.01%,AUC:85.5%,F1-Score:82.01%
	Naive Bayes	Accuracy:82.4%,Precision:86.9%,AUC:86.01%,F1-Score:83.3%
	KNN	Accuracy:72.5%,Precision:75.5%,AUC:76.81%,F1-Score:74.7%
	Random Forest	Accuracy:83.5%,Precision:88.8%,AUC:88.2%,F1-Score:84.2%
[10]	Naive Bayes	Accuracy:82.1%
	KNN	Accuracy:76.5%
	Random Forest	Accuracy:69.3%
[20]	Logistic Regression	Accuracy:72.4%
	Random Forest	Accuracy:69.4%
[21]	Logistic Regression	Accuracy:89.67%
	Naive Bayes	Accuracy:87.61%
	KNN	Accuracy:85.8%
[22]	Naive Bayes	Accuracy:82.1%,Precision:83.1%,Sensitivity:84.1%,F1-Score:85%
	Random Forest	Accuracy:78%,Precision:77%,Sensitivity:85%,F1-Score:80%
[23]	Logistic Regression	Accuracy:60%,Precision:62%,Recall:61%
	Naive Bayes	Accuracy:60%,Precision:72%,Recall:36%
	KNN	Accuracy:79%,Precision:74%,Recall:93%
	Random Forest	Accuracy:96%,Precision:93%,Recall:100%

	Voting Classifier	Accuracy:76%,Precision:79%,Recall:%

III. PROPOSED METHODOLOGY

A. Dataset Description

The dataset contains various patient health attributes including age, gender, cholesterol level, blood pressure and other clinical factors. The target variable indicates the presence or absence of heart disease. The dataset is split into features (X) and target variables (Y) and then further split into training and test sets for model training and evaluation.

Table-2: *Dataset Description*

S. No	Type	Count
1	Yes [1]	165
2	No [0]	138

B. Data Pre-processing

The dataset is loaded and examined for missing values. Descriptive statistics are calculated to understand the distribution of features. In the data pre-processing stage, we meticulously prepared the heart disease dataset to ensure its quality and integrity for machine learning. This involved meticulously cleaning the data by removing irrelevant or duplicate patient records. We addressed missing values using appropriate techniques, such as imputation or deletion, depending on the nature of the missing data and its impact on the analysis. Additionally, exploratory data analysis (EDA) was conducted to understand the distribution of features like age, blood pressure, cholesterol levels, etc. This analysis helped identify potential relationships between features and inform decisions regarding feature selection and model training for accurate heart disease prognosis.

C. Machine Learning Algorithms

To predict heart disease risk, machine learning dives into a toolbox of algorithms. Logistic regression calculates the odds of disease based on factors like age and cholesterol. Decision trees create a branching roadmap of questions about a patient's health, leading to a final classification of healthy or at-risk. Random forests build a team of decision trees, each analyzing different aspects, for a more robust prediction. Support Vector Machines draw a clear boundary between healthy and at-risk individuals based on their characteristics. Gradient boosting builds on prior models, focusing on their mistakes, to continuously improve accuracy. This diverse arsenal of algorithms empowers us to create powerful tools for identifying individuals who might be susceptible to heart disease.

Gaussian Naive Bayes: This probabilistic classifier proves valuable in heart disease prediction. It assumes independence between features, such as age and blood pressure. Based on this assumption, it efficiently calculates the probability of a patient having heart disease based on individual features and their distributions. This simplicity and computational efficiency make Naive Bayes a strong contender for initial screenings or resource-limited environments, allowing for rapid risk assessments.

