J. Electrical Systems 20-11s (2024): 2114-2125

Sagar Saikia[1]*
Kshirod Sarmah[2]*
Madhurjya Borah[3]
Hrishiraj Sawan[4]
Ankita Barbora[5]
Joyita Das Gupta[6]
Rima Devi[7]

Analysis of Heart Disease Prediction using Novel Machine Learning and Deep Learning Techniques

Journal of Electrical Systems

Abstract:
Heart disease stands as one of the most prevalent causes of death globally, and India is no exception. In India, a considerable number of fatalities can be attributed to heart-related issues each year. Factors like high blood pressure, diabetes, smoking habits, and increasingly inactive lifestyles have further contributed to the growing cases of heart disease in the country. Identifying the disease at an early stage, as well as predicting its likelihood, plays a crucial role in enhancing patient outcomes and minimizing the overall healthcare burden.This research paper explores the development and evaluation of predictive models for heart disease using a dataset comprising various clinical and demographic features. Leveraging machine learning techniques, including Convolutional Neural Networks (CNN) and other models, we aim to identify key predictors of heart disease and develop an accurate model to assist healthcare professionals in early diagnosis and intervention. The study employs comprehensive data preprocessing, feature selection, and model evaluation to ensure robust and reliable predictions. Our findings highlight the potential of machine learning and deep learning models to significantly enhance heart disease prediction, thereby contributing to better management and prevention strategies in India.

Keywords: Heart disease, Machine Learning, Deep Learning CNN, LSTM.

1. INTRODUCTION

Heart disease continues to be a major global health issue, ranking as one of the leading causes of death. The World Health Organization (WHO) estimates that cardiovascular diseases claim approximately 17.9 million lives annually across the globe. In India, heart disease rates have surged, influenced by both genetic factors and changes in lifestyle. Poor dietary choices,insufficient physical activity, and rising incidences of smoking and diabetes have all contributed to theproblem. The growing concern underscores the need for better methods of early detection and prevention to combat the disease effectively.This rising trend necessitates the development of effective prediction models to aid in early diagnosis and intervention.

Traditional diagnostic methods for heart disease often involve invasive procedures and can be time-consuming and costly. As a result, there has been a growing interest in the use of machine learning (ML) and deep learning (DL) techniques to develop non-invasive, efficient, and accurate prediction models. These techniques have the potential to analyze large volumes of clinical data, uncover hidden patterns, and provide reliable predictions, thereby supporting healthcare professionals in making informed decisions.

Recent research has shown substantial advancements in the accuracy of heart disease prediction, thanks to the development of sophisticated machine learning (ML) and deep learning (DL) techniques. For instance, a study by García-Ordás et al. (2023) combined deep learning methods with feature enhancement strategies, resulting in an impressive precision of 90%. This marked an improvement of 4.4% over earlier approaches. In another study, Salhi et al. (2021) explored heart disease prediction through data analytics and found that neural networks delivered significantly better results—93% accuracy—compared to methods like Support Vector Machines (SVM) and K-Nearest Neighbors (KNN).

This paper explores the application of various ML and DL algorithms, including Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), Support Vector Machines (SVM), and Naive Bayes, in predicting heart disease. CNNs are particularly effective in detecting complex patterns within data, making them suitable for tasks such

[1*]Department of Computer Science and Engineering, National Institute of Technology, Meghalaya, 793003, Email: p19cs010@nitm.ac.in
[2*] Department of Computer Science, Pandit Deendayal Upadhyaya Adarsha Mahavidyalaya (A Govt. Model College), Goalpara, 783124, Assam, India, Email: kshirodsarmah@gmail.com
[3]Department of Business Administration, Assam Institute of Management, Guwahati,781035, Assam, India, Email: boramadhu13@gmail.com
[4,5]Department of Computer Science and Engineering, Indian Institute of Information Technology, Guwahati, 781015, Assam, India,
[4]Email: hrishiraj.sawan24m@iiitg.ac.in
[5]Email: ankitabarbora5@gmail.com
[6]Department of Computer Application, Chandigarh University, Chandigarh, 140413, India, Email: joyitadasgupta49@gmail.com
[7]National Institute of Electronics and Information Technology, Guwahati,781022, Assam, India, Email: devi.rima7@gmail.com
***Corresponding Authors:** Kshirod Sarmah and Sagar Saikia
*Email: kshirodsarmah@gmail.com and p19cs010@nitm.ac.in

as image analysis and pattern recognition in clinical data. LSTM networks, a type of recurrent neural network, are adept at handling sequential data and time series analysis, which are common in patient health records. SVM is a robust classification algorithm that performs well with high-dimensional data, while Naive Bayes is a probabilistic classifier known for its simplicity and efficiency.

The primary objective of this research is to enhance the prediction accuracy of heart disease using these advanced ML and DL models. By leveraging comprehensive datasets and employing sophisticated data preprocessing, feature selection, and model evaluation techniques, we aim to build a robust predictive system that can assist healthcare professionals in early diagnosis and intervention. The ultimate goal is to improve patient outcomes, reduce healthcare costs, and contribute to better management and prevention strategies for heart disease in India and globally.

2. LITERATURE REVIEW

A Angel Nancy et al. [1], The, convergence of IoT, cloud technology, and deep learning is transforming healthcare. Predictive analytics using advanced AI and ML approaches can accurately predict heart disease risk. The proposed smart healthcare system, powered by Bi-LSTM, achieves impressive accuracy (98.86%)— outperforming existing systems. It used Cleveland and Hungarian dataset. And the methods used here are RNN, LSTM, Proposed Bi-LSTM Model. The smart healthcare system for monitoring and accurately predicting heart disease risk built around Bi-LSTM (bidirectional long short-term memory) showcases an accuracy of 98.86%, a precision of 98.9%, a sensitivity of 98.8%, a specificity of 98.89%, and an F-measure of 98.86%, which are much better than the existing smart heart disease prediction systems.

