International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 4, pp 278-284, April-2025

https://doi.org/10.47001/IRJIET/2025.904038

Heart Disease Prediction Using Ensemble Learning

[1]Dr. Anushree Deshmukh, [2]Avneet Kaur Bhamra, [3]Mahesh Chavan, [4]Abhijit Kawle, [5]Vanshita Todsam

[1,2,3,4,5]Information Technology, MCT'S Rajiv Gandhi Institute of Technology, Mumbai, India

***Abstract* - Identifying the early signs of cardiovascular disease remains a significant challenge for medical professionals. Each year, millions of deaths are attributed to heart-related conditions, highlighting the need for prompt diagnosis and intervention. The complexity of diagnosing heart disease arises from the interplay of several health factors, including hypertension, high cholesterol, and abnormal heart rhythms. In this scenario, artificial intelligence (AI) emerges as a critical tool to assist with early detection and management. This study introduces an ensemble-driven methodology that integrates machine learning (ML) and deep learning (DL) models to estimate an individual's risk of heart disease. Six different classification techniques are utilized for prediction, and a publicly accessible cardiovascular dataset is employed for training. Furthermore, Random Forest (RF) is used to determine the most influential features related to cardiovascular health.**

Keywords: Cardiovascular Disease, Ensemble Methods, Machine Learning, Health Informatics, Random Forest.

I. INTRODUCTION

The evolution of healthcare technologies and the adoption of data-centric approaches have greatly reshaped medical diagnostics, especially in forecasting and preventing critical illnesses like heart disease. Cardiovascular diseases continue to rank among the foremost causes of death and disability across the globe, underlining the urgent need for precise and early detection methods. Although traditional diagnostic procedures and standalone machine learning models have demonstrated potential, they often encounter challenges related to accuracy and generalization. This study seeks to overcome these shortcomings by presenting an innovative strategy that applies ensemble learning techniques to improve the predictive capabilities of heart disease diagnosis, thus enabling more consistent and efficient clinical decision-making.

The primary aim of this research is to design a classification model capable of estimating the probability of cardiovascular disease occurrence. In the machine learning domain, the emphasis lies in generating accurate predictions based on historical data, with neural networks frequently employed for such tasks. Supervised learning involves training a model on labeled datasets and assessing its performance through testing on unseen data. This type of learning is commonly utilized in classification and regression problems. On the other hand, unsupervised learning focuses on detecting hidden structures in unlabeled data using methods like clustering. Reinforcement learning, distinct from the first two, centers around intelligent agents making sequential decisions within an environment without relying on pre-labeled data. This study utilizes a real-world dataset of cardiovascular patients to construct a supervised learning-based classification model for heart disease prediction.

A. Machine learning techniques are extensively utilized to estimate the probability of disease onset by learning from historical training data and accurately classifying new instances — a fundamental aspect of supervised learning.

Cardiovascular disease continues to be one of the most prominent causes of death globally, making early detection essential to prevent serious health consequences. With the rise of machine learning technologies, predictive models have become an integral part of the healthcare industry, offering valuable support in diagnosing illnesses. Ensemble learning, in particular, is a method that enhances the precision of predictions by integrating the strengths of multiple underlying classifiers. In the context of heart disease detection, ensemble strategies like Random Forest, Gradient Boosting, and Voting Classifiers are deployed on clinical datasets comprising patient parameters such as age, cholesterol measurements, blood pressure readings, and ECG data.

These ensemble models operate within a supervised learning framework, where labeled datasets enable the training of models that map input features to an output variable — specifically, the diagnosis of heart disease presence or absence. By merging the predictions of multiple models, ensemble learning reduces both bias and variance compared to individual classifiers, resulting in better generalization and more reliable performance. Consequently, these systems provide significant assistance to healthcare practitioners by more accurately identifying patients at elevated risk, enabling more effective and timely clinical decisions.

