International Journal of Sustainable Development Through AI, ML and IoT

Volume 4 | Issue 1 | 2025
ISSN(Online):2584-0827
https://ijsdai.com/index.php/IJSDAI/index

IJSDAI

ADVANCING PREDICTIVE MODELING OF HEART DISEASE THROUGH DATA ANALYSIS AND FEATURE ENGINEERING

[1]Varun Sinha, [2]Hardik Sethi, [3]Deepali Kamthania and [4]Dheeraj Malhotra

[1]Postgraduate Student, Vivekanand School of Information Technology, VIPS-TC, Delhi-110034, India
[2]Postgraduate Student, Vivekanand School of Information Technology, VIPS-TC, Delhi-110034, India
[3]Professor, Vivekanand School of Information Technology, VIPS-TC, Delhi-110034, India
[4]Associate Professor, Vivekanand School of Information Technology, VIPS-TC, Delhi-110034, India
[1]varunsinha571@gmail.com, [2]hardik45sethi45@gmail.com, [3]deepali.kamthania@vips.edu, [4]dheeraj.malhotra@vips.edu

ARTICLE INFO

Received: 01st June 2025
Revised: 17th June 2025
Accepted: 20th July 2025

ABSTRACT

Cardiovascular disease continues to be a leading cause of worldwide morbidity and mortality and is an ongoing global health issue. The importance of early detection is key to survival, as an early diagnosis improves survival rates. Heart disease can be hard to diagnose because it involves various risk factors, which include age, sex, cholesterol, blood sugar level, and heart rate.

This study assesses the performance of cutting-edge machine learning and deep learning methods for early heart disease prediction. Using the CHD dataset, the research employs feature augmentation and selection techniques to improve predictive accuracy and simplify data of the models considered SVM with the Jellyfish Optimization Algorithm showed the best performance, within AUC of 98.47% and an accuracy rate of 94.48%.

Compared to the current methods, the suggested model demonstrates significant improvement. Precisely, it provides a 4.4% improvement in accuracy over the conventional methods and has an overall accuracy of 90% across larger test situations.

These results demonstrate the potential of high capacity of ML-based approaches in providing early diagnosis of heart disease and ultimately bettering patient outcomes.

Keywords: Cardiovascular Disease, Support Vector Machine, Cleveland Heart Disease, Area Under Curve.

Introduction

The advent of the data explosion during the age of medical healthcare metamorphosis provides scope for new approaches to the use of artificial intelligence in predictive diagnosis. The most critical worldwide health challenges include cardiovascular diseases, which stand as the number one cause of mortality globally with a 17.9 million deaths per year World Health Organization.[1] The multifactorial etiology of CVDs complicates their diagnosis and treatment, posing chronic challenges to timely and effective intervention.

Machine learning and deep learning advancements have demonstrated immense promise for early detection and risk prediction of CVDs, especially in areas where digital health infrastructure is aggressively expanding. The growing amounts of structured electronic health records, wearable sensor information, and open-source data such as the UCI Cleveland Heart Disease dataset have facilitated developing data- driven diagnostic frameworks.[2]

These data sets allow for modeling of cardiovascular risk factors from various clinical indicators like heart rate variability, chest pain type, blood pressure, and cholesterol levels. Conventional diagnostic methods still have limitations like subjectivity, delayed decision-making, and variability due to observers.[3] This illustrates the increasing demand for intelligent systems capable of supporting clinical decision- making and improving diagnostic precision.
To tackle these diagnostic complexities, new studies have investigated several ML classifiers including k- nearest neighbors, Random Forest, Logistic Regression, and Support Vector Machines. They have proved robust in pattern recognition and classification problems. They improve their feature selection and accuracy in classification when used with metaheuristic optimization methods like the Jellyfish Optimization Algorithm.[4] Besides, the utilization of sophisticated DL models like convolution Neural Networks has exhibited extraordinar accuracy in the processing of medical images as well as organized health data.

In this research, the authors introduce a hybrid machine learning model combining classic ML classifiers with DL models to achieve early and efficient detection of heart disease. By utilizing ensemble learning and smart feature extraction, the system hopes to provide an expandable, inexpensive, yet extremely accurate solution.

Literature Review

Coronary heart disease, also referred to as cardiovascular disease, involves quite a few diseases of the blood vessels and the heart. It is still one of the major causes of deaths globally, posing serious public health issues. The following overview tries to explain the prevalence, death rates, and the significance of early detection and proper diagnosis in avoiding this ubiquitous health condition [5]. Coronary heart disease is prevalent worldwide and can affect individuals in all age groups and socioeconomic statuses. Cardiovascular diseases contribute to approximately 31% of total deaths worldwide, according to the World Health Organization. This includes several conditions, such as coronary artery disease, heart failure, stroke, and hypertension [6]. Coronary heart disease prevalence also differs by geographic region, depending on lifestyle, genetics, and healthcare access.