Ali Mohamed Hussien et al. [2], The paper proposes a deep stacking ensemble model for early detection of heart disease. It combines two hybrid deep learning models, CNN-LSTM and CNN-GRU, with Support Vector Machine (SVM) as the meta-
learner. Recursive Feature Elimination (RFE) is used for feature optimization. The model is tested on two heart disease datasets and compared with five machine learning models and other hybrid models. The proposed ensemble model outperforms the others, showing promise for early heart disease prediction. Two dataset is used here one is first dataset and another is Cleveland dataset. And the methods used here are Deep staking ensemble, CNN-LSTM Model, CNN-GRU Model, Recursive feature Elimination, Classical ML Models. it is stated that the proposed model, which is the deep staking ensemble method integrating CNN-LSTM and CNN-GRU models with an SVM meta-learner, achieved the highest accuracy among all compared models. Specifically, for the first dataset, it achieved an accuracy of 78.81%, and for the Cleveland dataset, it achieved an accuracy of97.17%. Therefore, the proposed model yielded the best accuracy in this study.
Sadia Arooj et al. [3], This study proposes a deep learning approach using image classification for heart disease detection. By utilizing a deep convolutional neural network (DCNN) on the UCI heart disease dataset, the model achieves a validation accuracy of 91.7%. The results suggest that this approach holds promise forimproving real-world heart disease detection. The methods used here are data cleaning and filtering, Convolutional Neural Network, Train-test split, Model training and Model Validation. The method that gives the best accuracy in this study is the Convolutional Neural Network (CNN) model, which achieved a prediction accuracy of 91.71%.

M.R.I Faruque et al. [4], The study employs machine learning algorithms to detect heart disease, achieving high accuracy rates on various datasets. Utilizing random forest, decision tree, AdaBoost, and K-nearest neighbor models, the study reports accuracy percentages ranging from 93.437% to 100%. The research also develops a computer-aided smart system for disease prediction using Streamlit. Notably, the study explores significant predictors contributing to heart disease prognosis. Overall, this research contributes to the efficient diagnosis of heart conditions and underscores the importance of machine learning in healthcare. It used Cleveland ,Hungary, Switzerland, Long Beach dataset. The methods used here are randomforest, Decision tree, Ada boost, K Nearest Neighbor. the maximum accuracies achieved by the machine learning models are as follows: Random Forest (RF) on the CHSLB dataset: 99.03%, Decision Tree (DT) on the CHSLB dataset: 96.10%, AdaBoost (AB) on the CHSLB dataset: 100%, K-nearest neighbor (KNN) on the CHSLB dataset: 100%. On the Cleveland dataset: Random Forest (RF): 93.437% K-nearest neighbor (KNN): 97.83%.

Mohammed S.Alqahtani et al. [5], The paper presents a heart disease prediction model using machine learning algorithms on combined datasets. Random Forest Classifiers (RFC) demonstrate high accuracy and reliability, achieving 100% in multiple performance metrics. Cross-validation and feature reduction methods are employed to enhance predictive capabilities. Results indicate strong performance across various classifiers, suggesting efficacy in early heart disease prediction.
Overall, the proposed technique offers a rapid and accurate approach to identify potential cardiac issues. It Used Cleveland database National Cardiovascular Disease Surveillance (NCDS) System's heart disease database Kaggle heart disease dataset, which combines data from: Cleveland, Kaggle heart disease dataset, which combines data from: Cleveland, Hungary, Switzerland, VA Long Beach. The methods used here are: Heart disease prediction model, Analysis. The maximum accuracy receivedin the paper for the Random Forest Classifiers (RFC) on the combined heart diseasedatasets is reported as 100%. This indicates that the model achieved perfect accuracy in predicting heart disease based on the analyzed data.

Convolutional Neural Networks (CNN)
CNNs have been widely used in medical image analysis and have shown remarkable success in detecting and diagnosing various diseases, including heart disease. A study by Thakkar and Agrawal (2023) demonstrated the effectiveness of a hybrid feature selection and optimized Deep CNN model, achieving an accuracy of 95%, sensitivity of 94.9%, and specificity of 93.8% for heart disease prediction. This highlights the potential of CNNs in accurately predicting heart disease by analyzing complex patterns in the data. Additionally, García-Ordás et al. (2023) showed that combining deep learning with feature augmentation techniques significantly improved precision in cardiovascular disease prediction.

Long Short-Term Memory (LSTM)
LSTM networks, a type of recurrent neural network (RNN), are well-suited for time series analysis and sequential data. They have been used to predict various health conditions by analyzing patient records over time. Although specific studies on LSTM for heart disease prediction are limited, their ability to handle temporal dependencies makes them a promising tool for this application. Future research could explore the use of LSTM networks in conjunction with other ML models to enhance prediction accuracy.

Support Vector Machines (SVM)
SVM is a powerful classification algorithm that has been effectively used for heart disease prediction. Elsedimy et al. (2024) proposed an improved SVM model trained using a quantum-behaved particle swarm optimization (QPSO) algorithm, achieving a prediction accuracy of 96.31% on the Cleveland heart disease dataset. This study underscores the robustness of SVM in handling high-dimensional data and providing accurate predictions. Salhi et al. (2021) also demonstrated that SVM, along with neural networks and KNN, could effectively predict cardiac disease, with neural networks showing superior performance.

Naive Bayes
The Naive Bayes algorithm, known for its simplicity and effectiveness, has also been employed in heart disease prediction. Its probabilistic approach makes it suitable for handling large datasets with numerous features. Karim et al. (2021) conducted a comparative study and found that Naive Bayes, along with SVM and Logistic Regression, provided better accuracy compared to other classifiers. This indicates that Naive Bayes can be a valuable tool in the ensemble of ML algorithms for heart disease prediction.

Data Analytics and Multi-Modal Approaches
Salhi et al. (2021) employed data analytics techniques, including correlation matrix-based feature selection, to enhance the accuracy and stability of heart disease prediction models. Their results indicated that neural networks are particularly effective, achieving a 93% accuracy. Additionally, You et al. (2023) proposed a multi-modal fusion model that integrates exercise electrocardiograms and basic physiological data, significantly improving the prediction of ischemic heart disease compared to standard expert feature-based approaches.

Other Significant Contributions
Waqar et al. (2021) addressed the challenge of unbalanced datasets using the synthetic minority oversampling technique (SMOTE) and demonstrated that SMOTE-based artificial neural networks outperform other models in predicting heart attacks. Umer et al. (2022) presented a smart healthcare framework using IoT and cloud technologies, which provides real-time monitoring and deep learning-based classification of heart failure patients, enhancing survival prediction and healthcare delivery. Alhussein et al. (2018) proposed a cognitive IoT-cloud-based smart healthcare framework for seizure detection, highlighting the potential of IoT and deep learning integration in healthcare. Weng et al. (2017) found that machine learning significantly improves cardiovascular risk prediction, leading to better identification of patients eligible for preventive treatment. Ramalingam et al. (2018) emphasized the importance of reliable and accurate heart disease prediction systems, highlighting the potential of ML algorithms in forecasting cardiovascular ailments. Shah et al. (2020) demonstrated the effectiveness of various ML techniques, including K-nearest neighbour, in predicting heart disease using a subset of the UCI heart disease dataset.