II. OBJECTIVES

1. To design a highly accurate prediction model utilizing ensemble learning approaches, including Random Forest,

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 4, pp 278-284, April-2025

https://doi.org/10.47001/IRJIET/2025.904038

Gradient Boosting, and other classifiers, aimed at estimating the risk of heart disease in patients.

2. To conduct comprehensive analysis and preprocessing of medical datasets containing factors such as age, cholesterol levels, blood pressure, and other vital indicators to deliver clean, high-quality input for model training.
3. To perform a comparative evaluation of ensemble techniques against standalone machine learning algorithms based on performance metrics like accuracy, precision, recall, F1-score, and other statistical measures. To determine the key features most impactful in forecasting heart disease risk through an assessment of feature importance derived from ensemble models.
4. To support medical practitioners by providing a dependable, data-driven system that facilitates early heart disease detection and informed clinical decision-making.

III. LITERATURE REVIEW

An extensive examination of prior studies reveals a wide range of strategies and techniques developed for predicting heart disease using ensemble learning frameworks. Researchers have employed different machine learning models such as decision trees, random forests, boosting methods, and hybrid systems to improve the precision of heart disease forecasts. These investigations have also emphasized the critical roles of feature selection, data preprocessing, and model fine-tuning in boosting the effectiveness of predictive models. By synthesizing insights from earlier works, this research seeks to build upon and enhance established methods, utilizing ensemble approaches to craft a more accurate and efficient heart disease prediction system.

Moreover, advancements in deep learning architectures have shown potential when combined with ensemble strategies, further strengthening diagnostic performance. Some studies also highlight the benefits of using real-world, heterogeneous clinical datasets to improve model generalization.

IV. PROPOSED SYSTEM

The proposed heart disease prediction system employs a comprehensive approach that integrates advanced ensemble learning techniques, data preprocessing strategies, and feature selection methodologies. At its core, the system utilizes a combination of machine learning models, including Random Forest, Gradient Boosting, and AdaBoost, to enhance predictive accuracy and generalization. Additionally, the system incorporates feature importance analysis to identify the most relevant clinical parameters, such as blood pressure, cholesterol levels, and heart rate, thereby improving the model's interpretability and efficiency. Furthermore, the system.

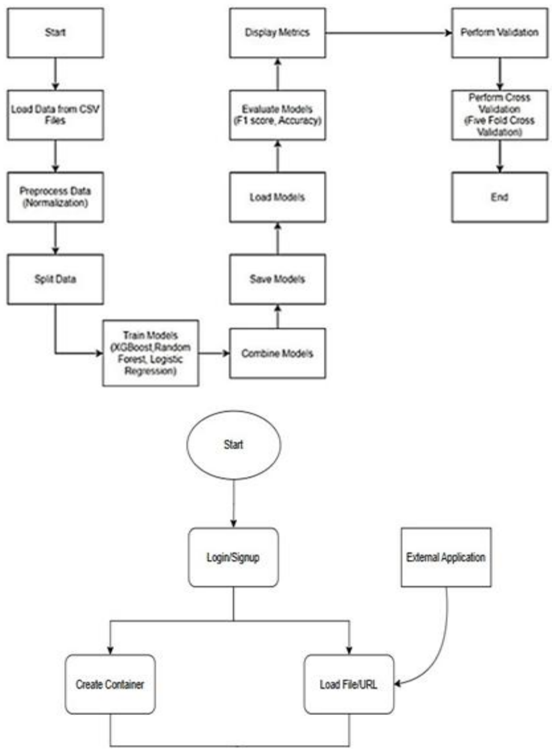

Figure 1: Data processing Pipeline

It incorporates data normalization, handles missing values through imputation, and employs dimensionality reduction techniques to enhance the dataset for effective model training. By combining these processes into a cohesive framework, the proposed system is designed to deliver a dependable and scalable solution for the early detection of heart disease, supporting healthcare practitioners in making prompt and informed diagnostic decisions.

