ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.02 No.04 (2024)

PREDICTIVE MODELING OF CARDIOVASCULAR DISEASE USING MACHINE LEARNING

Muhammad Aqeel Badar
Faculty of Computer Science and Information Technology, Superior University, Pakistan
Fakhar Ur Rehman
Faculty of Computer Science and Information Technology, Superior University, Pakistan
Javeed Ali
Faculty of Computer Science and Information Technology, Superior University, Pakistan
Fawad Nasim
Faculty of Computer Science and Information Technology, Superior University, Pakistan
Hijab Sehar
Riphah School of Computing and Innovation, Lahore
Corresponding Author: Muhammad Aqeel Badar . Email: aqeelbadar41@gmail.com,

Abstract:
The rising incidence of cardiac conditions is becoming a significant concern; therefore, it is crucial to anticipate these cases in advance. Consequently, arriving at this diagnosis presents a challenging endeavor that must be executed swiftly and with precision. The primary objective of this study is to assess an individual's likelihood of developing a cardiac condition by considering various medical factors. A tool was developed to predict the likelihood of a heart disease diagnosis, utilizing the patient's medical history as a basis for its predictions. We utilized several machine learning methods, including logistic regression and KNN, to forecast and categorize individuals diagnosed with heart disease. A systematic approach was utilized to regulate the model's implementation, aiming to enhance the precision of predicting heart attacks in individuals. The model demonstrated a significant level of precision in forecasting indicators of heart disease in a person through the application of KNN and Logistic Regression techniques. The accuracy achieved surpassed that of previous classifiers, such as naive Bayes. The implementation of the proposed model has notably alleviated stress by diminishing the chances of the classifier accurately and precisely detecting cardiac conditions. The proposed method for predicting cardiovascular disease reduces expenses and improves healthcare quality. This project collects substantial data that may aid in predicting individuals at risk of developing heart disease. Pynb files serve as the medium for its implementation.

Keywords: Heart diseases, logistic regression, random forest classifier, K nearest neighbors (KNN), Cardiovascular health.

1. Introduction

This study is based on past work in the area of cardiovascular disease prediction using machine learning algorithms. Relevant research works include the study of cybersecurity with blockchain technology [1] and the analysis of cybersecurity measures for social systems [2]. Nasim and colleagues have greatly contributed to this field in the following research areas: data-driven predictions in cricket [3], intelligent fault detection systems for vehicles [4], and image fusion techniques [5]. Such studies provide a solid base for the development and improvement of

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.02 No.04 (2024)

machine learning models that predict cardiovascular disease risk. Machine learning is characterized as the process of "manipulating and extracting implicit, previously unknown/known, and potentially useful information about data". The domain of machine learning encompasses a wide array of areas and continues to grow in its applications. To assess the precision of a specific dataset and generate forecasts, machine learning utilizes various classifiers such as ensemble learning, supervised learning, and unsupervised learning. The information will be invaluable to numerous individuals; therefore, we can apply it to our HDPS project. Currently, there is a significant occurrence of cardiovascular issues, which include various conditions that can impact the heart. According to projections by the World Health Organization, approximately 17.9 million people around the world die from cardiovascular diseases (CVDs).

Among adults, it stands as the leading cause of mortality. Our examination could assist in identifying individuals who are at a higher risk of being diagnosed with heart disease, taking into account their medical history [6]. It can identify individuals showing signs of heart disease, such as elevated blood pressure or chest discomfort, and assist in diagnosing the condition with fewer assessments and more efficient therapies, facilitating the appropriate treatment pathway. This study centres on three specific data mining methodologies: (1) Random Forest Classifier, (2) KNN, and (3) Logistic Regression. Our project attained an accuracy of 87.5%, surpassing the previous system that utilized only a single data mining technique. Consequently, the implementation of advanced data mining techniques led to an enhancement in the accuracy and efficiency of the HDPS. Logistic regression is applied within the framework of supervised learning. Logistic regression employs solely distinct values.

Based on several medical factors, including age, gender, chest pain, fasting blood sugar level, and more, this study aims to forecast a patient's chance of receiving a diagnosis of cardiovascular heart issues. A dataset from the UCI repository was used, which included patient attributes and medical history. Through the analysis of this data, we can determine the patient's risk level for developing cardiac disease. To predict the likelihood of a patient experiencing a heart condition, we categorize them based on 14 medical criteria. The medical attributes are analyzed through three distinct methodologies: Logistic regression, KNN, and random forest classifiers. KNN outperforms these algorithms, achieving an accuracy of 88.52%. Finally, we classify them based on potential risk factors for heart disease. This strategy is also very cost-effective.