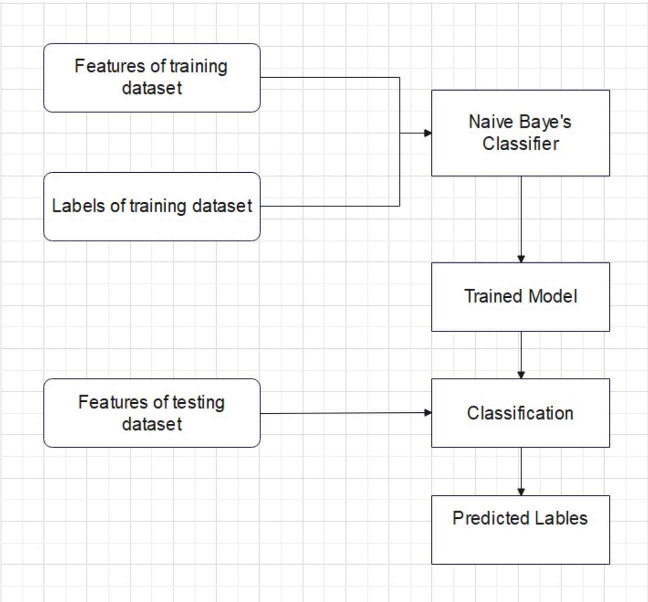


Figure-1: *Gaussian Naïve Bayes Architecture*

Logistic Regression: This workhorse algorithm serves as a cornerstone for heart disease prediction. Unlike Random Forest, which excels with complex data, Logistic Regression thrives in its simplicity and interpretability. It establishes a mathematical relationship between various factors like age, blood pressure, and cholesterol levels, and the likelihood of developing heart disease. This allows healthcare professionals to not only predict risk but also understand which factors contribute most significantly. This interpretability makes Logistic Regression a valuable tool for initial risk assessment and guiding further medical investigation.

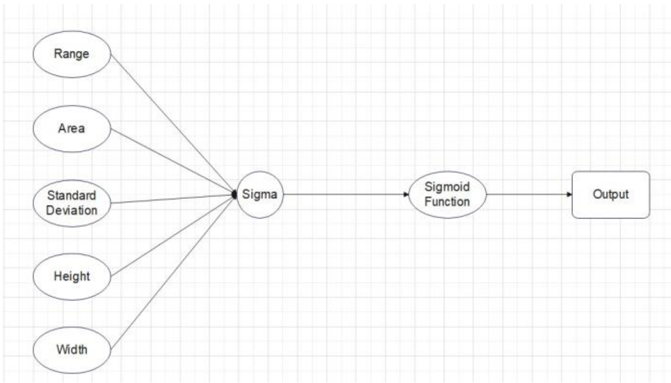


Figure-2: *Logistic Regression Architecture*

Voting Classifier: In heart disease prediction, a Voting Classifier acts as a conductor, harmonizing the predictions from multiple machine learning models. It doesn't create its own predictions, but instead, gathers predictions from various algorithms (like Logistic Regression and Random Forest) and makes the final call based on a voting scheme (e.g., majority vote). This approach leverages the strengths of different models, potentially leading to more robust and accurate predictions for heart disease risk compared to relying on a single model.

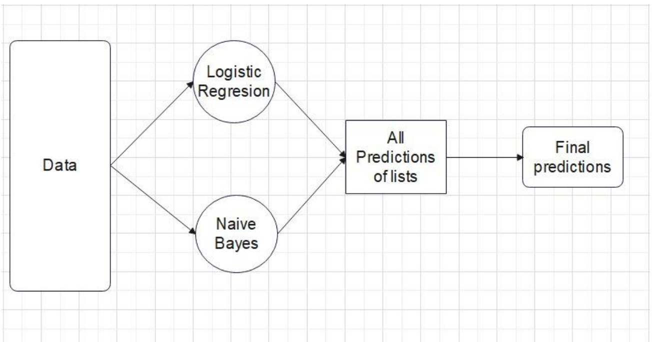


Figure-3: *Voting Classifier Architecture*

K-Nearest Neighbours: The K-Nearest Neighbours (KNN) algorithm for heart disease prediction. It begins by importing necessary libraries from scikit-learn, initializes a KNN model, and trains it on the training data. Predictions are then made on both the training and test datasets to assess model performance, and accuracy scores are calculated. Additionally, precision, recall, and F1-score metrics are computed to evaluate the model's effectiveness further. A confusion matrix is generated to visualize the classification results, offering insights into the model's performance across different classes.

