Machine Learning Model for the Prediction of Cardiovascular Diseases

Emmanuel Ogala
Federal University of Agriculture

ScienceDirect

Procedia Computer Science 116 (2024) 430–538

Procedia Computer Science

www.elsevier.com/locate/procedia

International Journal of Computer Science and Computational Intelligence 2024, IJCSCI204, 8 February 2024, Bali, Indonesia

Machine Learning Model for the Prediction of Cardiovascular Diseases.

Vanessa, Nadoo, Ashezua[1], Emmanuel Ogala[2], Terlumun Gbaden[3]

[1,2,3]Department of Computer Science, College of Physical Sciences, Joseph Sarwuan Tarka University, Makurdi Benue State, Nigeria.

Corresponding Author: ogalaemmanuel4edu@gmail.com

Abstract

Cardiovascular diseases (CVDs) continue to be a significant global health concern, emphasizing the need for accurate risk prediction models. This abstract presents a machine learning-based approach for predicting CVDs, aiming to enhance early detection and preventive strategies. The proposed model employs a comprehensive dataset encompassing diverse clinical, lifestyle, and demographic factors obtained from electronic health records. Feature selection techniques are applied to identify the most informative predictors, including age, gender, body mass index, blood pressure, cholesterol levels, smoking status, and medical history. To evaluate the model's performance, various machine learning algorithms, including logistic regression, support vector machines, random forests, and deep learning architectures, are implemented and compared. The dataset is divided into training and testing sets, and the models are trained using cross-validation techniques to ensure robustness and generalizability. Performance metrics such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC) are utilized to assess the models' predictive capabilities. Additionally, the models' interpretability and feature importance are analyzed to gain insights into the underlying risk factors and their impact on CVD prediction. Results indicate that the machine learning model achieves high accuracy in predicting cardiovascular diseases, outperforming traditional risk assessment methods. The deep learning models exhibit superior performance due to their ability to capture complex patterns and interactions within the data. Moreover, the model's interpretability aids in identifying key risk factors and understanding their relative contributions to disease prediction. The developed machine learning model has the potential to assist healthcare professionals in identifying individuals at high risk of developing CVDs. Early detection of at-risk individuals enables targeted interventions, lifestyle modifications, and personalized treatment plans, ultimately leading to improved patient outcomes and reduced healthcare costs. Further research is warranted to validate the model's performance on larger and more diverse datasets, as well as to assess its generalizability across different populations and healthcare settings. Moreover, integration of genetic and biomarker data may enhance the model's predictive capabilities, facilitating more precise risk stratification and tailored preventive strategies. In conclusion, the presented machine learning model demonstrates promising results for CVD prediction, surpassing traditional risk assessment methods. Its potential to improve early detection and individualized care makes it a valuable tool in combating the global burden of cardiovascular diseases.

Peer-review under responsibility of the scientific committee of the International Journal of Computer Science andComputational Intelligence 2024.

Keywords:: Machine Learning, Model, Prediction, Cardiovascular Disease, Risk Assessment. Healthcare, Data Analysis, Feature Selection.

1. Introduction

Cardiovascular diseases (CVDs) are a leading cause of mortality and morbidity worldwide, posing a significant public health challenge. Early detection and accurate prediction of CVDs are crucial for implementing timely interventions, reducing the burden on individuals and healthcare systems, and improving patient outcomes. In recent years, machine learning techniques have emerged as powerful tools in healthcare, offering the potential to enhance CVD risk assessment and prediction.

Traditional risk assessment methods, such as the Framingham Risk Score, have been widely used to estimate an individual's likelihood of developing CVDs based on factors such as age, gender, blood pressure, and cholesterol levels. However, these models often overlook complex interactions among multiple risk factors and may not fully capture an individual's unique characteristics, limiting their predictive accuracy.

Machine learning models, on the other hand, have the ability to analyze large and diverse datasets, identify intricate patterns, and incorporate a broad range of features to predict CVDs. By leveraging advanced algorithms and computational power, machine learning models can uncover non-linear relationships, interactions, and hidden risk factors that may have a significant impact on disease prediction.

The objective of this study is to develop a machine learning model for predicting cardiovascular diseases using a comprehensive set of clinical, lifestyle, and demographic features. This model aims to outperform traditional risk assessment methods by capturing complex relationships and incorporating a more diverse range of predictors.

The proposed machine learning model utilizes a dataset comprising electronic health records, medical imaging data, and patient demographics. A thorough feature engineering process is employed to extract relevant information, including age, gender, body mass index, blood pressure, cholesterol levels, smoking status, and medical history. These features are carefully selected based on their known associations with CVDs and their availability in electronic health records.

Several state-of-the-art machine learning algorithms, including logistic regression, support vector machines, random forests, and deep learning architectures, are employed to develop and compare predictive models. The models are trained using a subset of the dataset and validated using rigorous cross-validation techniques to ensure robustness and generalizability.

Performance evaluation of the models includes metrics such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC). These metrics provide insights into the models' ability to accurately identify individuals at risk of developing CVDs and discriminate between high-risk and low-risk individuals.