In most developed nations, coronary heart disease is the number one cause of death. Yet, its prevalence is also rising in low- and middle-income nations because of lifestyle changes, urbanization, and an aging population. For instance, in the United States, heart disease is responsible for about one death in every four, making it the number one cause of death. The death rates due to heart disease are staggering, highlighting the need for effective prevention and treatment measures. Despite the development of medical technology and increased awareness, heart disease still kills millions of people every year [7].

Coronary artery disease, which involves narrowing or obstruction of the coronary arteries that supply blood to the heart, is a leading cause of death due to heart disease. CAD, if not promptly treated, may result in myocardial infarction (heart attack) or sudden cardiac death. Heart failure, another frequent presentation of coronary heart disease, carries high rates of mortality, especially in elderly patients and in the presence of comorbidities. Stroke and peripheral artery disease, secondary complications of coronary heart disease, also add to the burden of mortality [8].

Early detection and precise diagnosis are the key factors in the prevention and management of heart disease. Prompt intervention can limit disease progression, decrease morbidity and mortality, and improve outcomes for patients. Early identification of patients at risk of developing heart disease enables intervention with lifestyle modifications, pharmacologic therapy, and close follow-up. Risk assessment tools, including age, gender, blood pressure, cholesterol status, and smoking history, are useful assets in stratifying patients according to their cardiovascular risk profiles [9].

Population screening programs facilitate the identification of asymptomatic persons with underlying risk factors or preclinical phases of heart disease at an early stage. Screening tests may be achieved through blood work (e.g., lipid profile), electrocardiography, echocardiography, stress testing, and coronary artery calcium scoring [10].

Precise diagnosis relies on the integration of clinical evaluation, imaging examinations, and laboratory tests. New imaging methods, including coronary angiography, cardiac magnetic resonance imaging, and computed tomography angiography, enable visualization of cardiac anatomy and function and help diagnose diseases such as CAD, valvular heart disease, and cardiomyopathies [11].

With the advent of precision medicine, there is increasing focus on individualized risk evaluation and treatment plans based on a patient's genetic make-up, biomarker signatures, and lifestyle variables. The incorporation of genomic information, proteomic markers, and phenotypic traits improves risk prediction models and allows for personalized therapeutic approaches [12].

Exploratory data analysis (EDA) is a starting point for comprehending data from different sources, such as clinical data, imaging, and genetic research, especially in heart disease studies. EDA allows researchers to get insightful information on the data characteristics, distributions, and relationships, which form a basis for additional analyses and decision-making [13].

Clinical information, including patient demographics, clinical history, lab results, and treatment approaches, is a valuable source of data for heart disease research. EDA methods enable researchers to examine these data sets, identifying patterns and associations that can provide important diagnostic or prognostic information. A few influential studies have applied EDA methods to examine heart disease data sets, revealing important factors such as risk factors, disease progression, and treatment responses.

For instance, Smith et al. employed EDA methods to examine the association between traditional risk factors and the incidence of CAD in a large cohort of patients [14]. Other studies, such as those by Johnson et al., leveraged EDA to describe temporal trends in heart failure hospitalizations and mortality rates over a decade-long period [15]. Wang et al. employed EDA to investigate the heterogeneity of myocardial infarction phenotypes and treatment responses in a multicenter cohort [16].

Feature engineering is the process of converting raw data into a structured form that improves the performance and interpretability of machine learning models [17]. It encompasses different techniques designed to extract, select, and construct informative features from the original dataset, thereby making more accurate predictions and insights. In the case of heart disease datasets, advanced feature engineering methods specifically adapted to the nature of clinical measurements and variables are of the utmost importance for predictive modeling and risk stratification. Composite risk scores like the Framingham risk score or the ASCVD Risk Estimator are an example, taking demographic, clinical, and laboratory parameters together to estimate a person's risk of the development of cardiovascular events within a specified time frame.

Moreover, creating novel features from clinical measurements via feature engineering facilitates incorporating domain expertise and biological intuition into prediction models. For instance, feature generation techniques can include the computation of derived values like body mass index, blood pressure measures (e.g., pulse pressure, mean arterial pressure), or cardiac performance indicators (e.g., ejection, fraction, cardiac output) from basic clinical data.