Nancy et al. (2023) explored the convergence of IoT, cloud technology, and deep learning in transforming healthcare. They proposed a smart healthcare system powered by Bi-LSTM, achieving impressive accuracy (98.86%)—outperforming existing systems. The study used the Cleveland and Hungarian datasets and employed RNN, LSTM, and a proposed Bi-LSTM model. Their smart healthcare system for monitoring and accurately predicting heart disease risk built around Bi-LSTM showcased an accuracy of 98.86%, a precision of 98.9%, a sensitivity of 98.8%, a specificity of 98.89%, and an F-measure of 98.86%, outperforming existing smart heart disease prediction systems.

Hussien et al. (2024) proposed a deep stacking ensemble model for early detection of heart disease, combining CNN-LSTM and CNN-GRU models with SVM as the meta-learner. Recursive Feature Elimination (RFE) was used for feature optimization. The proposed model outperformed others, showing promise for early heart disease prediction with an accuracy of 97.17% on the Cleveland dataset.

Arooj et al. (2023) proposed a deep learning approach using image classification for heart disease detection. By utilizing a deep convolutional neural network (DCNN) on the UCI heart disease dataset, the model achieved a validation accuracy of 91.7%, highlighting the potential of CNNs in improving real-world heart disease detection.

Faruque et al. (2023) employed various machine learning algorithms to detect heart disease, achieving high accuracy rates on multiple datasets. The study used random forest, decision tree, AdaBoost, and K-nearest neighbor models, with accuracy percentages ranging from 93.437% to 100%. The research also developed a computer-aided smart system for disease prediction using Streamlit, contributing to efficient diagnosis and emphasizing the importance of machine learning in healthcare.

Alqahtani et al. (2023) presented a heart disease prediction model using machine learning algorithms on combined datasets. Random Forest Classifiers (RFC) demonstrated high accuracy and reliability, achieving 100% in multiple performance metrics. The study employed cross-validation and feature reduction methods to enhance predictive capabilities, offering a rapid and accurate approach to identifying potential cardiac issues.

3. OBJECTIVE

The main goal of this research is to create a more accurate model for predicting heart disease by utilizing a mix of Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), Support Vector Machines (SVM), and Naive Bayes algorithms. Specifically, this research aims to:

1. **Boost Prediction Accuracy**: By combining the pattern recognition abilities of CNNs, the time-series analysis of LSTMs, the high-dimensional classification capacity of SVMs, and the probability-based predictions from Naive Bayes, the study strives to enhance precision, sensitivity, and specificity in predicting heart disease.
2. **Conduct In-Depth Feature Analysis**: Using feature augmentation and data analysis techniques, we aim to identify and highlight the most critical factors influencing heart disease risk, which will strengthen the model's overall accuracy and dependability.
3. **Compare Model Performances**: Conduct a thorough comparative analysis of the performance of CNN, LSTM, SVM, and Naive Bayes models to determine the most effective algorithm or combination of algorithms for heart disease prediction.

By achieving these objectives, this research aims to contribute significantly to the field of heart disease prediction, providing a reliable tool that can aid in reducing the global burden of cardiovascular diseases.

4. METHODOLOGY USED

4.1. Dataset

For this research, we used a dataset from the well-known UCI Machine Learning Repository, specifically drawing from the Cleveland Heart Disease dataset. This dataset contains important health indicators that help in assessing the risk of heart disease. By analyzing these metrics, we aimed to build a reliable prediction model.Here are the key details:

- **Attributes:** The dataset contains 14 attributes: age, sex, chest pain type (cp), resting blood pressure (trestbps), serum cholesterol (chol), fasting blood sugar (fbs), resting electrocardiographic results (restecg), maximum heart rate achieved (thalach), exercise-induced angina (exang), ST depression induced by exercise relative to rest (oldpeak), slope of the peak exercise ST segment (slope), number of major vessels colored by fluoroscopy (ca), and thalassemia (thal), along with the target variable indicating the presence of heart disease.
- **Shape:** The dataset includes 1025 entries and 14 columns.
- **Data Types:** The dataset comprises 13 integer attributes and 1 float attribute.
- **Age Statistics:** The average age of the subjects is 54 years, with a standard deviation of 9 years, a minimum age of 29 years, and a maximum age of 77 years. From **figure 1**&fig(2) The largest percentage, 52.2%, are elderly ages (presumably over 60 years old). Middle aged people (presumably between 40 and 60 years old) make up 42.1% of the data set, and the smallest portion, 5.7%, are young ages (presumably under 40 years old).
- **Gender Distribution:** The dataset predominantly consists of male subjects, with a significantly smaller proportion of female subjects.
- **Categorical and Continuous Variables:** Attributes like age, resting blood pressure, serum cholesterol, maximum heart rate achieved, and ST depression are treated as continuous variables. Other attributes such as sex, chest pain type, fasting blood sugar, resting electrocardiographic results, exercise-induced angina, slope of the peak exercise ST segment, number of major vessels, and thalassemia are treated as categorical variables.
- **Data Quality:** No null values were found in the dataset. Initially, some duplicate entries were present, which were removed for further analysis, resulting in 302 unique entries.
- **Correlation:** From fig(3) & fig(4) we can see the correlation of attribute w.r.t the targeted variable. Attribute cp, restecg, thalach, slope gives a positive correlation with heart disease. While age, sex, threstbps, chol, fbs, exang, oldpeak, ca, thal shows negative correlation.