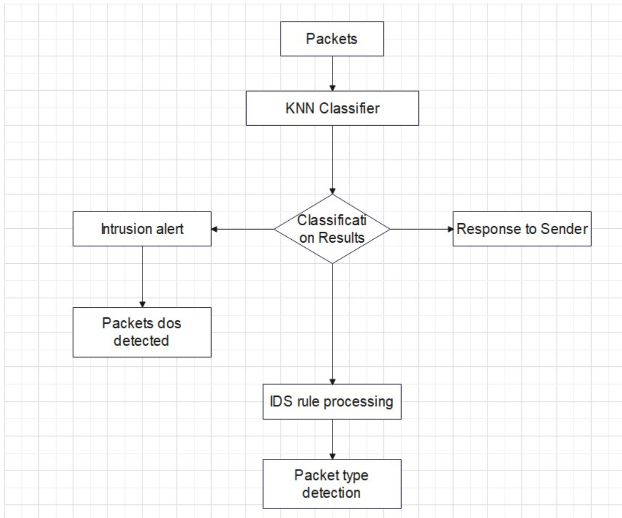
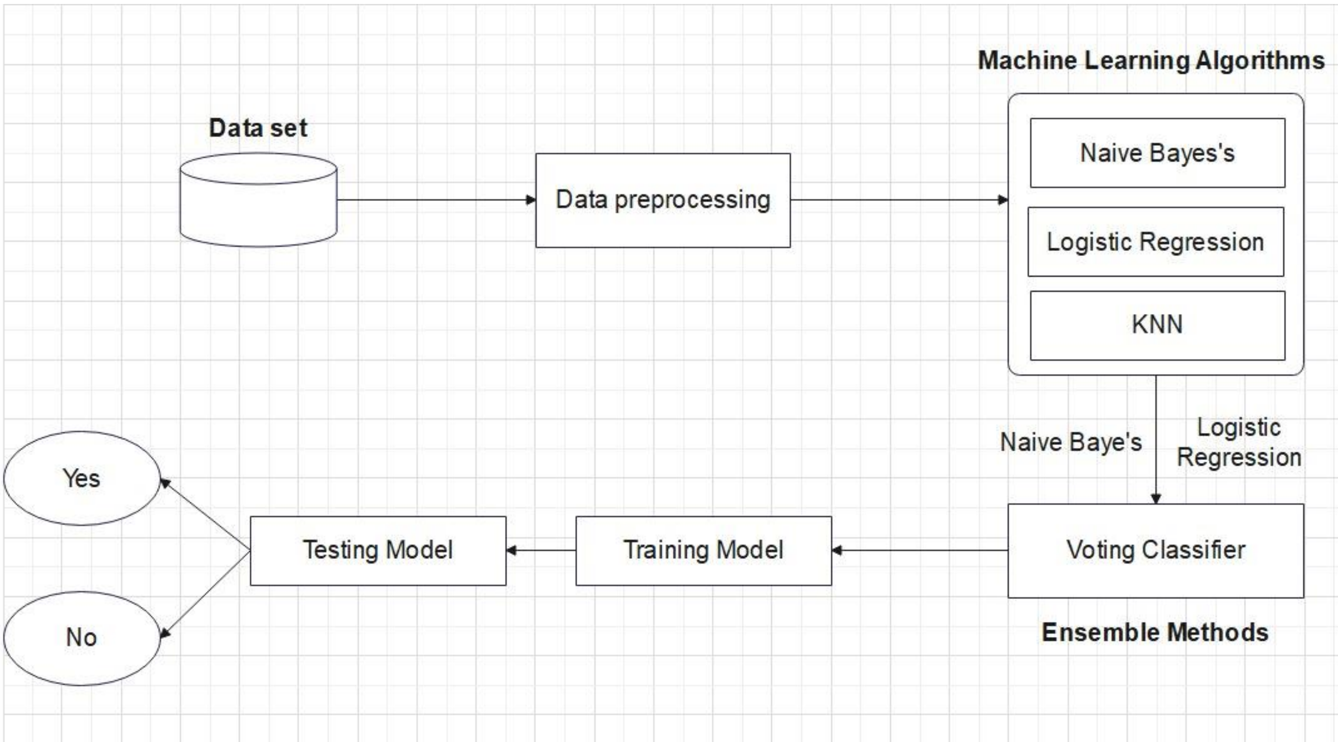


Figure-4: *KNN Architecture*

IV. Complete Architecture Diagram



V. Findings and Discussions

The confusion matrix, encapsulates the classification outcomes of the logistic regression model, showcasing its ability to discern between positive and negative instances. The matrix's diagonal elements represent correct predictions, where instances of no heart disease (true negatives) and instances of heart disease (true positives) were accurately identified. However, off-diagonal elements denote misclassifications: instances of heart disease were wrongly classified as negative (false negatives), and instances were incorrectly identified as positive (false positives). This matrix serves as a comprehensive snapshot of the model's performance, detailing its strengths and areas for refinement. Meanwhile, the Receiver Operating Characteristic (ROC) curve provides a graphical representation of the model's trade-off between true positive rate and false positive rate across varying threshold values.