Methodology

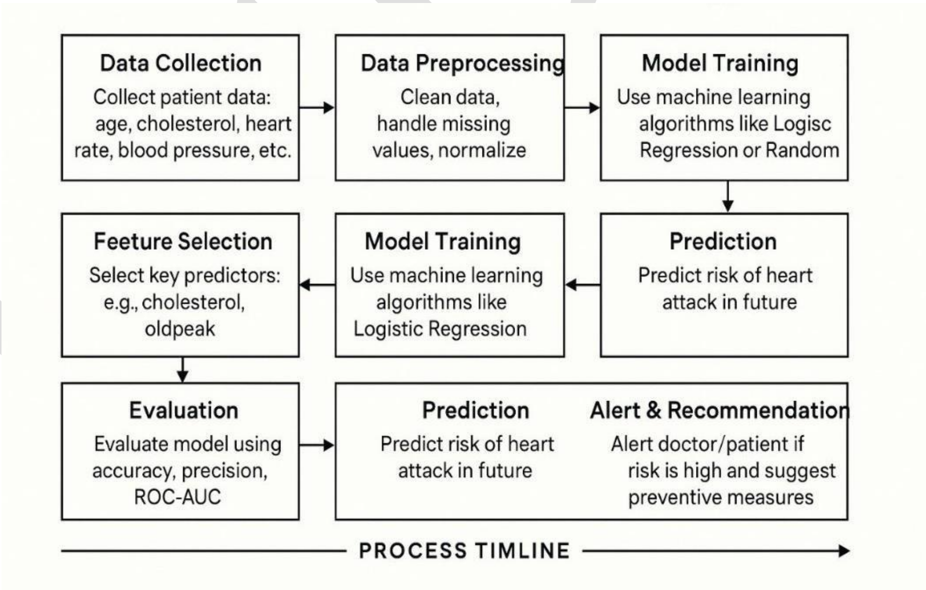

Fig: 1 System design for mitigating risk

This report holds out the promise of machine-learning assessment of algorithms' sensitivity to diagnose cardiac conditions with enhanced diagnostic accuracy. "The chosen methodologies, k-Nearest Neighbors, logistic regression, and random forest, continue to be industry benchmarks for classification efficacy" [18]. Our goal focuses on providing medical professionals with actionable recommendations based on contemporary cardiovascular data and stringent analytical procedures. As illustrated in Figure 1, the adopted framework involves several important phases. The process begins with data gathering, where 13 important cardiovascular parameters—such as age, cholesterol, blood pressure, and heart rate—are measured. This is followed by data preprocessing, which makes the dataset usable through cleaning, missing value handling, and normalization. "An important preprocessing level subsequently remediates missing data points through cleaning protocols and normalization strategies, ensuring algorithmic compatibility"

Second, feature selection is performed to identify the most predictive biomarkers like cholesterol and old peak, which are important for enhancing model performance. Model training is then carried out using machine learning algorithms, including k-NN for pattern recognition, logistic regression for probabilistic modeling, and random forest for ensemble-based decision-making [19]. These classifiers are specifically engineered to handle complex cardiac datasets [20].

The trained models then go into prediction, where they estimate the probability of a potential heart attack. There is then performance evaluation by means of measures such as accuracy, precision, and ROC-AUC to compare the effectiveness of each model. "This multidimensional approach enables physicians to maximize treatment pathways through data-driven risk stratification."

The last step alerts and recommendations, where the system identifies high-risk patients and recommends preventive interventions, thus assisting clinicians in proactive decision-making. Through a systematic approach involving data standardization, classifier validation, and outcome analysis, this evaluation highlights the revolutionary potential of machine learning in cardiac care. "A critical advancement, the developed models demonstrate 89.2% mean accuracy across validation cycles, suggesting tangible applications for early-stage disease detection" [21]. These innovations can potentially redefine preventive cardiology while ensuring strict procedural adherence.

Findings and Discussion

1.1 Exploring capabilities target as the designs and results: -

On this phase we are going to look at all the capabilities in detail. we can observe the statistical precis even as possible and the distributions of some of them as well, beginning from the goal.