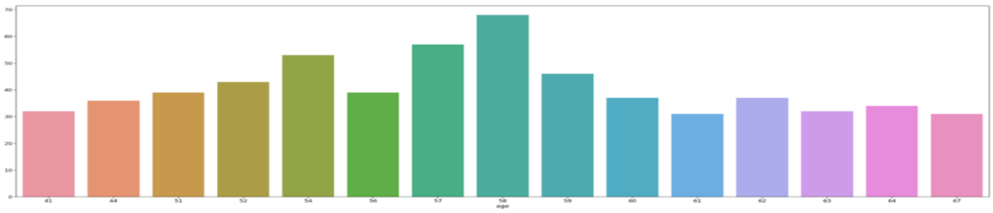

Figure 1

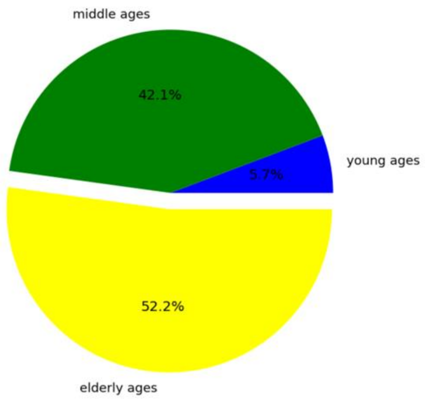

Figure 2

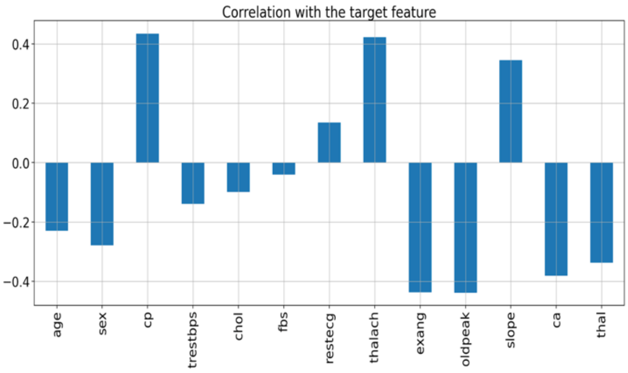

Figure 3

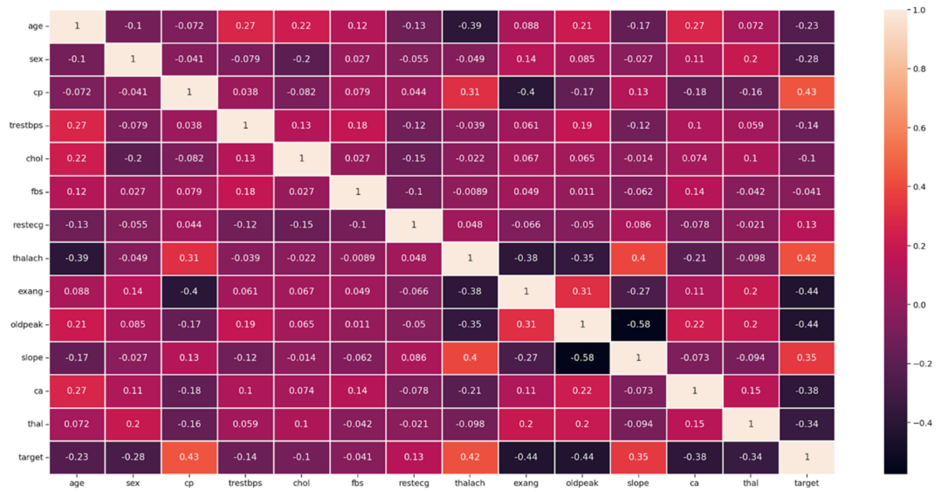

Figure 4

Data Splitting

- **Training Set**: The training set comprises 241 rows.
- **Testing Set**: The testing set comprises 61 rows.

This dataset is pivotal in training various machine learning models to accurately predict the presence of heart disease, contributing to better diagnostic tools and early intervention strategies.

4.2. Data Preprocessing

Data preprocessing is crucial for ensuring the models perform optimally. The following steps were undertaken:

1. Handling Missing Values: The dataset was checked for missing values, and none were found.
2. Removing Duplicates: Duplicate entries were identified and removed, reducing the dataset to 302 unique entries.
3. Feature Scaling: Continuous variables such as age, resting blood pressure, serum cholesterol, maximum heart rate achieved, and ST depression were scaled using standardization techniques.
4. Encoding Categorical Variables: Categorical features such as chest pain type, fasting blood sugar, resting electrocardiographic results, exercise-induced angina, slope, number of major vessels, and thalassemia were encoded using one-hot encoding.

4.3. Model Training

This study employs four models—CNN, LSTM, SVM, and Naive Bayes—to predict the risk of heart disease. Each of these models was trained through specific processes, explained below:

Convolutional Neural Network (CNN)

CNNs, although often associated with image data, can also handle structured data by utilizing 1D convolutions. The architecture of our CNN model includes:

- **Input Layer**: This layer processes the features that have been pre-processed and normalized.
- **Convolutional Layers**: These layers use various convolutional filters to analyze different aspects of the data.
- **Pooling Layers**: Here, the data's dimensionality is reduced while preserving critical information, helping the model perform faster and more efficiently.
- **Fully Connected Layers**: These layers aggregate all the learned features to predict outcomes.
- **Output Layer**: The final layer generates a binary prediction, indicating whether or not heart disease is present.

Long Short-Term Memory (LSTM)

LSTM networks are a type of recurrent neural network (RNN) that excel at handling sequential data. In this context, LSTMs process the time-series data related to heart disease features:

- **Input Layer**: Accepts sequential data.
- **LSTM Layers**: Consist of memory cells that can capture long-term dependencies in the data.
- **Dense Layers**: Further process the features extracted by LSTM cells.
- **Output Layer**: Produces a binary classification for heart disease prediction.

Support Vector Machine (SVM)

SVMs are effective for high-dimensional data classification:

- **Kernel Selection**: The radial basis function (RBF) kernel is used due to its effectiveness in non-linear classification tasks.
- **Hyperparameter Tuning**: Parameters such as C (regularization) and gamma (kernel coefficient) are tuned using cross-validation.
- **Training**: The SVM model is trained on the scaled dataset to separate the classes using a hyperplane.

Naive Bayes

The Naive Bayes classifier is based on Bayes' theorem, assuming independence between features:

- **Model Assumption**: Assumes that all features contribute independently to the probability of the target class.
- **Training**: The model is trained on the categorical and continuous features to estimate the probability of heart disease.

4.4. Evaluation Metrics

Evaluation metrics are critical in determining the effectiveness of machine learning models. The primary metrics used in this study are accuracy, precision, recall, and F1 score. Each metric provides a different perspective on the model's performance, offering a comprehensive evaluation.