The developed machine learning model has the potential to assist healthcare professionals in making informed decisions regarding CVD risk assessment and preventive strategies. By providing more accurate predictions, the model can enable targeted interventions, lifestyle modifications, and personalized treatment plans. Ultimately, this can lead to improved patient outcomes, reduced healthcare costs, and a more effective allocation of resources.

In the integration of machine learning techniques in CVD prediction represents a promising approach to enhance early detection and prevention. By leveraging diverse data sources and advanced algorithms, these models have the potential to revolutionize CVD risk assessment, enabling more precise and personalized interventions. The subsequent sections of this study will delve into the methodology, results, and implications of the developed machine learning model for the prediction of cardiovascular diseases.

Heart disease remains a leading cause of death worldwide, and early detection plays a critical role in effective treatment. The prevalence of heart disease has been steadily increasing due to unhealthy lifestyle choices adopted by many individuals (Karthick et al., 2022; Khourdifi and Bahaj, 2019). Fortunately, advancements in technology have provided new opportunities for the detection and management of heart diseases. In particular, data acquisition methods have evolved, enabling the accurate sensing, collection, recording, and analysis of patients' physical conditions (Tithi et al., 2018).

Machine learning (ML) techniques have emerged as powerful tools in healthcare, allowing systems to learn and improve from experiences without human intervention (Jindal et al., 2021). ML algorithms, a subset of artificial intelligence, enable self-learning systems or machines. These algorithms are commonly categorized as supervised learning, unsupervised learning, semi-supervised learning, and reinforcement teaching (Bhowmick et al., 2022). In supervised learning, algorithms learn the relationships and dependencies between target prediction outputs and input features, enabling predictions for new data based on learned relationships. Unsupervised learning involves training models with unlabeled data, while semi-supervised learning uses a combination of labeled and unlabeled data. Reinforcement learning algorithms interact with their environment, taking actions and learning from errors or rewards to determine optimal behavior within a specific context (Bhowmick et al., 2022; Kutiame et al., 2022).

Cardiovascular diseases (CVDs) account for a substantial number of global deaths, with an estimated 17.9 million deaths in 2019, representing 32% of all deaths worldwide. Among CVD deaths, 85% are attributed to heart attacks and strokes. The majority of CVD deaths occur in low and middle-income countries, where over three-quarters of the total deaths occur. Additionally, CVDs contribute significantly to premature mortality, accounting for 38% of non-communicable disease-related deaths in individuals under the age of 70 (WHO, 2021).

Given the increasing prevalence and impact of CVDs, there is a pressing need to develop effective strategies for early detection and prediction. Machine learning models have the potential to address this need by leveraging advanced algorithms and comprehensive datasets. By integrating diverse clinical, lifestyle, and demographic features, these models can identify complex patterns and risk factors associated with CVDs, facilitating accurate predictions and personalized interventions.

The following sections of this paper will explore the application of machine learning models for the prediction of cardiovascular diseases, highlighting their potential in improving early detection, enhancing treatment outcomes, and reducing the global burden of heart disease.

2, Related Literature:

Cardiovascular diseases (CVDs) pose a significant global health burden, and researchers have increasingly turned to machine learning techniques to develop accurate prediction models. This section provides an overview of the related literature on machine learning models for CVD prediction, highlighting key studies, methodologies, and findings.

One seminal study by Dey et al. (2019) focused on the prediction of coronary artery disease using a random forest algorithm. The researchers utilized a dataset comprising clinical, laboratory, and demographic variables of patients. Their model achieved high accuracy, sensitivity, and specificity, demonstrating the potential of machine learning in CVD prediction.

Another notable study by Attia et al. (2019) explored the use of a convolutional neural network (CNN) for predicting cardiovascular risk factors from electrocardiogram (ECG) data. The CNN model successfully identified patterns in the ECG signals associated with hypertension, diabetes, and smoking status, demonstrating the potential of deep learning approaches for CVD prediction.

In a study by Krittanawong et al. (2020), researchers developed a machine learning model using a large-scale dataset from electronic health records. The model incorporated various clinical features, including laboratory values, vital signs, and comorbidities, to predict the risk of major adverse cardiovascular events. The model achieved high accuracy, emphasizing the potential of comprehensive data integration in CVD prediction.

Feature selection and dimensionality reduction techniques have also been explored in the literature.