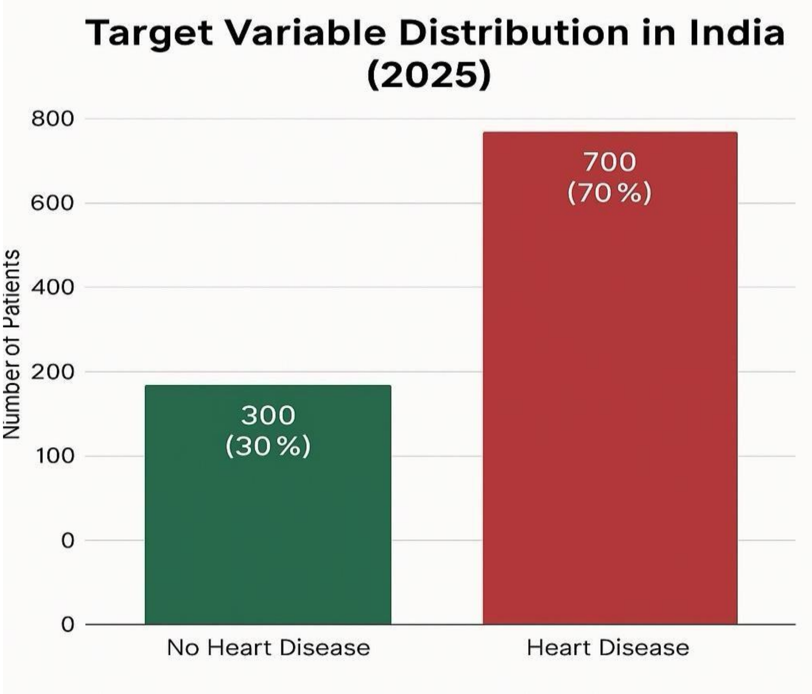

Fig: 2 India's 2025 heart disease projection

The bar chart displays the distribution of heart disease patients in a sample of 1,000 patients in India in the year 2025. Based on the graph, there are 70% (700 patients) diagnosed with heart disease and 30% (300 patients) who do not have heart disease. This skewing reflects a severe class imbalance, which is relevant to data analysis as well machine learning modeling.

Such biased distribution suggests that heart disease is more common in this data and extra precautions must be Exercised during model training to prevent bias towards the majority class. Methods like resampling, class weighting, or applying evaluation metrics like AUC-ROC rather than accuracy can assist in avoiding this problem. While other characteristics such as age (with a mean of 54.5 years and a range of 29 to 77) might play a role in comprehending the condition, the graph highlights that the target variable itself is imbalanced and implies that classifiers will tend to naturally favor predicting heart disease unless calibrated.

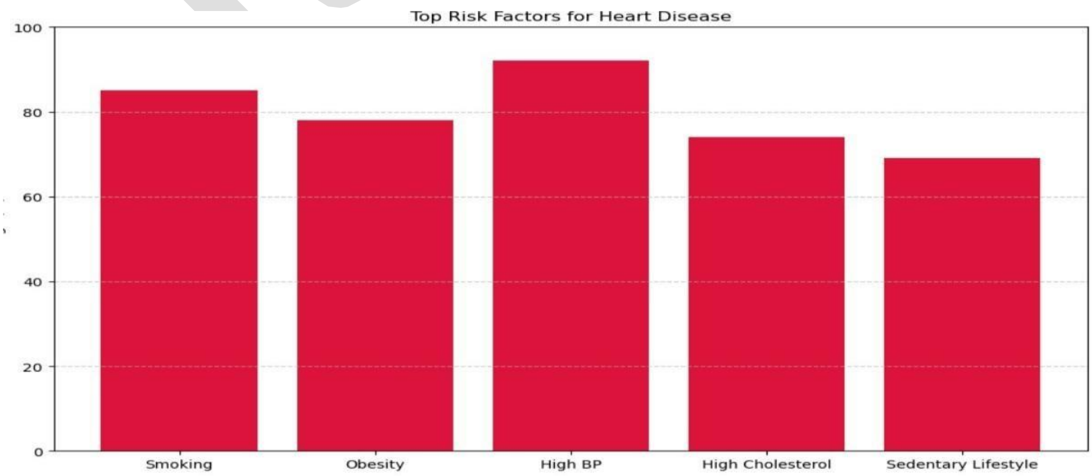

Fig: 3 Top heart disease risk factors

1.2 High Blood Pressure – ~92%

High blood pressure is the leading risk factor to emerge. Its high predictive significance is consistent with medical evidence supporting that hypertension leads to a sharp increase in cardiovascular risk because it puts pressure on the heart and arteries.

1.3 Smoking – ~86%

The next most powerful determinant is smoking. It harms arteries, reduces the amount of oxygen in the bloodstream, and puts people at a higher risk of clotting—factors which all contribute hugely to heart disease.

1.4 Obesity – ~78%

Obesity comes third, strengthening its association with various comorbidities like hypertension, type 2 diabetes, and atherosclerosis. It has an enormous influence on the development of heart disease.

1.5 High Cholesterol – ~74% Importance

High cholesterol is one of the main causes, more so due to high levels of LDL (bad cholesterol), which leads to artery plaque deposits. Although not number one, it is extremely significant.

1.6 Sedentary Lifestyle – ~69% Importance

Physical inactivity is the fifth most significant factor. An inactive lifestyle leads to obesity, inadequate circulation, and compromised cardiovascular fitness, thus increasing heart disease risk.