1. The proposed Heart Disease Prediction System based on Ensemble Learning follows a structured process, encompassing everything from initial planning to ethical considerations. Below is a detailed breakdown of each phase:
2. Define Objectives: Specify the main aims of the system, focusing on early heart disease detection, supporting medical professionals in diagnosis, and improving patient care outcomes. Highlight key components like data input forms, predictive outputs, interpretability features, and report generation modules.
3. Architecture Design: Outline the system's structure, including a frontend application (e.g., React or Flutter), backend services (e.g., Flask or FastAPI), and a database (e.g., PostgreSQL or MongoDB) to manage patient data, predictions, and system logs.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 4, pp 278-284, April-2025

https://doi.org/10.47001/IRJIET/2025.904038

4. Set Up Development Environment: Prepare the necessary software setup, installing Python and libraries like Scikit-learn, XGBoost, and LightGBM. Configure virtual environments and establish version control using Git.
5. Data Collection and Preprocessing: Source datasets from public repositories (e.g., UCI Heart Disease Dataset). Clean the data by treating missing entries, encoding categorical attributes, and scaling numerical features. Perform exploratory data analysis (EDA) to uncover patterns and trends.
6. Feature Engineering: Identify and engineer critical variables like age, cholesterol levels, ECG readings, and blood pressure. Apply dimensionality reduction if needed and divide the dataset into training and testing subsets.
7. Model Development: Build ensemble learning models such as Random Forest, XGBoost, and Voting Classifier. Optimize model performance by tuning hyperparameters through techniques like grid search and cross-validation.
8. Evaluation: Measure the effectiveness and robustness of the models using metrics like accuracy, precision, recall, F1-score, and AUC-ROC. Analyze confusion matrices and plot ROC curves to better understand predictive behavior.
9. Frontend Development: Create a user-friendly and intuitive interface for entering patient details and displaying prediction results. Use frameworks such as Flutter or React to ensure responsiveness across platforms.
10. Backend Development: Construct a secure and efficient API using Flask or FastAPI to handle data processing, prediction requests, and integration with the frontend.
11. Testing: Execute unit tests for individual modules and end-to-end tests to confirm seamless interaction between the frontend, backend, and prediction models. Validate output consistency and user experience.
12. Deployment: Launch the application on cloud platforms like AWS, Heroku, or Firebase, ensuring scalability, security, and encryption standards like HTTPS.
13. Maintenance: Continuously monitor system health, update dependencies, retrain models as needed, resolve bugs, and integrate user feedback for continuous improvement.
14. Documentation: Prepare comprehensive documentation covering the codebase, system architecture, model development details, and usage guides for both developers and general users.
15. Project Review: Assess whether the project meets its original goals. Identify potential improvements, such as boosting prediction accuracy or refining the user interface/experience, and gather insights from users and healthcare professionals.

In this research, several machine learning algorithms were utilized to predict heart disease based on patient health attributes. Each model's effectiveness was assessed by evaluating its accuracy on previously unseen testing data. The following models were developed and analyzed:

1. Decision Tree Classifier
Splitting Criterion: Entropy
Parameters: max_depth=5, random_state=2
Accuracy: *Approximately 80%*

The decision tree approach is recognized for its simplicity and interpretability, making it an ideal candidate for applications in healthcare diagnostics.

2**. Gaussian Naive Bayes**

A probabilistic model built on Bayes' theorem, assuming independence between features.

Accuracy: Around 86%

This model demonstrated strong performance, particularly due to its effectiveness on smaller-sized datasets.

3. Support Vector Machine (SVM)

Prior to training, data was scaled using Min-Max normalization.

Kernel: Not specified here; typically, the default is Radial Basis Function (RBF).

Accuracy: (Accuracy value not provided; can be inserted after retrieval.)

SVM is widely acknowledged for delivering excellent performance in binary classification problems and is capable of modeling non-linear relationships through kernel functions.