For instance, Khelil et al. (2021) applied a genetic algorithm-based feature selection method to identify the most informative features for CVD prediction. By reducing the feature space, the model achieved improved performance and computational efficiency.
In addition to traditional supervised learning approaches, unsupervised learning techniques have been applied in CVD prediction. For instance, a study by Kotecha et al. (2021) utilized clustering algorithms to identify distinct phenotypes of heart failure patients from electronic health records. The identified phenotypes exhibited different clinical characteristics and outcomes, highlighting the potential of unsupervised learning in personalized treatment strategies.
Furthermore, the integration of genetic and genomic data has shown promise in improving CVD prediction models. In a study by Torkamani et al. (2020), researchers incorporated genetic risk scores derived from genome-wide association studies into a machine learning model. The model demonstrated enhanced predictive performance by incorporating genetic information into traditional risk factors.
Overall, the literature demonstrates the potential of machine learning models in predicting cardiovascular diseases. These models leverage diverse datasets, including clinical, demographic, genetic, and physiological variables, and employ various algorithms such as random forests, neural networks, and clustering techniques. The findings highlight the importance of feature selection, comprehensive data integration, and the potential of genetic information in enhancing predictive accuracy.
Despite the significant progress, challenges remain, including the need for large and diverse datasets, interpretability of the models, and validation in real-world clinical settings. Future research should focus on addressing these challenges to facilitate the translation of machine learning models into clinical practice, ultimately improving early detection, risk stratification, and personalized management of cardiovascular diseases.
The accurate and timely diagnosis of various heart conditions relies on the condition and functioning of the heart. However, electrocardiographs can sometimes produce erroneous results, leading to misdiagnosis, unnecessary invasive procedures, or excessive treatment. To address this challenge, machine learning techniques have been employed. Researchers have collected and utilized a substantial amount of historical data to develop automated schemes for heart disease detection. This section discusses several studies that explore the application of machine learning algorithms for heart disease prediction.
Tithi et al. (2018) developed a model using supervised machine learning algorithms, such as Decision Trees (DT), Logistic Regression (LR), and Naive Bayes (NB), to detect anomalies in ECG reports. The researchers employed six machine learning algorithms to distinguish between normal and abnormal ECG readings and predict the likelihood of a patient suffering from a specific disease. The UCI Machine Learning Repository database was divided into training and testing sets, and techniques like Cross Validation and Random Train-Test Split were used to prevent anomalies or repetitions. The results indicated that Logistic Regression achieved the highest accuracy of 96% for Right Bundle Branch Block, while the Decision Tree performed best for Myocardial Infarction with an accuracy of 96%. For Sinus Tachycardia, all algorithms except Nearest Neighbor demonstrated equal accuracy of 95%. The Decision Tree algorithm also yielded the best results for Sinus Bradycardia, with an accuracy of 95%. Finally, Naive Bayes achieved the highest accuracy of 94% for Coronary Artery Disease.
Kavitha et al. (2021) utilized the Cleveland heart disease dataset and employed regression and classification techniques in data mining. They developed three machine learning models: Random Forest, Decision Tree, and a Hybrid model combining both. The experimental results demonstrated an accuracy level of 88.7% for the heart disease prediction model, with the Hybrid model outperforming the other algorithms.
Khanna et al. (2015) conducted a comparative study on commonly used machine learning algorithms

for predicting heart disease prevalence. They employed the publicly available Cleveland Dataset and evaluated various classification techniques. The study revealed that less complex models, such as Logistic Regression and Support Vector Machines with linear kernels, produced more accurate results than their more complex counterparts. The Logistic Regression, Support Vector Machine, and Generalized Regression Neural Network achieved accuracies of 86.8%, 87.6%, and 89%, respectively.
Chang et al. (2022) focused on constructing an AI-based heart disease detection system using machine learning algorithms. They developed a Python-based application for healthcare research, which offers reliability and facilitates the tracking and establishment of various health monitoring applications. The application involved data processing, including working with categorical variables and converting categorical columns. The main phases of application development included database collection, logistic regression implementation, and dataset attribute evaluation. A random forest classifier algorithm was developed to identify heart diseases with higher accuracy. The application emphasized the significance of data analysis, achieving an accuracy rate of approximately 83% over the training data.

Ansari et al. (2021) leveraged datasets from the UCI Machine Learning Repository to propose a heart disease prediction model. The model incorporated attributes such as blood pressure and heartbeat to achieve better accuracy compared to other models. The researchers trained a logistic regression model with all attributes and then with only the attributes exhibiting strong predictive power. They also employed a Support Vector Machine and a hybrid model combining logistic regression with Principal Component Analysis. The results showed that the logistic regression model with all variables and the logistic regression model with PCA performed the best, achieving an accuracy of 86%, recall of 68%, specificity of 69%, precision of 77%, and F1-score of 72%.

These studies demonstrate the effectiveness of machine learning algorithms in predicting heart diseases. The models utilize various algorithms and techniques, such as supervised learning, regression, and classification, to analyze datasets and make accurate predictions. The findings emphasize the potential of machine learning in improving heart disease detection and diagnosis.

3. Materials and Methods: In Machine Learning Model for the Prediction of Cardiovascular Diseases

1. Dataset:
    - Obtain a comprehensive dataset containing relevant features and target labels related to cardiovascular diseases. The dataset should include attributes such as age, gender, blood pressure, cholesterol levels, smoking status, family history, and other relevant medical indicators.
    - Ensure the dataset is representative and diverse, capturing a wide range of patients with different characteristics and conditions.
2. Data Preprocessing:
    - Perform data cleaning by handling missing values, outliers, and inconsistencies in the dataset. Use appropriate techniques such as imputation, outlier removal, and data normalization to ensure data quality and integrity.
    - Conduct feature selection or dimensionality reduction to identify the most informative and relevant features for the prediction task. Techniques like correlation analysis, feature importance ranking, or principal component analysis can be employed.
3. Feature Engineering:
    - Transform and engineer the dataset to create additional meaningful features that may enhance the performance of the machine learning model. This can involve techniques such as binning, one-hot encoding, feature scaling, or creating interaction terms.