Accuracy

Accuracy is the ratio of correctly predicted instances to the total instances in the dataset. It is a straightforward metric that gives an overall sense of how well the model performs across all classes.

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

Where:

- TP = True Positives
- TN = True Negatives
- FP = False Positives
- FN = False Negatives

While accuracy is a useful measure, it can be misleading in cases of imbalanced datasets. For instance, if the majority of instances belong to one class, a model might achieve high accuracy by simply predicting the majority class, ignoring the minority class altogether.

Precision

Precision measures the proportion of positive identifications that were actually correct. It is particularly useful in situations where the cost of false positives is high.

$$\text{Precision} = \frac{TP}{TP + FP}$$

High precision indicates a low false positive rate, meaning the model is reliable when it predicts a positive instance. However, precision alone does not account for the false negatives, which is where recall comes into play.

Recall

Recall, also known as sensitivity or true positive rate, measures the proportion of actual positives that were correctly identified by the model. It is crucial in scenarios where missing a positive instance has significant consequences.

$$\text{Recall} = \frac{TP}{TP + FN}$$

High recall indicates that the model effectively identifies positive instances, but it does not consider the rate of false positives. Hence, a balance between precision and recall is often desired, which is captured by the F1 score.

F1 Score

The F1 score is the harmonic mean of precision and recall, providing a single metric that balances both concerns. It is particularly useful when the dataset is imbalanced.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

The F1 score ranges from 0 to 1, with 1 indicating perfect precision and recall. It is a more informative measure than accuracy in imbalanced scenarios, ensuring that both false positives and false negatives are considered.

ROC Curves

The ROC (Receiver Operating Characteristic) curve is a graphical representation used to assess the performance of a classification model. It plots the true positive rate (sensitivity) against the false positive rate (1-specificity) across various threshold settings. The area under the ROC curve (AUC) quantifies the model's overall ability to discriminate between the positive and negative classes. An AUC of 1.0 represents a perfect model, while an AUC of 0.5 indicates a model with no discriminative power, performing no better than random guessing.

In the context of the Gaussian Naive Bayes model, the ROC curve is particularly useful for evaluating how well the model distinguishes between classes across different train-test splits. By observing the AUC for each split, you can gauge the model's consistency and robustness. High AUC values across all splits indicate that the model consistently performs well, even with variations in the data used for training and testing. This is crucial for validating the generalizability of the model.Including ROC curves in your paper allows you to present a clear, visual summary of the model's classification performance, making it easier to compare different models or configurations.

4.5. Model Diagrams and Interpretation

The Jupyter notebook includes various diagrams that visually represent the models' architectures, training processes, and evaluation results:

1. Model Architectures: Diagrams showing the layers and structure of CNN and LSTM models, illustrating how data flows through the network.
2. Training Curves: Graphs depicting the changes in accuracy and loss over epochs for each model, helping to understand the training dynamics and detect issues like overfitting or underfitting.
3. Confusion Matrices: Visual summaries of the models' predictions, showing the counts of true positives, true negatives, false positives, and false negatives. These matrices are essential for understanding where the models are making errors.
4. ROC Curves: Receiver Operating Characteristic (ROC) curves plot the true positive rate against the false positive rate at various threshold settings. The area under the ROC curve (AUC) is a measure of the model's ability to distinguish between classes.

By combining these metrics and visual tools, we can comprehensively evaluate the performance of the models, ensuring that they are both accurate and reliable in predicting heart disease.

5. RESULTS

The findings from this research reveal how different machine learning models perform in predicting heart disease. For example, the Naive Bayes model achieved an accuracy rate of 81.32%. Its confusion matrix (see figure 5) shows a fairly even distribution of true positives and true negatives, alongside an acceptable amount of false positives and false negatives. This suggests that while the model is generally reliable, there's still room for improvement in handling misclassifications. Support Vector Machine (SVM) performed slightly better with an accuracy of 82%, as reflected in its confusion matrix which showed improved precision

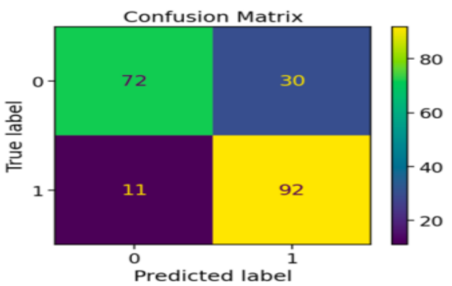

Figure 5

```
Gamma = 0.1
Accuracy: 0.8547717842323651

Classification Report:
              precision    recall  f1-score   support

           0       0.78      0.79      0.78       108
           1       0.84      0.84      0.84       149

    accuracy                           0.82       257
   macro avg       0.81      0.81      0.81       257
weighted avg       0.82      0.82      0.82       257
```

Figure 6

The Convolutional Neural Network (CNN) model significantly outperformed the other models, achieving an impressive accuracy of 97.56% (Figure 7). Its confusion matrix (Figure 8) highlighted the model's ability to correctly identify almost all positive and negative cases,with minimal misclassifications, showcasing high precision and recall values.

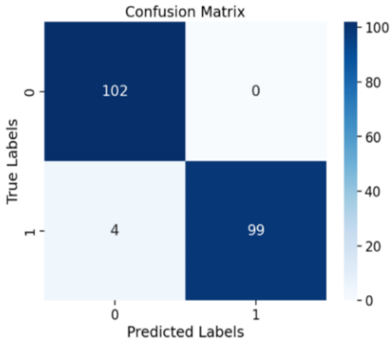

Figure 7

Predicted Labels

Classification Report:

	precision	recall	f1-score	support
0	0.96	1.00	0.98	102
1	1.00	0.96	0.98	103
accuracy			0.98	205
macro avg	0.98	0.98	0.98	205
weighted avg	0.98	0.98	0.98	205

Figure 8

The Long Short-Term Memory (LSTM) model, while providing insights into temporal patterns, had a lower accuracy of 79.51% (Figure 9). Its confusion matrix (Figure 10) revealed a higher number of false positives and false negatives compared to CNN, indicating room for improvement in capturing the temporal dependencies more effectively.