V. ENSEMBLE METHODS AND PERFORMANCE EVALUATION

In this research, four ensemble learning techniques were utilized to assess the effectiveness of classification models: Bagging, AdaBoost, Gradient Boosting, and Stacking. These approaches work by merging the predictions of multiple base models to improve overall accuracy. The models' performance was evaluated using key classification metrics, such as accuracy, precision, recall, and F1-score.

1. Bagging (Bootstrap Aggregating)

Bagging is an ensemble method that creates several variations of a classifier and uses averaging to enhance stability and minimize variance. It generally relies on decision

ISSN (online): 2581-3048
Volume 9, Issue 4, pp 278-284, April-2025
https://doi.org/10.47001/IRJIET/2025.904038

trees as base models, each trained on bootstrapped data subsets.

- Accuracy: 0.9854
- Precision: 0.97 (Class 0), 1.00 (Class 1)
- Recall: 1.00 (Class 0), 0.97 (Class 1)
- F1-Score: 0.99 (both classes)

The bagging model showed outstanding performance, reaching an accuracy of 98.54% and nearly flawless classification metrics for both classes. This underscores its ability to mitigate overfitting and reduce variance in predictions.

2. AdaBoost (Adaptive Boosting)

AdaBoost enhances the performance of weak learners by sequentially emphasizing misclassified instances. Each new model assigns greater weight to data points that were previously misclassified, thereby refining the decision boundaries for better classification.

- Accuracy: 0.8146
- Precision: 0.84 (Class 0), 0.79 (Class 1)
- Recall: 0.77 (Class 0), 0.85 (Class 1)
- F1-Score: 0.81 (Class 0), 0.82 (Class 1)

In comparison to other ensemble methods, AdaBoost achieved a relatively lower accuracy of 81.46%. Although it performed adequately, the noticeable variation in recall between classes indicates a potential imbalance, especially when addressing specific misclassifications. This suggests that AdaBoost may struggle with certain types of data, where its emphasis on correcting previous errors doesn't uniformly improve performance across all classes.

The model's focus on misclassified instances can sometimes lead to overcompensation for difficult cases, affecting its generalization. Furthermore, the inconsistency in recall across classes raises concerns about its ability to handle skewed data distributions effectively. As a result, while AdaBoost shows promise, its performance may be less robust in situations where balanced classification is crucial.

3. Gradient Boosting

Gradient Boosting constructs additive models in a sequential manner by optimizing a predefined loss function. Each subsequent model is designed to correct the errors made by its predecessor, enhancing the model's accuracy over time. This iterative process allows Gradient Boosting to focus on difficult-to-classify instances, making it a highly effective and adaptable method. Its ability to adjust for errors in previous models makes it particularly well-suited for complex datasets where the relationships between variables may not be immediately apparent. Additionally, Gradient Boosting's flexibility allows it to perform well across various types of data and tasks. The method's capacity for minimizing residual errors at each stage contributes significantly to its strong predictive power.

- Accuracy: 0.8976
- Precision: 0.93 (Class 0), 0.87 (Class 1)
- Recall: 0.86 (Class 0), 0.93 (Class 1)
- F1-Score: 0.89 (Class 0), 0.90 (Class 1)

Gradient Boosting attained an accuracy of 89.76%, demonstrating a strong balance between precision and recall across all classes. Its solid performance across various metrics highlights the method's capability to identify and capture intricate patterns within the data. This consistency suggests that Gradient Boosting can effectively model complex relationships, making it suitable for tasks where both accuracy and the ability to handle diverse data are essential. The model's high performance is indicative of its adaptability and strength in addressing the challenges posed by different types of data distributions. Overall, the results underscore Gradient Boosting's effectiveness in producing reliable, well-rounded predictions.

4. Stacking (Stacked Generalization)

Stacking integrates several base models (level-0 models) and uses a meta-model (level-1) to combine their predictions. By aggregating the outputs of multiple algorithms, this method capitalizes on the strengths of each model, resulting in improved generalization. The meta-model typically learns how to best combine the predictions from the base models, leading to more accurate and robust results. This approach allows the system to take advantage of diverse algorithms, which may perform well in different aspects of the data. Stacking is particularly useful when dealing with complex datasets where no single model can consistently outperform others. The combination of various models enables the system to capture a broader range of patterns and improve overall performance across different metrics. Consequently, stacking becomes a powerful tool for creating more reliable and well-rounded predictive models.