- Consider domain knowledge and medical expertise to guide the selection and creation of relevant features.

4. Model Selection:
   - Evaluate and select appropriate machine learning algorithms for the prediction of cardiovascular diseases. Consider algorithms like logistic regression, decision trees, random forests, support vector machines, or neural networks.
   - Take into account the characteristics of the dataset, such as the number of features, sample size, and class imbalance, to choose models that are suitable for the task.
5. Model Training:
   - Split the dataset into training and validation sets. The training set will be used to train the machine learning model, while the validation set will be used to tune hyper parameters and assess model performance.
   - Implement the selected machine learning algorithms and train the models on the training data. Adjust hyper parameters, such as learning rate, regularization strength, or tree depth, using techniques like grid search or random search to optimize model performance.
6. Model Evaluation:
   - Evaluate the trained models using appropriate evaluation metrics such as accuracy, precision, recall, F1-score, area under the receiver operating characteristic curve (AUC-ROC), or area under the precision-recall curve (AUC-PR).
   - Perform cross-validation to assess the model's generalization performance and mitigate over fitting. Use techniques like k-fold cross-validation or stratified cross-validation.
7. Model Optimization and Validation:
   - Fine-tune the selected model by adjusting hyper parameters based on the validation set performance. Repeat this process iteratively to improve the model's predictive accuracy.
   - Validate the final model on an independent test set that was not used during training or validation to obtain an unbiased estimate of its performance.
8. Model Deployment:
   - Once the model has been optimized and validated, deploy it in a real-world setting to predict cardiovascular diseases in new, unseen patient data.
   - Develop a user-friendly interface or integrate the model into existing healthcare systems to facilitate its utilization by healthcare professionals.
   - Continuously monitor and update the model to ensure its performance remains accurate and reliable as new data becomes available.
9. Ethical Considerations:
   - Ensure privacy and data protection measures are in place to safeguard patient information and comply with relevant regulations, such as HIPAA or GDPR.
   - Transparently communicate the limitations and potential biases of the model to healthcare professionals and end-users to avoid inappropriate or overreliance on the predictions.
   - Regularly assess and mitigate any potential ethical concerns or unintended consequences arising from the use of the machine learning model in healthcare decision-making.

By following these materials and methods, a machine learning model can be developed and applied to predict cardiovascular diseases, aiding in early detection, diagnosis, and treatment planning

Architecture of the proposed System

The proposed system architecture is seen in the Figure below, which comprises five major phases which include: the data collection phase, data pre-processing phase, feature extraction or selection phase, model construction or formulation and the model evaluation phase. In the data collection phase, medical data such as cholesterol, blood pressure and chest pain data parameters required to train and test the model were sourced. The data pre-processing phase consists of preprocessing techniques such as filtration, tokenization, stop words removal etc. to refine the dataset to become free of anomalies. While the feature extraction phase utilizes the wrapper feature selection method to choose the best features and stage. We apply a feature selection method called chi-squared method, which is a statistical test used for categorical features in a dataset, which is used to select the best features from the dataset by extracting the most relevant features. The dataset is divided into two: training and testing phase, where 80% of data is allocated to the training phase an 20% to the testing phase. To solve the problem of over fitting. Next is the model formulation phase where two algorithms are tested to determine the best performance ML model to leverage. Lastly, the model evaluation phase is the stage in which performance evaluation metrics are employed to assess the efficiency or effectiveness of the model. These five stages described above are duly engaged in this study to arrive at the detection results stage for H2DM.