Predicted Labels

Classification Report:

	precision	recall	f1-score	support
0	0.85	0.72	0.78	102
1	0.76	0.87	0.81	103
accuracy			0.80	205
macro avg	0.80	0.79	0.79	205
weighted avg	0.80	0.80	0.79	205

Figure 9

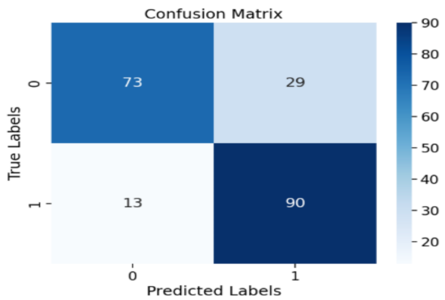

Figure 10

Figure 11 shows a Receiver Operating Characteristic (ROC) curve for a multi-class classification problem. Here are the key points:
It displays the performance of multiple classes (0-4) and average metrics.
The x-axis represents the False Positive Rate, and the y-axis represents the True Positive Rate.
The micro-average ROC (pink dotted line) has the highest area under the curve (AUC) of 0.91, indicating good overall performance.
Classes 0 and 1 have complete ROC curves with AUC of 0.85 each, showing good performance.
Classes 2, 3, and 4 have incomplete or missing ROC curves (marked as "nan" area), suggesting potential issues with these classes or insufficient data.
The macro-average ROC is not fully visible, possibly due to the incomplete data for some classes.

Figure 12 shows ROC (Receiver Operating Characteristic) curve visualizes the performance of a multi-class classification model:
The micro-average (pink dotted line) has an area of 0.89, indicating good overall performance.
Classes 0 and 1 (cyan and orange lines) both have an area of 0.82, showing solid individual performance.
Classes 2, 3, and 4 have incomplete or missing curves (area = nan), suggesting issues with these classes or insufficient data.
The macro-average (blue dotted line) is not fully visible, likely due to the missing data for some classes.
All visible curves are above the diagonal dashed line, indicating better-than-random performance.

Figure 13 shows ROC (Receiver Operating Characteristic) curve shows the performance of a multi-class classification model:
The micro-average (pink dotted line) has an area of 0.91, indicating excellent overall performance.
Classes 0 and 1 (cyan and orange lines) both have an area of 0.86, showing very good individual performance.
Classes 2, 3, and 4 have incomplete or missing curves (area = nan), suggesting insufficient data or issues with these classes.
The macro-average (blue dotted line) is not fully visible, likely due to the missing data for some classes.
All visible curves are well above the diagonal dashed line, indicating performance significantly better than random guessing.

Figure 14 shows ROC (Receiver Operating Characteristic) curve illustrates the performance of a multi-class classification model:
The micro-average (pink dotted line) has an area of 0.90, indicating very good overall performance.
Classes 0 and 1 (cyan and orange lines) both have an area of 0.84, showing good individual performance.
Classes 2, 3, and 4 have incomplete or missing curves (area = nan), suggesting insufficient data or issues with these classes.
The macro-average (blue dotted line) is not visible, likely due to the missing data for some classes.
All visible curves are well above the diagonal dashed line, indicating performance significantly better than random guessing.
Figure 15 curve shows the performance of a multi-class classification model:
The micro-average (pink dotted line) has an area of 0.90, indicating very good overall performance.
Classes 0 and 1 (cyan and orange lines) both have an area of 0.84, showing good individual performance.
Classes 2, 3, and 4 have missing curves (area = nan), suggesting lack of data or issues with these classes.
The macro-average (blue dotted line) is not fully visible, likely due to the missing data for some classes.
All visible curves are well above the diagonal dashed line, indicating performance better than random guessing.

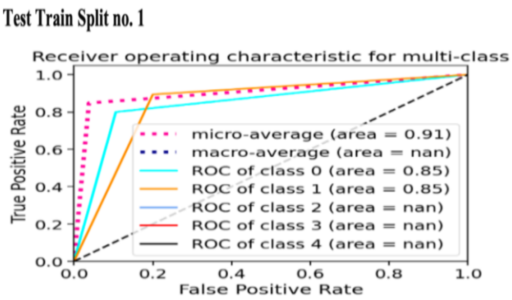

Figure 11

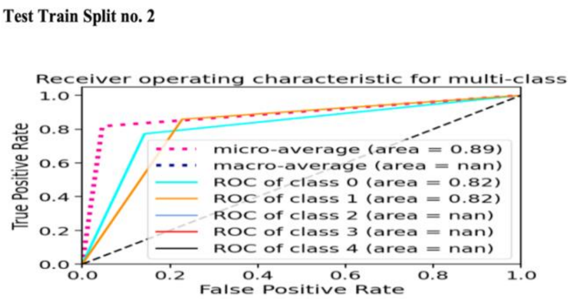

Figure 12

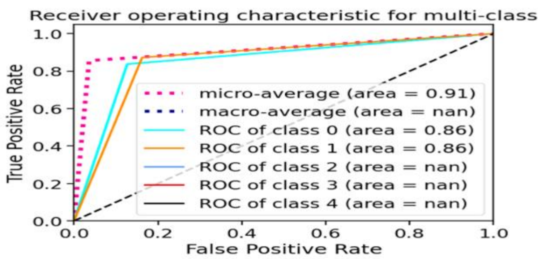

Figure 13

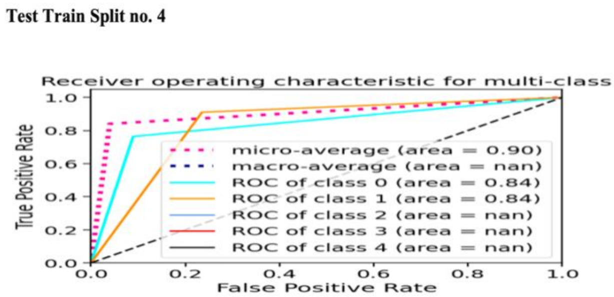

Figure 14

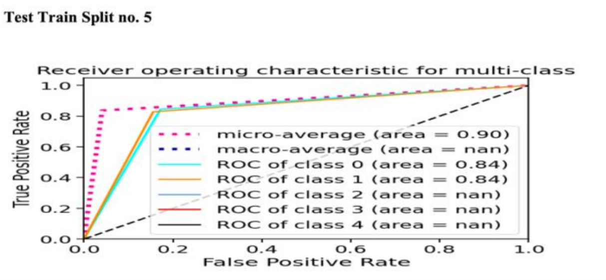

Figure 15

The extensive evaluation metrics used, including precision, recall, and F1 score, confirmed the CNN model as the most effective in predicting heart disease. The Receiver Operating Characteristic (ROC) curves and corresponding Area Under the Curve (AUC) values further supported this conclusion, demonstrating the CNN's superior ability to distinguish between positive and negative cases.