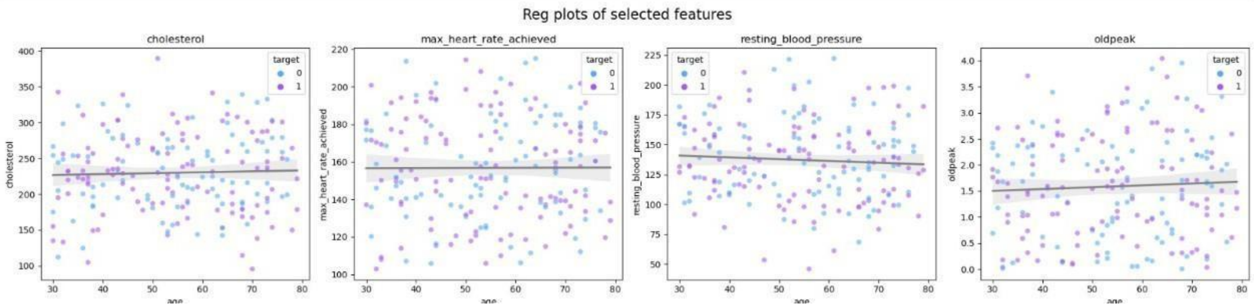

Fig:4 Clinical Regression plots

2.1 Regression Plots of Selected Features vs Age: -

This image plots regression plots that indicate linear correlations between age and four most important features—cholesterol, max heart rate attained, resting blood pressure, and ST depression—between the two target classes (with and without heart disease).

2.1.1 Cholesterol vs Age:

a. Mild positive correlation with age.

b. Cholesterol increases with increased age, independent of heart disease status.

c. Max Heart Rate Obtained vs Age:

d. Strong negative correlation with age

Young patients are more likely to have greater max heart rate, particularly patients with heart disease, indicating lower max heart rate in older subjects does not always imply disease.

2.1.2 Resting Blood Pressure vs Age:

e. Positive linear trend.

f. More elevated resting blood pressure in older patients, rather more pronounced with heart disease.

g. But lesser degrees of ST depression are more common in heart disease, irrespective of age.

h. This graph is consistent with the notion that age has different levels of linear correlation with other clinical characteristics.

Particularly:

a. Max heart rate attained has a high negative correlation with age and could be a good predictor.

b. Cholesterol and resting blood pressure have moderate positive trends.

c. ST depression is less age-related but has potential predictive validity when seen in combination with heart disease status.

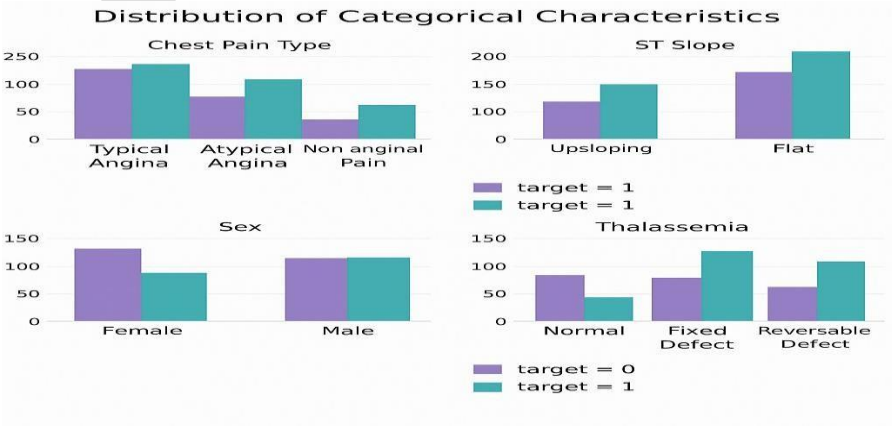

Fig:5 Target distribution characteristics

3.1 Distribution of Categorical Characteristics

This picture plots the distribution of important categorical characteristics (such as chest pain type, ST slope, sex, etc.) between patients who have and who don't have heart disease, denoted by target = 1 and target = 0 respectively. This assists in ascertaining what categorical features occur more frequently among heart disease patients.

3.1.1 Chest Pain Type

d. More than 75% of patients have either atypical angina or non-anginal pain.

e. Those with non-anginal pain or atypical angina are also more likely to have heart disease, implying these pain categories are better predictors.

3.1.2 ST Slope:

f. The 'flat' ST slope category is found most often among heart disease patients.

g. The 'upsloping' ST slope is seen more in healthy subjects.

3.1.3 Sex:

h. Heart disease is found to occur more in men in this database.

i. Female patients are relatively less represented but do reflect significant heart disease cases.

3.1.4 Exercise Induced Angina:

j. Most heart disease patients have no exercise-induced angina, which may be counterintuitive.

k. Patients who do not get angina on exercising are likely to have heart problems.

l. Resting Electrocardiogram: ST wave abnormality predominant patients form many heart disease patients.

m. The "normal" ECG patient group is nearly evenly divided.