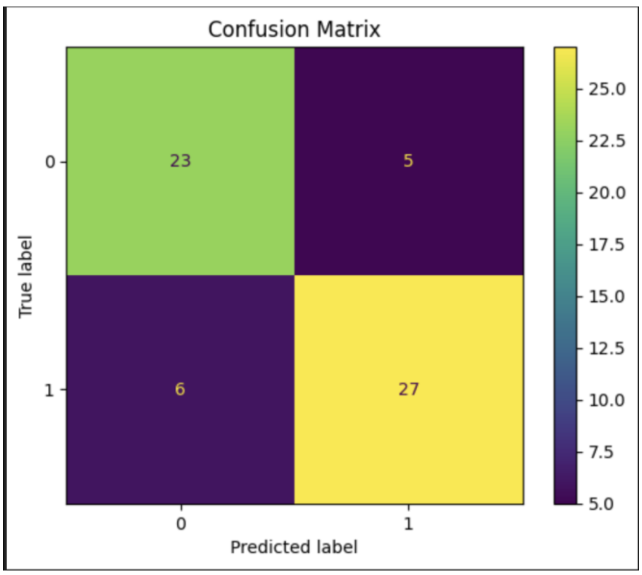


Figure-4: *Confusion Matrix for Logistic Regression*

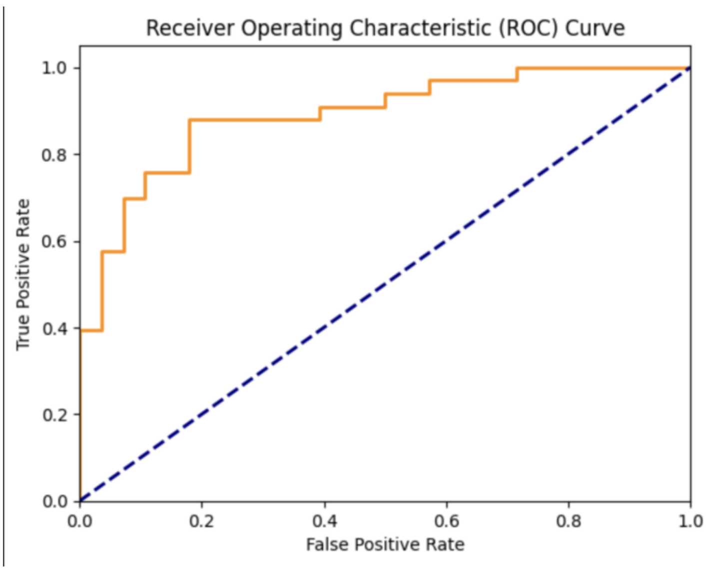


Figure-5: *ROC Curve for Logistic Regression*

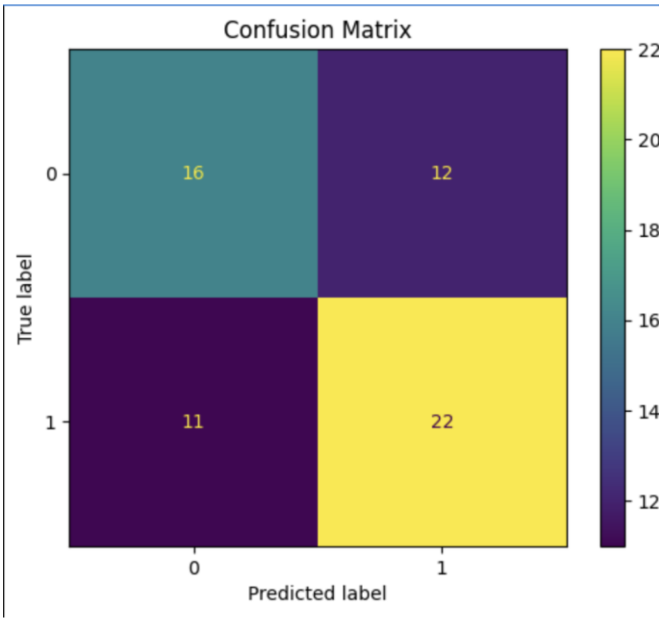


Figure-6: *Confusion Matrix for KNN*

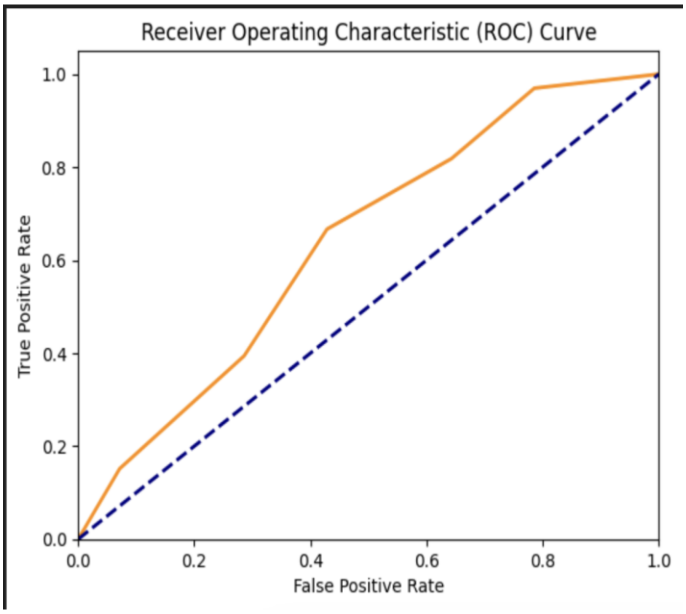


Figure-7: *ROC Curve for KNN*

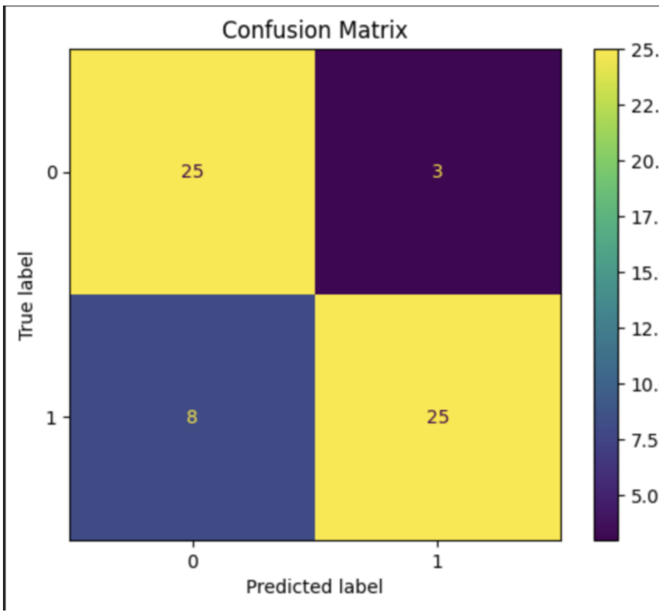


Figure-8: *Confusion Matrix for Naïve Bayes*

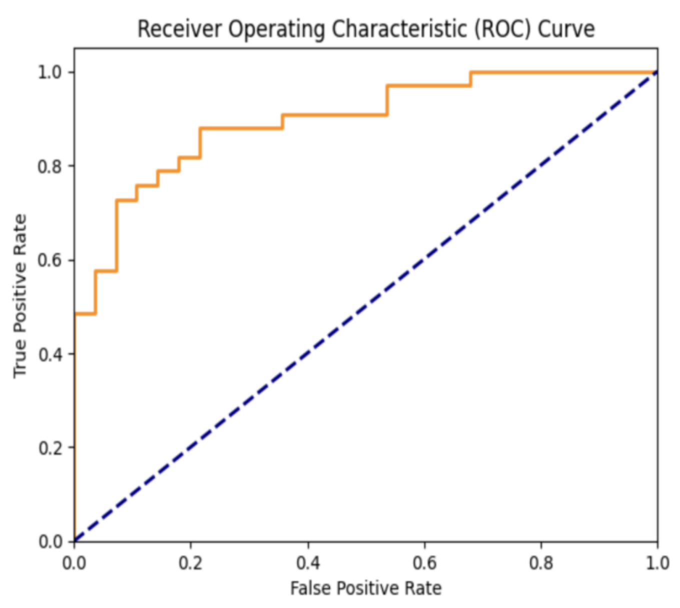


Figure-9: *ROC Curve for Naïve Bayes*

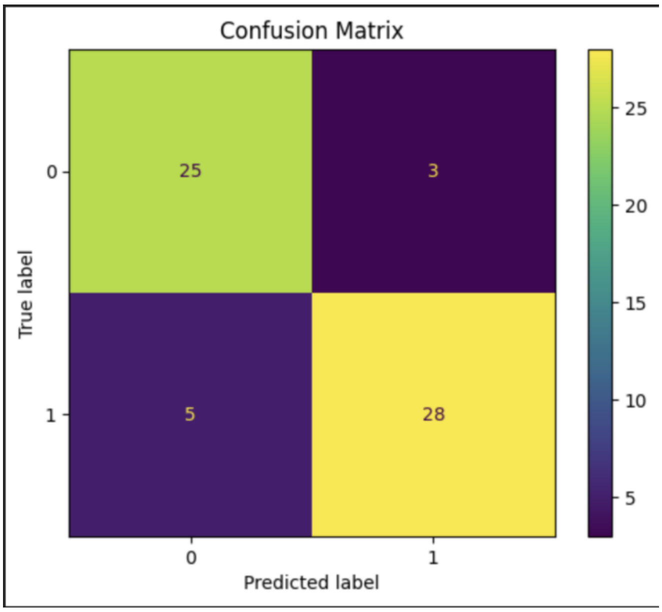


Figure-10: *Confusion Matrix for Voting Classifier*

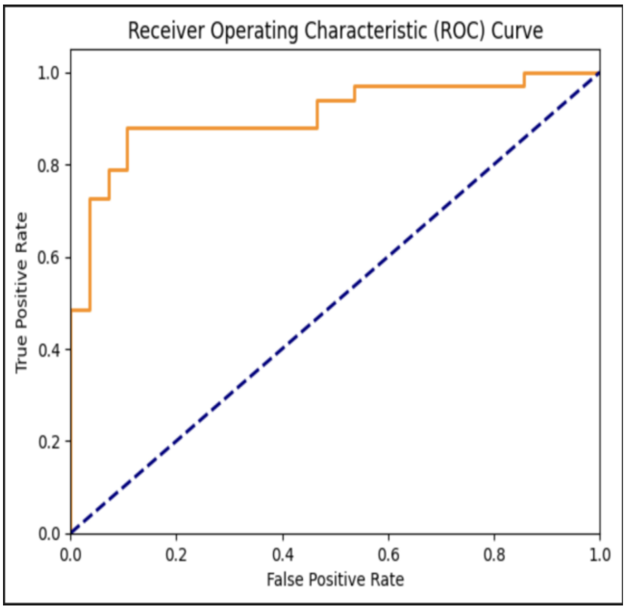


Figure-11: *ROC Curve for Voting Classifier*

V. Results

The table presents a comparative evaluation of four machine learning models for heart disease prediction: Logistic Regression, Naïve Bayes, K-Nearest Neighbours (KNN), and an ensemble method employing a Voting Classifier. Each model's performance metrics, including accuracy, precision, recall, and F1-score, are reported. Logistic Regression exhibits a commendable accuracy of 85.1% along with balanced precision, recall, and F1-score.