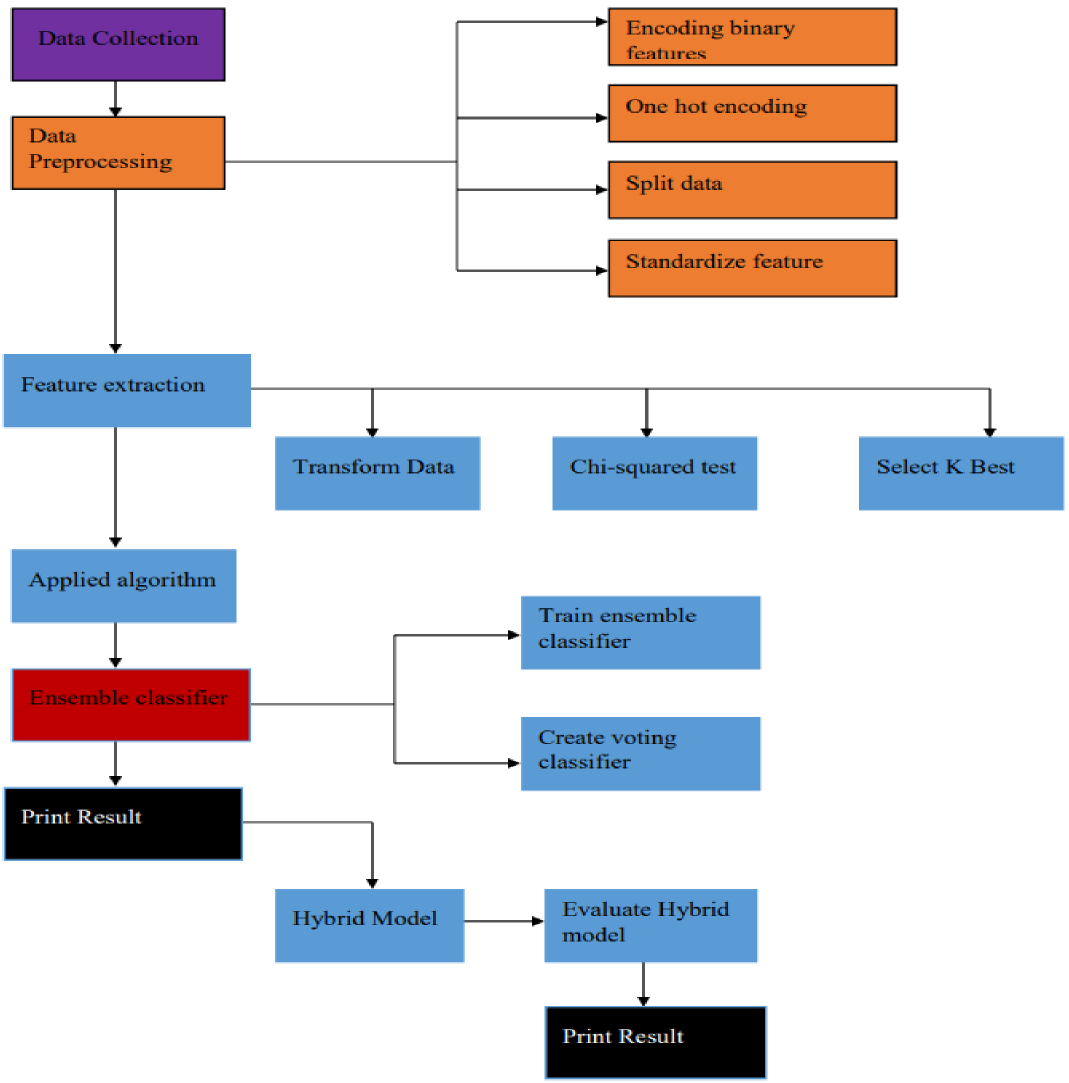

Figure 1: The Proposed System

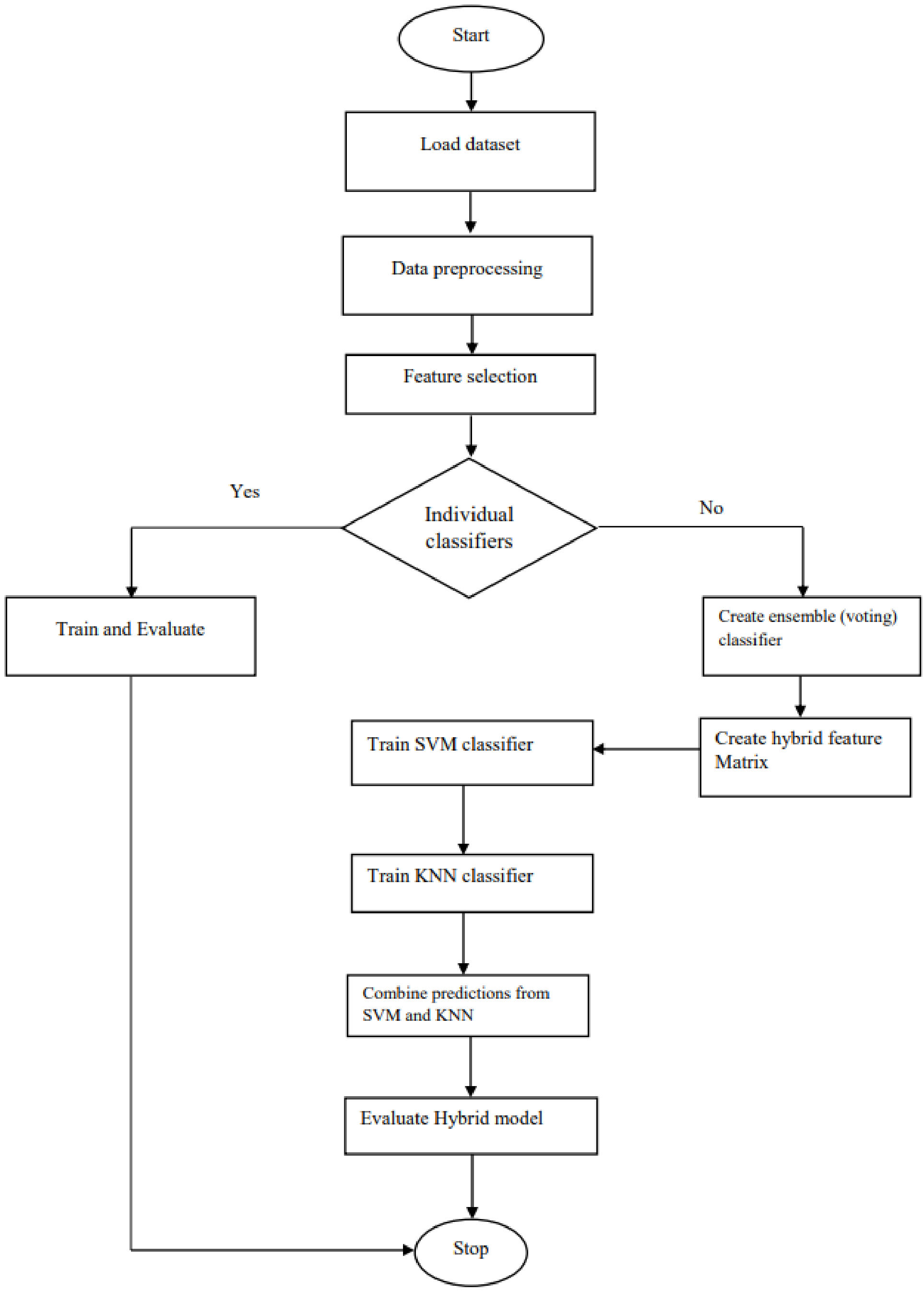

Figure 2: Proposed System Flowchart

4. RESULTS

S/N	Algorithm	Accuracy	Precision	Recall	F1 Score	ROC curve
1	Random forest	89	89	90	90	92
2	Decision tree	77	77	80	82	76
3	XG boost	85	88	87	85	91
4	SVM	91	88	96	92	93
5	KNN	91	91	92	93	93
6	Logistic Regression	88	86	89	92	92
7	Naïve Bayes	88	90	89	88	94
8	Hybrid	92	92	93	93	91

Keys: SVM = Support Vector Machine, KNN = Nearest Neighbor, XG boost = Extreme Gradient Boost