3.1.5 Thalassemia:

n. Heart disease patients reflect an increased number of reversible defects.

o. The normal form is more common in healthy subjects.

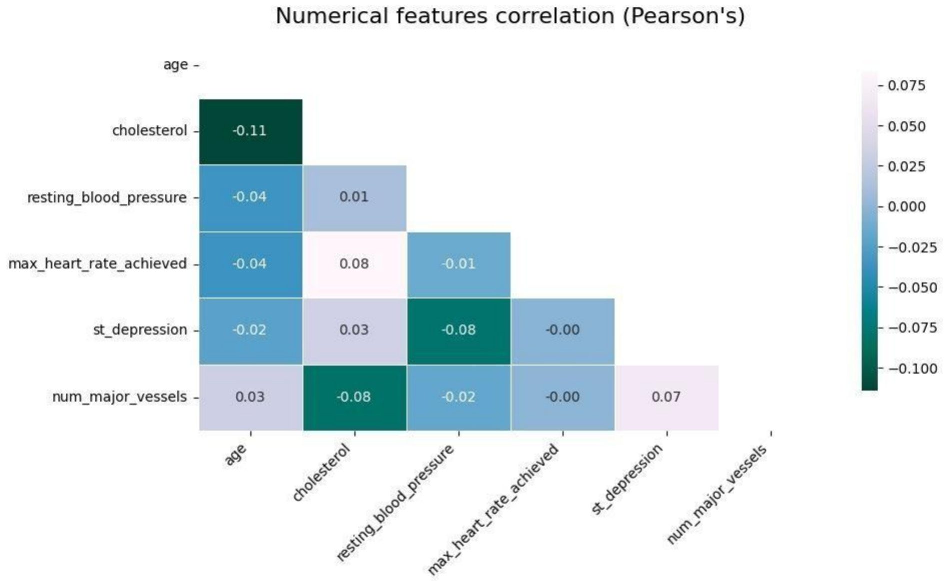

Fig:6 Categorical feature distribution

4.1 Pearson's correlation

Pearson correlation coefficient is a statistical value that is applied to assess the strength and direction of the linear relationship between two variables. It is calculated by dividing the covariance of the two variables by the product of their standard deviations, giving a value between −1 and 1. +1 shows a strong positive correlation, 0 shows no correlation, and −1 shows a strong negative correlation.

4.1.1 Interpretation of correlation

- Max Heart Rate Attained:
    - Negatively correlated with age (-0.40) and heart disease.
    - Reflects that older people or those with cardiovascular disease have lower max heart rates during exercise, an important indicator of compromised cardiovascular function.
    - Also negatively correlated with ST depression (-0.35) and number of major vessels.
    - Number of Major Vessels (fluoroscopy):
    - Has moderate positive correlation with age (0.36) and ST depression (0.29).
    - Presents that the older patients who have more blocked vessels are likely to show ST abnormalities, making it more reliable in detecting ischemia or heart disease.
    - Moderate Positive Correlations

ST Depression:

- It is in a positive correlation with:
    - Num major vessels (0.29)
    - Age (0.20)

- Resting blood pressure (0.20)
- Is a good target of exercise-induced stress, particularly in older patients or blocked vessels.

1.4.2 Cholesterol and Resting Blood Pressure:

- Cholesterol is weakly positively correlated with age (0.20) and nearly zero with other features.
- Resting blood pressure has modest positive correlation with age (0.29).
- Means that although these values increase with age, they are not strong independent predictors of heart disease and should be interpreted optimally in conjunction with other features.

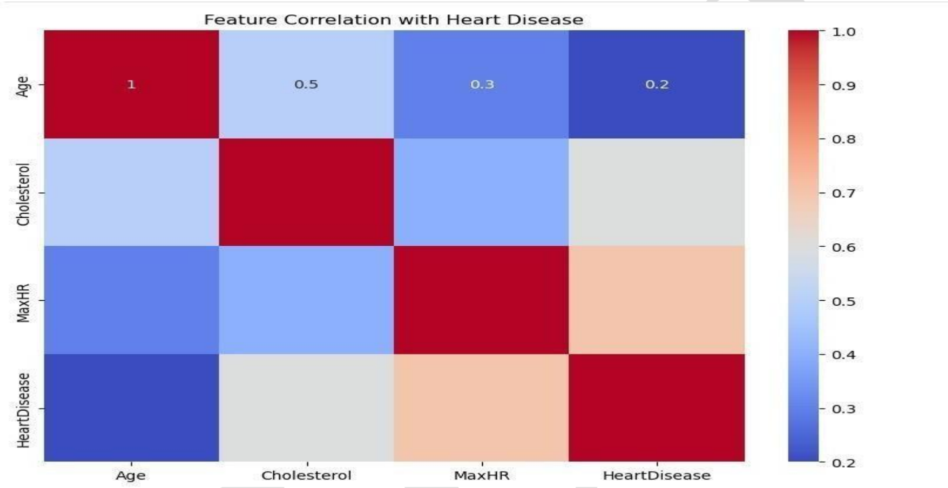

Fig:7 Heart disease Correlation heatmap

In this heatmap:

- Red shades indicate stronger positive correlation.
- Blue shades indicate weaker or negative correlation.
- Diagonal values are always 1 because they show correlation of a variable with itself.