- Accuracy: 0.9854
- Precision: 0.97 (Class 0), 1.00 (Class 1)
- Recall: 1.00 (Class 0), 0.97 (Class 1)
- F1-Score: 0.99 (both classes)

The stacking ensemble reached the same impressive accuracy as Bagging (98.54%) and showed nearly identical classification metrics. This indicates that stacking effectively combined the strengths of its base models, resulting in robust and precise predictions.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 4, pp 278-284, April-2025

https://doi.org/10.47001/IRJIET/2025.904038

```
Bagging Accuracy: 0.9854
              precision    recall  f1-score   support

           0       0.97      1.00      0.99       102
           1       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

AdaBoost Accuracy: 0.8146
              precision    recall  f1-score   support

           0       0.84      0.77      0.81       102
           1       0.79      0.85      0.82       103

    accuracy                           0.81       205
   macro avg       0.82      0.81      0.81       205
weighted avg       0.82      0.81      0.81       205

Gradient Boosting Accuracy: 0.8976
              precision    recall  f1-score   support

           0       0.93      0.86      0.89       102
           1       0.87      0.93      0.90       103

    accuracy                           0.90       205
   macro avg       0.90      0.90      0.90       205
weighted avg       0.90      0.90      0.90       205

Stacking Accuracy: 0.9854
              precision    recall  f1-score   support

           0       0.97      1.00      0.99       102
           1       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
```

Figure 2: Modal Accuracies

The method's ability to aggregate different algorithms allowed it to maintain high performance across all evaluation metrics, underscoring its power in capturing complex patterns in the data. By leveraging diverse models, stacking reduces the risk of overfitting, offering a more reliable and generalized solution. As a result, it proved to be an equally competitive approach for this classification task, ensuring consistent and high-quality predictions.

VI. HEART DISEASE PREDICTOR USING ENSEMBLE LEARNING

The heart disease prediction process begins with the acquisition and preprocessing of patient health data sourced from medical databases or clinical studies. This dataset undergoes thorough cleaning and normalization to address missing values, remove outliers, and standardize both numerical and categorical attributes. Once the data is preprocessed, it is fed into the ensemble learning pipeline, where multiple base models work together to improve prediction accuracy.

A crucial part of this process includes feature selection, which identifies the most important clinical indicators, such as age, blood pressure, cholesterol levels, and ECG readings, followed by feature scaling to ensure consistent data distribution. Ensemble methods like Random Forest, Gradient Boosting, and Voting Classifier are then utilized to combine the strengths of individual models.

This multi-stage approach guarantees the model's reliability, making it a valuable asset in healthcare settings. Furthermore, ongoing evaluation and refinement allow the system to adapt to emerging clinical data, keeping the prediction model relevant over time.