Naïve Bayes, although less accurate at 78%, demonstrates reasonable precision and recall. KNN achieves a high accuracy of 84.7% with notably high precision but lower recall, indicating potential challenges in correctly identifying all positive instances. The ensemble method outperforms individual models with an accuracy of 86.8%, showcasing improved precision, recall, and F1-score, highlighting the effectiveness of combining multiple models for enhanced predictive performance.

Table-3: *Performance Metrics*

S. No	**Model Name**	**Accuracy**	**Precision**	**Recall**	**F1-score**
1.	Logistic Regression	0.851	0.843	0.818	0.830
2.	Naïve Bayes	0.847	0.647	0.666	0.656
3.	KNN	0.780	0.892	0.757	0.819
4.	Ensemble (Voting Classifier)	0.868	0.903	0.848	0.875

VI. CONCLUSION

Our study aimed to develop an effective model for predicting heart disease using ensemble machine learning algorithms, specifically Logistic Regression and Naive Bayes. Through rigorous experimentation and evaluation, we achieved a notable accuracy of 86.8%, showcasing the efficacy of our approach in identifying individuals at risk of heart disease. Our findings underscore the importance of leveraging ensemble techniques, which combine the strengths of multiple classifiers to enhance predictive performance. By integrating the complementary characteristics of Logistic Regression and Naive Bayes, our ensemble model demonstrated superior accuracy compared to individual algorithms, thereby offering a robust solution for heart disease prediction. Furthermore, the high accuracy achieved in our study holds significant implications for clinical practice and public health initiatives. Early detection of heart disease risk factors can facilitate timely intervention and preventive measures, ultimately improving patient outcomes and reducing healthcare burdens. Our project underscores the potential of ensemble machine learning algorithms in heart disease prediction and emphasizes the importance of continued research and innovation in leveraging artificial intelligence for improving healthcare outcomes.