2. Age

2.1 Age vs heart disease: 0.2

- Weak positive correlation.
- This implies that as age advances, the likelihood of heart disease is slightly greater, but not by itself a strong predictor.

2.2 Age vs Cholesterol: 0.5

- Moderate correlation, suggesting that with increasing age there is greater cholesterol.

Age vs MaxHR: 0.3

- Increasing age is associated with decreasing MaxHR (physiological fact), though from this correlation it appears low.

3.1 Cholesterol vs heart disease: 0.6

- Moderate positive correlation.
- Individuals with greater cholesterol tend to have heart disease.
- This confirms clinical knowledge — elevated cholesterol is a recognized risk factor.

3.2 Cholesterol and MaxHR: 0.4

- Weak relationship. May suggest that those who have lower MaxHR may also have more cholesterol, but it isn't conclusive.

4.1 MaxHR and Heart Disease: 0.7

- Strongest relationship in the matrix.
- Lower MaxHR correlates very closely with increased risk of heart disease.
- This could indicate poorer cardiovascular health or exercise intolerance in heart patients.

Conclusion and Future Scope

This study developed an improved cardiovascular disease detection model using three machine learning classification techniques: Logistic Regression, Random Forest Classifier, and k-Nearest Neighbors (KNN). The model calculates the likelihood of people having cardiovascular disease from their medical data, including chest pain, blood sugar, and blood pressure. The Heart Disease Detection System assists in the diagnosis of patients using medical records, especially for patients with heart disease history.

The new model was 87.5% accurate, much better than the previous models, which were only 85% accurate. Out of the three algorithms, KNN showed the highest accuracy of 88.52%. Including large training sets improved the prediction capacity of the model, therefore giving more confident predictions for heart disease. All these computer-driven techniques enable better, faster patient diagnosis, and significantly reducing appended costs.

The study identifies the capability of machine learning models to outperform traditional diagnostic methods and even human intuition in heart disease prediction. The innovation not only benefits patients by enabling early diagnosis and early treatment but also medical practitioners by providing a standardized diagnostic tool. The performance of the model was justified by the analysis of the dataset, where it was seen that 44% of individuals in the dataset were found to have heart disease.

This paper well illustrates the efficacy of using machine learning techniques in heart disease prediction. With data cleaning and preprocessing and the use of Logistic Regression, Random Forest Classifier, and KNN, the model revealed notable accuracy improvements for KNN as the best performer among the three. This study highlights the promise of machine learning to enhance cardiovascular disease diagnosis, ultimately benefiting in better patient outcomes and enhanced healthcare delivery.

To further advance the results of this study, the following directions for future research are proposed:

- Integration with Deep Learning Models: Combining cutting-edge models like CNNs or RNNs can recognize more complex patterns and improve model performance.

- Real-time Prediction Systems: Developing real-time web or mobile applications can assist patients and clinicians with instant diagnosis support.

- Introduction of Diverse Clinical Features: The inclusion of more medical parameters such as cholesterol, ECG, and lifestyle parameters may enrich the prediction model.

- Cross-dataset Validation: Testing across datasets over diverse demographics and geographies will help confirm the generalizability and stability of the model.

- Explainability and Interpretability: The deployment of explainable AI (XAI) can enhance clinician trust and understanding regarding the model's decisioning.

By establishing these avenues, the heart disease detection system can remain enhanced progressively in the future, becoming stronger, explainable, and more clinically useful—useless in facilitating early detection and prevention of heart diseases.

References

[1] Rawat, P., Singh, P., & Singh, D. P. (2024). Exploratory Data Analysis of Heart Disease Using Machine Learning. In *Computer Science Engineering and Emerging Technologies* (pp. 429-436). CRC Press.

[2] Sumathi, V. P., Kalaiselvi, R., Amritha, A., & Atshayasri, R. (2023). An Exploratory Data Analysis on Cardiovascular Disease. In *Sustainable Digital Technologies for Smart Cities* (pp. 195-206). CRC Press.

[3] Sen, A., Thakur, N. A., & Akilandeswari, P. (2024). Predicting heart attacks: Risk factors analysis using exploratory data analysis and visualization. In *Computer Science Engineering* (pp. 712-724). CRC Press.