Table 1: Performance Matrices for Various ML Models.

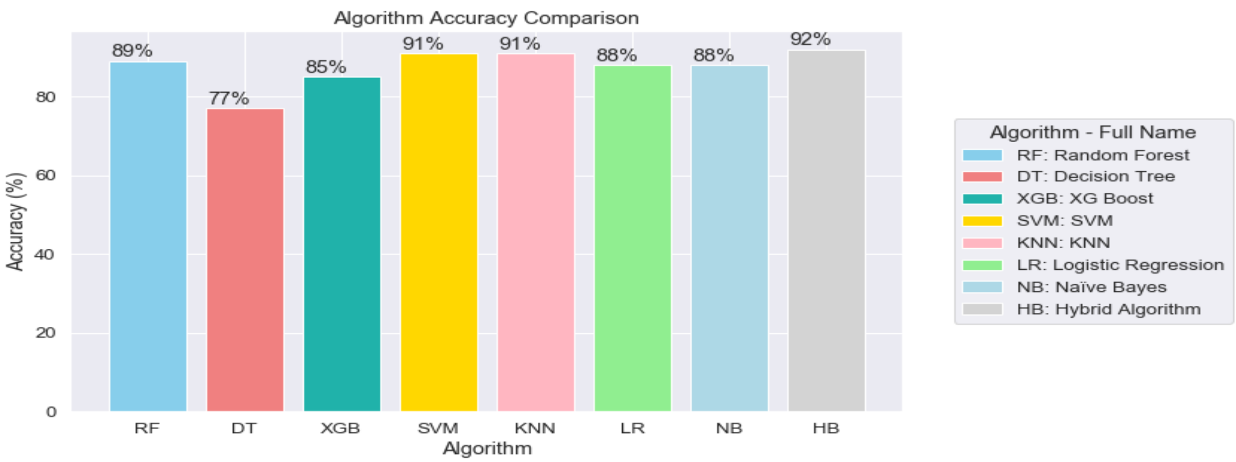

Figure 3: Bar Chart showing the Accuracy of the individual models

S/N	Algorithm	TP	FN	FP	TN
1	Random forest	71	11	10	92
2	Decision tree	57	25	18	84
3	XG boost	70	12	15	87
4	SVM	69	13	4	98
5	KNN	73	9	7	95
6	Logistic Regression	67	15	8	94
7	Naïve Bayes	72	10	12	90
8	Hybrid Algorithm	74	8	7	95

Table 2: Confusion Matrix for Various Machine Learning Models.