VII. REFERENCES

[1] Shuge Ouyang, Research of Heart Disease Prediction Based on Machine Learning, 2022

[2] Narendra Mohan, Vinod Jain, Gauranshi Agrawal, Heart Disease Prediction Using Supervised Machine Learning Algorithms, 2021

[3] A. Lakshmi & R. Devi, Heart Disease Prediction Using Enhanced Whale Optimization Algorithm Based Feature Selection With Machine Learning Techniques, 2023

[4] Mamta Gagoriyal & Mukesh Kumar Khanelwal, Heart Disease Prediction Using Hybrid Machine Learning Approach, 2023

[5] Edupalli Greeshmanth Kumar, Madan Lal Saini, Syed Abbas Khadar Ali, Beeram Bhanu Teja, A Clinical Support System for Prediction of Heart Disease Using Ensemble Learning Techniques, 2023

[6] Krishna Battula, Raghupatruni Durgadinesh, Killi SuryaPratap, Gedela Vinaykumar, Use of Machine Learning Techniques in the Prediction of Heart Disease, 2021

[7] S. Chua, V. Sia, P. N.E. Nohuddin, Comparing Machine Learning Models for Heart Disease Prediction, 2022

[8] Rahul Katarya & Polipireddy Srinivas, Predicting Heart Disease at Early Stages Using Machine Learning: A Survey, 2020

[9] P. Sujatha & K. Mahalakshmi, Performance Evaluation of Supervised Machine Learning Algorithms in Prediction of Heart Disease, 2020

[10] Halima El Hamdaoui, Said Boujraf, Nour El Houda Chaoui, Mustapha Maaroufi, A Clinical Support System for Prediction of Heart Disease Using Machine Learning Techniques, 2020

[11] Goutam Kumar Sahoo, Keerthana Kanike, Santos Kumar Das, Poonam Singh, Machine Learning-Based Heart Disease Prediction: A Study for Home Personalized Care, 2022

[12] R Sivaprasad, M. Hema, Bharati N Ganar, Sunil D M, Vaishali Mehta, Mohammad Fahlevi, Heart Disease Prediction and Classification Using Machine Learning and Transfer Learning Model, 2022

[13] M. Kavitha, G Gnaneswar, R Dinesh, Y Rohith, R Sai Suraj, Heart Disease Prediction Using Hybrid Machine Learning Model, 2021

[14] Meghavi Rana, Mohammad Zia Ur Rehman, Srishti Jain, Comparative Study of Supervised Machine Learning Methods for Prediction of Heart Disease, 2022

[15] Nirmala S, Veena K, Indu B, Jagadevi N. Kalshetty, Heart Disease Prediction Using Artificial Intelligence Ensemble Network, 2022

[16] R Sivaprasad, M. Hema, Bharati N Ganar, Sunil D M, Vaishali Mehta, Mohammad Fahlevi, Heart Disease Prediction and Classification Using Machine Learning and Transfer Learning Model, 2022

[17] P Yadav, Prabhav Saini, Pragya Mittal, Feature Optimization-Based Heart Disease Prediction Using Machine Learning, 2021

[18] Muntasir Mamun, Milon Uddin, Vivek Kumar Tiwari, Asm Mohaimenul Islam, Ahmed Ullah Ferdous, MLHeartDis: Can Machine Learning Techniques Enable to Predict Heart Diseases?, 2022

[19] Vijeta Sharma, Shrinkhala Yadav, Manjari Gupta, Heart Disease Prediction Using Machine Learning Techniques, 2020

[20] Sachin R Jadhav, Rohan Kulkarni, Aditya Yendralwar, Prajot Pujari, Swapnil Patwari, Monitoring and Predicting of Heart Diseases Using Machine Learning Techniques, 2023

[21] Aman Solanki, Anand Vardhan, Aman Jharwal, Narendra Kumar I, Heart Disease Prediction Using Machine Learning, 2023

[22] Md Razu Ahmed, S M Hasan Mahmud, Md Altab Hossin, Hosney Jahan, Sheak Rashed Haider Noori, A Cloud-Based Four-Tier Architecture for Early Detection of Heart Disease with Machine Learning Algorithms, 2018

[23] A Lakshmanarao, A Srisaila, T Srinivasa Ravi Kiran, Heart Disease Prediction Using Feature Selection and Ensemble Learning Techniques, 2021

[24] Suhitha Katari, Thanguturu Likith, Muthineni Phani Sai Sree, Venubabu Rachapudi, Heart Disease Prediction Using Hybrid Machine Learning Algorithms, 2023

[25] V Selvakumar, Advaith Achanta, Nithin Sreeram, Machine Learning-Based Chronic Disease (Heart Attack) Prediction, 2023

[26] Debabrata Swain, Santosh Kumar Pani, Debabala Swain, A Metaphoric Investigation on Prediction of Heart Disease Using Machine Learning, 2024

[27] Puneet, Deepika, Pritpal Singh, Rahul Bansal, Saket Sharma, Coronary Heart Disease Prediction Using Voting Classifier Ensemble Learning, 2021