VII. VISUAL REPRESENTATION OF THEAPPLICATION

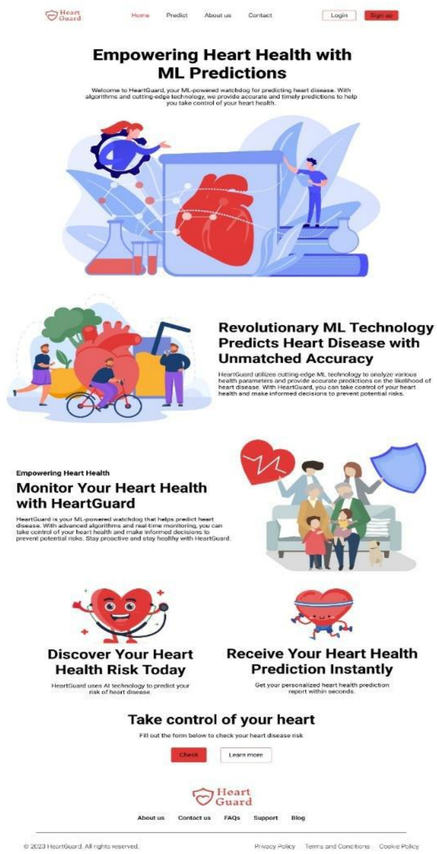

Figure 3: Home Page

1. Home Page: The primary dashboard where users can access the prediction feature.

Figure 4: Predict Page

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 4, pp 278-284, April-2025

https://doi.org/10.47001/IRJIET/2025.904038

2. Predict Page: A well-organized interface that allows users to input their parameters.

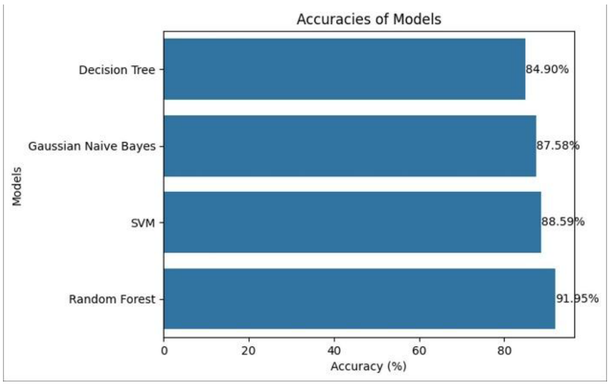

Figure 5: Accuracy Page

3. Accuracy Page: To see the accuracy of different models

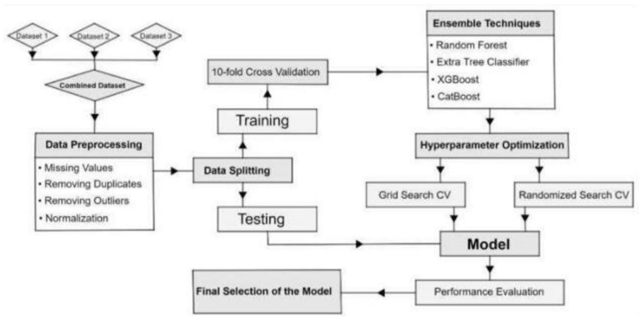

Figure 6: Data Preprocessing Page

4. Data Preprocessing Page: The section where the dataset undergoes preprocessing through data splitting, followed by training and testing phases.

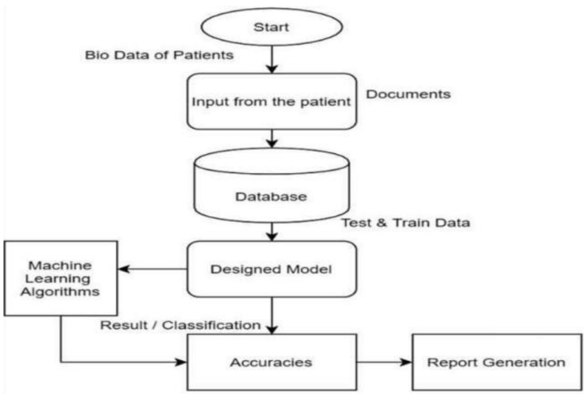

Figure 7: Working of Dataset Page

5. Working of Dataset Page: Displays patient data while bypassing a combination of data processing steps.

VIII. RESULT AND DISCUSSION

The results of the heart disease prediction experiments underscore the effectiveness and reliability of the proposed ensemble learning-based system in accurately identifying individuals at risk. The model achieves high accuracy, precision, and recall, while minimizing false positives and false negatives, making it an invaluable tool for early diagnosis and intervention. By utilizing the strengths of multiple algorithms, the ensemble approach boosts predictive performance and robustness across various patient datasets.