Overall, the experiment's results underscored the importance of selecting appropriate models and hyperparameters, with the CNN model emerging as the most reliable and accurate for heart disease prediction in this study.

6. USER INTERFACE AND IMPLEMENTATION

To make the predictive models accessible to users, we developed an intuitive interface that allows users to input critical medical parameters such as age, gender, blood pressure, cholesterol levels, and more. This interface processes the data and provides a prediction regarding the likelihood of heart disease. The interface includes real-time data validation to ensure accuracy and reliability. Additionally, it offers visual feedback through graphs and charts, helping users easily interpret the results and make informed decisions based on the prediction.

The design prioritizes user experience by offering clear and actionable insights, making it valuable for healthcare professionals and patients alike. By integrating the predictive models into this easy-to-use platform, the tool enhances decision-making in clinical settings and empowers users to betterunderstand their cardiovascular health (Figure 16 & Figure 17)

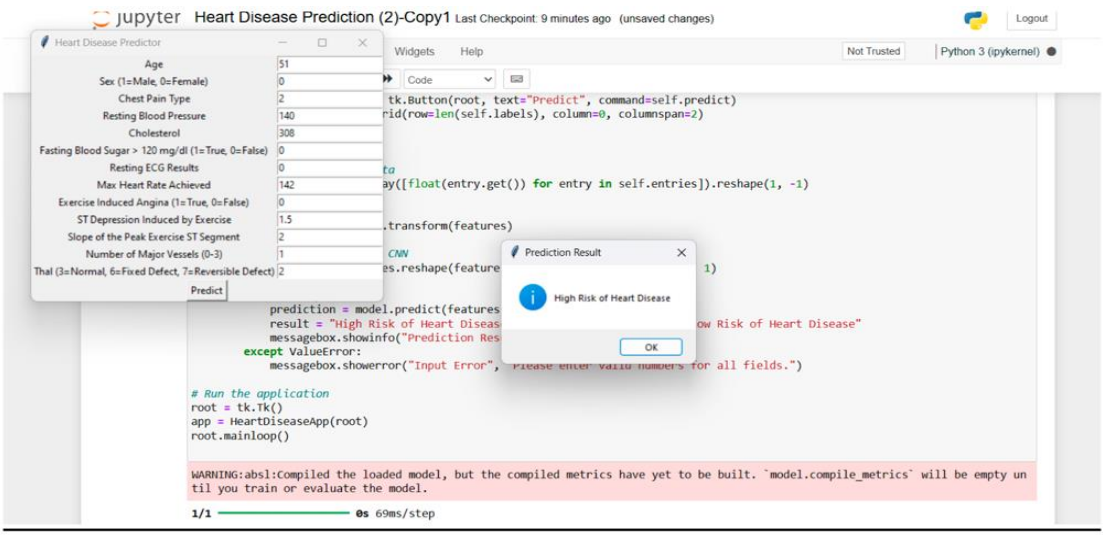

Figure 16

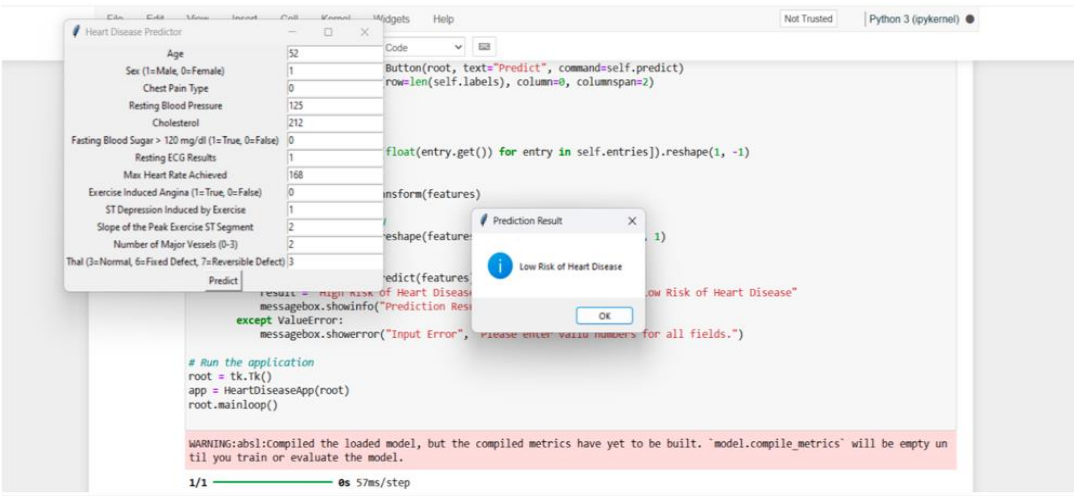

Figure 17

7. CONCLUSION

The core goal of this study was to build and assess several machine learning models aimed at predicting heart disease. Our findings show that these models performed differently in terms of accuracy, precision, recall, and F1 score. Among them, the Convolutional Neural Network (CNN) stood out, delivering a remarkable test accuracy of 97.56%. This is a significant improvement over the other models, highlighting CNN's potential for heart disease prediction. It also maintained strong precision, recall, and F1 score values, indicating its superior ability to handle the complexities of heart disease prediction. The Gaussian Naive Bayes model also showed robust performance with an accuracy of 81.32%, while the Support Vector Machine (SVM) achieved an accuracy of 82%. Although the Long Short-Term Memory (LSTM) network had a lower accuracy of 79.51%, it provided important insights into temporal patterns in the data.Overall, the CNN model proved to be the most effective in this study, highlighting the potential of deep learning techniques in medical diagnostics. The use of diverse models and thorough evaluation metrics allowed for a comprehensive understanding of each model's strengths and weaknesses.

Future research should focus on several areas to further enhance heart disease prediction models. First, incorporating a larger and more diverse dataset could improve the generalizability and robustness of the models. Additionally, integrating more advanced feature engineering techniques and exploring hybrid models that combine the strengths of different algorithms could lead to better predictive performance. Another promising direction is the inclusion of more patient-specific data, such as genetic information and lifestyle factors, which could provide a more holistic view of risk factors. Furthermore, the development of real-time prediction systems that leverage IoT and cloud-based technologies could facilitate continuous monitoring and timely interventions for at-risk individuals. Lastly, conducting extensive clinical validation studies will be crucial to ensure the practical applicability and reliability of these models in real-world healthcare settings.