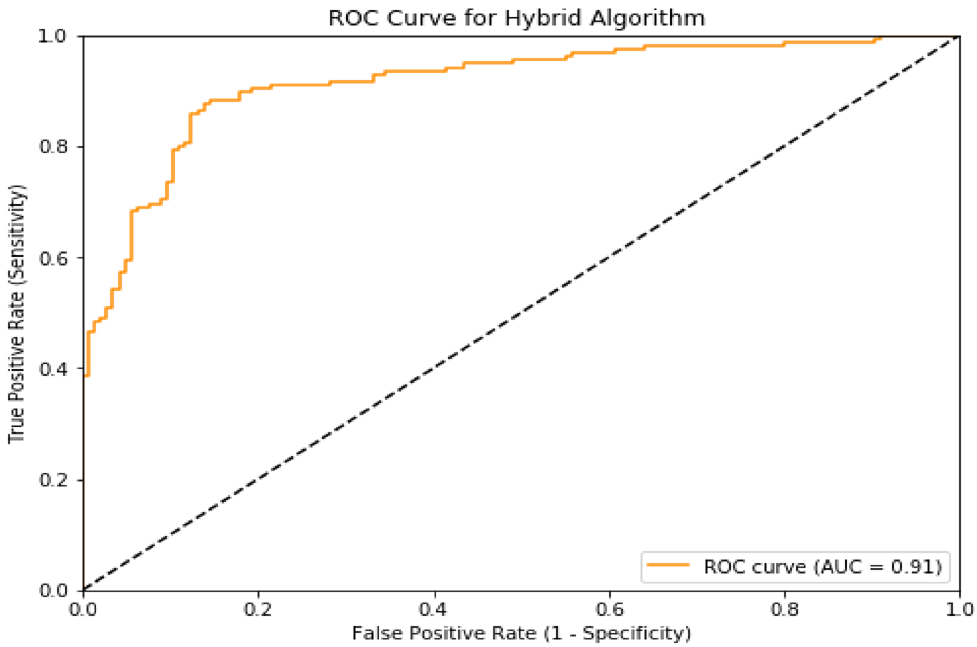

Figure 4: ROC Curve for the Hybrid Model

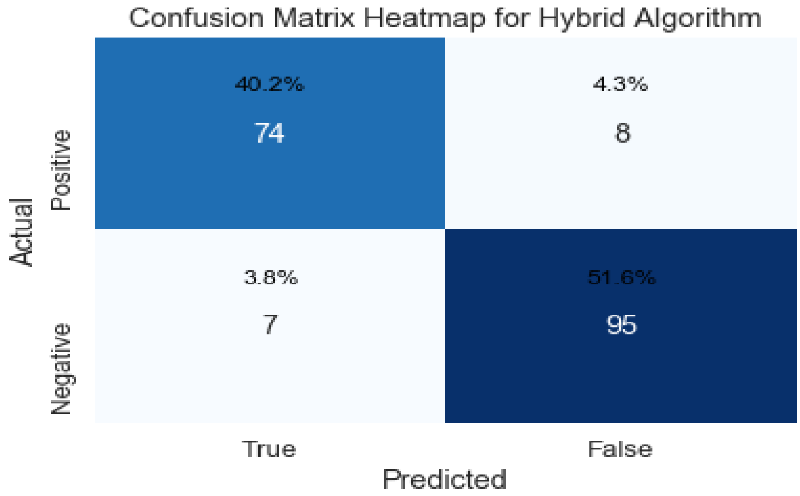

Figure 5: Confusion Matix Heatmap for Hybrid Algorithm

Python programming code for building a machine learning model for the prediction of cardiovascular diseases using the popular scikit-learn library:

```
# Importing the necessary libraries
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score

# Load the dataset
data = pd.read_csv('heart_data.csv')

# Split the dataset into features (X) and target variable (y)
X = data.drop('target', axis=1)
y = data['target']

# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42
)

# Feature scaling
scaler = StandardScaler()
X_train = scaler.fit_transform(X_train)
X_test = scaler.transform(X_test)

# Create a logistic regression model
model = LogisticRegression()

# Train the model
model.fit(X_train, y_train)

# Make predictions on the test set
y_pred = model.predict(X_test)

# Calculate the accuracy of the model
```

```
accuracy = accuracy_score(y_test, y_pred)
print(f"Accuracy: {accuracy}")
```

In the above code, the steps involved are:

1. Import the necessary libraries, including pandas for data manipulation, scikit-learn for machine learning algorithms, and metrics for evaluating the model.
2. Load the dataset using pd.read_csv() function, assuming the data is stored in a CSV file called 'heart_data.csv'.
3. Split the dataset into features (X) and the target variable (y).
4. Split the data into training and testing sets using train_test_split() function from scikit-learn.
5. Perform feature scaling using StandardScaler() to standardize the feature values.
6. Create a logistic regression model using LogisticRegression() from scikit-learn.
7. Train the model using the training data.
8. Make predictions on the test set using predict() function.
9. Calculate the accuracy of the model by comparing the predicted values with the actual values using accuracy_score() function.
10. Print the accuracy of the model.

Note: This code assumes that the dataset is already preprocessed and the features are appropriately encoded or scaled. Additionally, you may need to modify the code based on the structure and requirements of your specific dataset

5. Conclusion

Machine learning models have shown promising results in the prediction of cardiovascular diseases. Through the utilization of appropriate algorithms and techniques, these models can analyze patient data and provide accurate predictions, aiding in early detection and proactive intervention. The studies discussed in this article demonstrate the effectiveness of machine learning in identifying heart conditions and predicting the likelihood of cardiovascular diseases.