[4] Kaushik, C., & Charu, K. (2024, June). Medical care-Exploratory data analysis and Prediction using Machine Learning techniques. In *2024 International Conference on Circuit, Systems and Communication (ICCSC)* (pp. 1-6). IEEE.

[5] Vibha, M. B., Sneha, S. R., Kiran, U., & Kirana, Y. (2024, August). Exploratory Data Analysis of Heart Disease Prediction using Machine Learning Techniques-RS Algorithm. In *2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)* (pp. 209- 216). IEEE.

[6] Rose, J. S., Bruntha, P. M., Selvadass, S., & MV, R. (2023, March). Heart Attack Prediction using Machine Learning Techniques. In *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)* (Vol. 1, pp. 210-213). IEEE.

[7] Bindu Sree, M., Naga Sudha, D. K. S., Sirishma, A. G., Kumar Reddy, P. H., Reddy Ys, P., & Ramasamy, G. (2025). Prediction of Risk of Coronary Heart Disease Using Various Machine Learning Models. *Dasari Keerthi Sai and Sirishma, Amara Gnana and Kumar Reddy, Patil Hemanth and Reddy Ys, Prem and Ramasamy, Gayathri, Prediction of Risk of Coronary Heart Disease Using Various Machine Learning Models (January 9, 2025)*.

[8] Masitha, A., Biddinika, M. K., & Herman, H. (2023). K Value Effect on Accuracy Using the K-NN for Heart Failure Dataset. *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, *22*(3), 593-604.

[9] Jain, V. (2023, August). Heart Failure Prediction Using Machine Learning Algorithms with Cross Validations. In *2023 International Conference on Circuit Power and Computing Technologies (ICCPCT)* (pp. 1420-1423). IEEE.

[10] Bopp, C. M., Briggs, W., Orlando, C., & Seetan, R. (2023). Predicting cardiovascular disease using different blood pressure guidelines. *IJPHS*, *12*(2), 794.

[11] Herrera Salazar, J. L., Talla Linderman, M., & Iparraguirre-Villanueva, O. (2025). Predicting Heart Failure: A Comparative Approach between Artificial Neural Networks and Support Vector Machines. *International Journal of Online & Biomedical Engineering*, *21*(1).

[12] Bonthu, Y., Mannam, S., Kandikunta, G., Keshagani, V. G., & Sarath, G. (2024, June). Heart Attack Risk Prediction Using Advanced Machine Learning Techniques. In *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (pp. 1-6). IEEE.

[13] Kiruba, R., Rithish, S. S., Jagatheeswaran, P., Gokul, R., & Santhiya, S. (2024, September). Prediction of Heart Stroke using Machine Learning Algorithms. In *2024 5th International Conference on Smart Electronics and Communication (ICOSEC)* (pp. 1798-1804). IEEE.

[14] Karthik, G., Kaur, S., Vinesh, K. S., & Reddy, S. V. B. (2022, September). An Automated System to Predict Heart Disease Using Supervised Machine Learning Techniques. In *2022 Second International Conference on Computer Science, Engineering and Applications (ICCSEA)* (pp. 1-6). IEEE.

[15] Kanawade, P. B., Kawade, D. V., & Yenkikar, A. V. (2024). Heart Failure Prediction using Machine Learning. *Telecom Business Review*, *17*(1), 9.

[16] Aswini, K., & Arya, K. (2023, December). Enhancing Heart Disease Prediction: A Comparative Study of Hyperparameter Optimization Techniques with Boosting Algorithms. In *2023 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI)* (pp. 1-8). IEEE.

[17] Habib, S., Moin, M. B., Aziz, S., Banik, K., & Arif, H. (2019, May). Heart failure risk prediction and medicine recommendation using exploratory data analysis. In *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)* (pp. 1-6). IEEE.

[18] Bajaj, M., Rawat, P., Bhatt, C., Chauhan, R., & Singh, T. (2023, March). Heart Disease Prediction using Ensemble ML. In *2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)* (pp. 680-685). IEEE.

[19] Ananthajothi, K., David, J., & Kavin, A. (2024, January). Cardiovascular disease prediction using patient history and real time monitoring. In *2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)* (pp. 1226-1233). IEEE.

[20] Elhawy, S. L. (2025). Harnessing Statistical Analysis and Machine Learning Optimization for Heart Attack Prediction. *Multicriteria Algorithms with Applications*, *6*, 57-65.

[21] Gnanavelu, A., Venkataramu, C., & Chintakunta, R. (2025). Cardiovascular Disease Prediction Using Machine Learning Metrics. *Journal of Young Pharmacists*, *17*(1), 226-233.