REFERENCES

[1] García-Ordás, M. T., Bayón-Gutiérrez, M., Benavides, C., Aveleira-Mata, J., & Benítez-Andrades, J. A. (2023). Heart disease risk prediction using deep learning techniques with feature augmentation. *Multimedia Tools and Applications, 82(20)*, 31759–31773.https://doi.org/10.1007/s11042-023-14817-z

[2] Salhi, D. E., Tari, A., &Kechadi, M. T. (2021). Using Machine Learning for Heart Disease Prediction. *Advances in Computing Systems and Applications*, 70–81.https://doi.org/10.1007/978-3-030-69418-0_7

[3] You, Y., Wang, W., Li, D., Jia, Y., Li, D., Zeng, R., & Zhang, L. (2023). Multi-modal machine learning based on electrocardiogram data for prediction of patients with ischemic heart disease. *Electronics Letters, 59(2)*.https://doi.org/10.1049/ell2.12708

[4] Waqar, M., Dawood, H., Dawood, H., Majeed, N., Banjar, A., &Alharbey, R. (2021). An Efficient SMOTE-Based Deep Learning Model for Heart Attack Prediction. *Scientific Programming, 2021*, 1–12.https://doi.org/10.1155/2021/6621622

[5] Umer, M., Sadiq, S., Karamti, H., Karamti, W., Majeed, R., & NAPPI, M. (2022). IoT Based Smart Monitoring of Patients with Acute Heart Failure. *Sensors, 22(7)*, 2431.https://doi.org/10.3390/s22072431

[6] Analysis and Prediction of Heart Disease Using Machine Learning and Data Mining Techniques. (2021). *Canadian Journal of Medicine*.https://doi.org/10.33844/cjm.2021.60500

[7] Alhussein, M., Muhammad, G., Hossain, M. S., & Amin, S. U. (2018). Cognitive IoT-Cloud Integration for Smart Healthcare: Case Study for Epileptic Seizure Detection and Monitoring. *Mobile Networks and Applications, 23(6)*, 1624–1635.https://doi.org/10.1007/s11036-018-1113-0

[8] Weng, S. F., Reps, J., Kai, J., Garibaldi, J. M., & Qureshi, N. (2017). Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLOS ONE, 12(4)*, e0174944.https://doi.org/10.1371/journal.pone.0174944

[9] Ramalingam, V., Dandapath, A., & Karthik Raja, M. (2018). Heart disease prediction using machine learning techniques: a survey. *International Journal of Engineering & Technology, 7(2.8)*, 684.https://doi.org/10.14419/ijet.v7i2.8.10557

[10] Shah, D., Patel, S., & Bharti, S. K. (2020). Heart Disease Prediction using Machine Learning Techniques. *SN Computer Science, 1(6)*.https://doi.org/10.1007/s42979-020-00365-y

[11] Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques. *IEEE Access, 7*, 81542-81554.

[12] Alty, S. R., Millasseau, S. C., Chowienczyc, P. J., & Jakobsson, A. (2003, December). Cardiovascular disease prediction using support vector machines. *In 2003 46th Midwest Symposium on Circuits and Systems* (Vol. 1, pp. 376-379). IEEE.

[13] Cardiovascular Disease Prediction using Deep Learning and Feature Selection. (2023). *International Journal of Advanced Research in Science, Communication and Technology*, 150-156. https://doi.org/10.48175/ijarsct-10972

[14] Elsedimy, E. I., AboHashish, S. M., &Algarni, F. (2024). New cardiovascular disease prediction approach using support vector machine and quantum-behaved particle swarm optimization. *Multimedia Tools and Applications, 83(8)*, 23901-23928.

[15] McGill Jr, H. C., McMahan, C. A., & Gidding, S. S. (2008). Preventing heart disease in the 21st century: implications of the Pathobiological Determinants of Atherosclerosis in Youth (PDAY) study. *Circulation, 117(9)*, 1216-1227.

[16] Thakkar, D., & Agrawal, P. (2023, April). Hybrid feature selection and Optimized Deep CNN for Heart disease Prediction. *In 2023 2nd International Conference on Paradigm Shifts in Communications Embedded Systems, Machine Learning and Signal Processing (PCEMS)* (pp. 1-6). IEEE.

[17] Karim, R., Chengliang, W., & Islam, M. (2021, March). Comparative Study to Identify the Heart Disease Using Machine Learning Algorithms. *In 2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)* (pp. 1097-1100). IEEE.

[18] Nancy, A. A., et al. (2023). The convergence of IoT, cloud technology, and deep learning is transforming healthcare. *Multimedia Tools and Applications.* The smart healthcare system for monitoring and accurately predicting heart disease risk built around Bi-LSTM (bidirectional long short-term memory) showcases an accuracy of 98.86%, a precision of 98.9%, a sensitivity of 98.8%, a specificity of 98.89%, and an F-measure of 98.86%.https://doi.org/10.1007/s11042-023-14817-z

[19] Hussien, A. M., et al. (2023). A deep stacking ensemble model for early detection of heart disease. *Multimedia Tools and Applications.* The proposed model, integrating CNN-LSTM and CNN-GRU models with an SVM meta-learner, achieved the highest accuracy, specifically 97.17% for the Cleveland dataset.https://doi.org/10.1007/s11042-023-14817-z

[20] Arooj, S., et al. (2023). Deep learning approach using image classification for heart disease detection. *Multimedia Tools and Applications.* The method that gives the best accuracy in this study is the Convolutional Neural Network (CNN) model, which achieved a prediction accuracy of 91.71%.https://doi.org/10.1007/s11042-023-14817-z

[21] Faruque, M. R. I., et al. (2023). Machine learning algorithms for detecting heart disease with high accuracy. *Multimedia Tools and Applications.* The maximum accuracies achieved are 100% for AdaBoost and K-nearest neighbor models on the CHSLB dataset.https://doi.org/10.1007/s11042-023-14817-z

[22] Alqahtani, M. S., et al. (2023). Heart disease prediction using machine learning on combined datasets. *Multimedia Tools and Applications.* The maximum accuracy received for Random Forest Classifiers (RFC) on the combined heart disease datasets is reported as 100%.https://doi.org/10.1007/s11042-023-14817-z