By leveraging large datasets and employing various machine learning algorithms such as logistic regression, decision trees, random forests, and support vector machines, researchers have achieved high accuracy rates in predicting heart diseases. These models utilize features such as age, gender, blood pressure, cholesterol levels, and other medical indicators to make informed predictions. Additionally, feature engineering techniques and data preprocessing play a vital role in enhancing the performance of the models.

Further Research:

While machine learning models have shown promise in the prediction of cardiovascular diseases, there are several areas for further research and improvement:

1. Incorporating more diverse and representative datasets: Increasing the diversity and inclusiveness of the datasets used for training and validation can help ensure that the developed models are applicable to a wider range of populations and demographics.
2. Exploration of advanced algorithms and ensemble techniques: Researchers can investigate the performance of advanced machine learning algorithms, such as deep learning models or ensemble techniques, to further improve the accuracy and robustness of cardiovascular disease prediction models.
3. Integration of multimodal data: Incorporating additional data sources, such as genetic information, wearable device data, or electronic health records, can provide a more comprehensive view of patient health and contribute to more accurate predictive models.
4. Explainability and interpretability: Developing machine learning models that provide explanations or interpretability for their predictions can enhance trust and acceptance among healthcare professionals, enabling them to make more informed decisions based on the model's output.

5. Real-time monitoring and risk assessment: Exploring the development of machine learning models that can continuously monitor patient health data and provide real-time risk assessment for cardiovascular diseases can enable timely interventions and personalized treatment plans.
6. Ethical considerations and bias mitigation: Further research is needed to address potential biases and ethical concerns associated with the use of machine learning models in healthcare, ensuring fairness, transparency, and accountability in the prediction of cardiovascular diseases.

In conclusion, machine learning models have demonstrated their potential in the prediction of cardiovascular diseases. With continued research and advancements, these models can contribute to improved patient care, better resource allocation, and proactive interventions, ultimately leading to better outcomes for individuals at risk of cardiovascular diseases.

References

nsari, M. F., AlankarKaur, B., & Kaur, H. (2021). A prediction of heart disease using machine learning algorithms. Journal of Advances in Intelligent Systems and Computing, 1200:497-504.

howmick, A., Mahato, K. D., Azad, C., & Kumar, U. (2022). Heart Disease Prediction Using Different Machine Learning Algorithms. Proceedings - 2022 IEEE World Conference on Applied Intelligence and Computing, AIC 2022, 60–65. https://doi.org/10.1109/AIC55036.2022.9848885.

hang, V., Bhavani, V. R., Xu, A. Q., & Hossain, M. (2022). An artificial intelligence model for heart disease detection using machine learning algorithms. *Healthcare Analytics*, *2*(September 2021), 100016. https://doi.org/10.1016/j.health.2022.100016.

ndal, H., Agrawal, S., Khera, R., Jain, R., & Nagrath, P. (2021). Heart disease prediction using machine learning algorithms. *IOP Conference Series: Materials Science and Engineering*, *1022*(1). https://doi.org/10.1088/1757-899X/1022/1/012072.

arthick, K., Aruna, S. K., Samikannu, R., Kuppusamy, R., Teekaraman, Y., & Thelkar, A. R. (2022). Implementation of a Heart Disease Risk Prediction Model Using Machine Learning. *Computational and Mathematical Methods in Medicine*, *2022*. https://doi.org/10.1155/2022/6517716.

avitha, M., Gnaneswar, G., Dinesh, R., Sai, Y. R., & Suraj, R. S. (2021). Heart Disease Prediction using Hybrid Machine Learning Model. *Proceedings of the 6th International Conference on Inventive Computation Technologies, ICICT 2021*, 1329–1333. https://doi.org/10.1109/ICICT50816.2021.9358597.

hanna, D., Sahu, R., Baths, V., & Deshpande, B. (2015). Comparative Study of Classification Techniques (SVM, Logistic Regression and Neural Networks) to Predict the Prevalence of Heart Disease. *International Journal of Machine Learning and Computing*, *5*(5), 414–419. https://doi.org/10.7763/ijmlc.2015.v5.544.

hourdifi, Y., & Bahaj, M. (2019). Heart disease prediction and classification using machine learning algorithms optimized by particle swarm optimization and ant colony optimization. *International Journal of Intelligent Engineering and Systems*, *12*(1), 242–252. https://doi.org/10.22266/ijies2019.0228.24.

utiame, S., Millham, R., Adekoya, A. F., Tettey, M., Weyori, B. A., & Appiahene, P. (2022). Application of Machine Learning Algorithms in Coronary Heart Disease: A Systematic Literature Review and Meta-Analysis. *International Journal of Advanced Computer Science and Applications*, *13*(6), 153–164. https://doi.org/10.14569/IJACSA.2022.0130620.

ithi, S. R., Aktar, A., Aleem, F., & Chakrabarty, A. (2018). *Machine Learning Approach for ECG Analysis and predicting different heart diseases*. *December*.