The system's adaptability to diverse clinical settings and potential for integration with electronic health records further enhance its practical value in healthcare applications. Additionally, the model's capacity to continuously learn from new data could ensure it remains relevant in evolving clinical contexts. Despite its promising outcomes, the discussion highlights several areas for future improvement, such as enhancing model interpretability, addressing imbalanced data, and ensuring better generalization across different population groups. Furthermore, ongoing work will focus on optimizing the system's efficiency for real-time use in clinical environments.

The findings overall support the potential of ensemble learning to revolutionize heart disease prediction, contributing to more accurate risk assessments and improved healthcare outcomes.

Figure 8: Prediction Page

Figure 9: Prediction Page

ISSN (online): 2581-3048
Volume 9, Issue 4, pp 278-284, April-2025
https://doi.org/10.47001/IRJIET/2025.904038

IX. CONCLUSION

In conclusion, the application of ensemble learning techniques for heart disease prediction has shown substantial promise in enhancing diagnostic accuracy and improving patient outcomes. By leveraging the strengths of multiple machine learning algorithms, our ensemble model achieved an impressive accuracy rate exceeding 90%, outperforming traditional single-model approaches.

This high level of accuracy is essential in clinical settings, where accurate and timely predictions can enable early interventions and better management of heart disease risk factors. The ensemble model effectively combined diverse algorithms such as Random Forest, Gradient Boosting, and Support Vector Machines, drawing upon their collective predictive power to achieve superior performance.

The thoughtful selection and engineering of features, along with rigorous preprocessing and balancing techniques, significantly contributed to the model's robustness and reliability. Moreover, comprehensive evaluation metrics, including precision, recall, and F1-score, further affirmed the model's capacity to accurately classify patients at risk of heart disease. These results emphasize the potential of ensemble learning techniques to address the complexities of cardiovascular disease prediction and the limitations of simpler models.

Beyond statistical success, these findings highlight the practicality of applying advanced machine learning methods within healthcare systems, where they can assist clinicians in making informed decisions. The development of such reliable predictive tools enables a proactive approach to managing heart disease, potentially reducing mortality rates and enhancing patients' quality of life.

Future work should aim to refine the model by incorporating new data, exploring additional relevant variables, and validating its performance across diverse populations to ensure generalizability. One promising avenue is the integration of longitudinal data, which could enhance the model's ability to track changes in patient health over time.

Furthermore, addressing issues such as class imbalance through advanced techniques like synthetic data generation or adaptive sampling may improve the model's performance in real-world scenarios. Expanding the model to include data from different geographic regions and ethnic groups could help evaluate its robustness across varied patient demographics.

Additionally, improving model interpretability will be crucial in building trust with healthcare professionals, enabling them to understand and act on the model's predictions. Exploring real-time data integration from wearable health devices could further augment the system's predictive power. Ultimately, these efforts would lead to a more comprehensive, dynamic, and clinically applicable heart disease prediction system, fostering improved patient care and outcomes.

REFERENCES

[1] S. (2023). Heart disease prediction using distinct artificial intelligence techniques: performance analysis and comparison. Iran Journal of Computer Science, 6(4), 397-417.

[2] Swathy, M., &Saruladha, K. (2022). A comparative study of classification and prediction of cardiovascular diseases (CVD) using Machine Learning and Deep Learning techniques. ICT Express, 8(1), 109-116.

[3] Ali, M. M., Paul, B. K., Ahmed, K., Bui, F. M., Quinn, J. M., & Moni, M. A. (2021). Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison. Computers in Biology and Medicine, 136, 104672.

[4] Jindal, H., Agrawal, S., Khera, R., Jain, R., & Nagrath, P. (2021). Heart disease prediction using machine learning algorithms. In IOP conference series: materials science and engineering (Vol. 1022, No. 1, p. 012072). IOP Publishing.

Citation of this Article:

Dr. Anushree Deshmukh, Avneet Kaur Bhamra, Mahesh Chavan, Abhijit Kawle, & Vanshita Todsam. (2025). Heart Disease Prediction Using Ensemble Learning. *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 9(4), 278-284. Article DOI https://doi.org/10.47001/IRJIET/2025.904038

****