[28] Arunima Jaiswal, Mounika Singh, Nitin Sachdeva, Empirical Analysis of Heart Disease Prediction Using Deep Learning, 2023

[29] Nazim Nisar Itoo & Vijay Kumar Garg, Heart Disease Prediction Using a Stacked Ensemble of Supervised Machine Learning Classifier, 2022

[30] Arkadeep Bhowmick, Kapil Dev Mahato, Chandrasekhar Azad, Uday Kumar, Heart Disease Prediction Using Different Machine Learning Algorithms, 2022

[31] Gurpreet Singh, Kalpna Guleria, Shagun Sharma, Machine Learning and Deep Learning Models for Early Detection of Heart Disease, 2023

[32] Mandadi Sai Gangadhar, Kalyanam Venkata Sree Sai, Salem Hruthik Sai Kumar, Kanaparti Anil Kumar, M Kavitha, S S Aravinth, Machine Learning and Deep Learning Techniques on Accurate Risk Prediction of Coronary Heart Disease, 2023

[33] Senthil G A, R Prabha, M Razmah, T Veeramakali, Sridevi S, Yashini R, Machine Learning Heart Disease Prediction Using KNN and RTC Algorithm, 2022

[34] Karthikeyan G, Komarasamy G, Daniel Madan Raja S, Modelling an Efficient Heart Disease Prediction System Using Norm- and Regularization-Based Learning Approach, 2022

[35] Aditi Gavhane, Gouthami Kokkula, Isha Pandya, Kailas Devadkar, Prediction of Heart Disease Using Machine Learning, 2018

[36] Atharv Nikam, Sanket Bhandari, Aditya Mhaske, Shamla Mantri, Cardiovascular Disease Prediction Using Machine Learning Models, 2020

[37] Owana Marzia Moushi, Nishat Ara, Md Helaluddin, Himadri Shekhar Mondal, Enhancing the Accuracy and Explainability of Heart Disease Prediction Models Through Interpretable Machine Learning Techniques, 2023

[38] S Hemalatha, T Kavitha, K Chitra, B Babitha, N Kavya, R Keshika, Extensive Review on Predicting Heart Disease Using Machine Learning and Deep Learning Techniques, 2023

[39] Taher M Ghazal, Amer Ibrahim, Ali Sheraz Akram, Zahid Hussain Qaisar, Sundus Munir, Sanza Islam, Heart Disease Prediction Using Machine Learning, 2023

[40] Sulistyo Damas Prakoso, Adhistya Erna Permanasari, Azkario Rizky Pratama, Heart Disease Prediction Using Machine Learning: A Systematic Literature Review, 2023

[41] M Priya, A Anitha, M K Nallakaruppan, J Raju, Raghavan R, Divyansh Srivastava, Heart Disease Prediction Using Machine Learning Algorithms, 2023

[42] Sunitha Guruprasad, Valesh Levin Mathias, Winslet Dcunha, Heart Disease Prediction Using Machine Learning Techniques, 2021

[43] Nahida Nazir, Aditya Saraogi, Kanksha Kaur, Vinod Bera, Akash Kumar Singh, Omdev Dahiya, Navlesh Kumar Singh, Kapil Bhardwaj, Suman Avdhesh Yadav, Heart Disease Prediction Using Machine Learning, 2023

[44] Anurag, Shreya Kalta, Jatin Thakur, Himanshu Bhardwaj, Yogesh Banyal, Optimized Ensemble Model for Heart Disease Prediction Using Machine Learning, 2023

[45] Ansh Bahuguna, Meenu Gupta, Rakesh Kumar, Statistical Analysis and Prediction of Heart Disease Using Machine Learning, 2023

[46] Shuying Shen, A Multi-Source Based Healthcare Method for Heart Disease Prediction by Machine Learning, 2021

[47] Ch Raja Shaker, Anisetti Sidhartha, Anto Praveena, A Christy, B Bharati, An Analysis of Heart Disease Prediction Using Machine Learning and Deep Learning Techniques, 2022

[48] Ahmed Al Ahdal, Manik Rakhra, Sumit Badotra, Taha Fadhaeel, An Integrated Machine Learning Techniques for Accurate Heart Disease Prediction, 2022

[49] Ambika Sekhar, Amrutha Babu, Jayalekshmi V K, Adithya Udayan, Machine Learning Based Heart Disease Prediction, 2022

[50] Dolley Srivastava, Himanshu Pandey, Ambuj Kumar Agarwal, Richa Sharma, Opening the Black Box: Explainable Machine Learning for Heart Disease Patients, 2023