Related Work

A significant amount of investigation into the application of machine learning algorithms for identifying cardiovascular heart disease has motivated this study. This document presents a concise overview of the existing literature. Various algorithms, including the Random Forest Classifier, KNN, and Logistic Regression, have been utilized to accurately predict cardiovascular disease. The results demonstrate that the ability of each algorithm to capture the specified objectives differs [7]. By employing both traditional and contemporary machine learning and deep learning models, the model incorporating IHDPS successfully determined the decision boundary. This provided access to crucial and fundamental information, such as a family history of heart disease. However, the IHDPS model's precision was significantly lower than that

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.02 No.04 (2024)

of the more recent models that use artificial neural networks and other machine and deep learning approaches to predict heart disease.

McPherson et al. [8] employed an internal implementation strategy utilizing specific neural network techniques to pinpoint the risk factors associated with atherosclerosis or coronary heart disease. The only aspect they could reliably predict was the presence or absence of the specific condition in the test participant.

R. Subramanian et al. [24] presented neural networks for diagnosing and predicting heart disease, blood pressure, and other features. When applying the model to the Test Dataset, a deep neural network was created using the designated disease-related variables to guarantee accurate results about the existence of heart disease. Utilizing the output perceptron, this network successfully produced an output that nearly incorporated 120 hidden layers.

The utilization of the supervised network for diagnosing heart illness has been suggested [16]. During the testing conducted by a physician, an unidentified dataset was employed, and the model, which had been trained on previously acquired data, was utilized to forecast the outcome, assessing the precision of the given model.

Information Source

A group of individuals was chosen based on a structured dataset that considered various medical conditions and their previous history of cardiac issues [2]. Cardiovascular conditions encompass a variety of disorders associated with the heart. Cardiovascular diseases are the biggest cause of death among middle-aged persons, according to the World Health Organization (WHO). A comprehensive data set comprising 304 distinct patients' medical records spanning multiple age groups is utilized. To diagnose cardiac disease, this dataset contains critical patient medical information like age, resting blood pressure, and fasting blood sugar level. This dataset comprises thirteen medical attributes from 304 individuals, aiding in the differentiation between patients at risk and those not at risk of developing heart disease. The dataset on heart disease is from the UCI repository. This information enables the identification of the patterns that influence the likelihood of individuals developing heart disease. The records are categorized into two distinct sections: Testing and Training. The dataset comprises 303 entries organized into 14 columns, with each entry corresponding to a distinct row. Each feature has a detailed description in "Table 1."

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.02 No.04 (2024)

1. TABLE

S.No	Observation	Description	Values
1	Age	Age in Years	Continuous
2	Sex	Sex of the Subject	Female/Male/ Transjender
3	CP	Chest Pain	Are Four Types
4	Trestbps	Blood Pressure At Rest	Continuous
5	Chol	Serum Cholesterol	Continuous
6	FBS	Fasting blood sugar levels	<, or > 120 mg/dl
7	Restecg	Electrocardiogram at Rest	Are Five Values
8	Thalach	The highest heart rate attained	Continuous
9	Exang	Angina Induced by Exercise	Yes/No
10	Oldpeak	Exercise-induced ST Depression in relation to rest duration	Continuous
11	Slope	Slope of the Peak Exercise ST section	up/Down/Flat
12	Ca	Provides the number of major vessels that have been fluoroscope-colored.	0-3
13	Thal	Types of defects	Reversible/Fixed/Normal
14	Num(Disorder)	Heart Disease	Among the four main categories, present or not present.

Table 1. Various Attribute Used are Listed

Method

This study explores numerous machine-learning methods, like K nearest neighbours (KNN), logistic regression, and random forest classifiers. These algorithms have the potential to assist healthcare professionals. The medical analysts accurately identify cardiac disease. This process involves a thorough examination of journals, published studies, and the most recent statistics regarding cardiovascular disease. The proposed model is underpinned by a robust framework derived from the methodology [13]. The methodology consists of a series of procedures that convert provided data into discernible patterns that users can understand. The proposed approach (Figure 1) is divided into several phases: the initial phase involves gathering data; the subsequent phase focuses on extracting key values; and the third phase entails analyzing the data. Data preparation addresses absent values, the dataset, and standardizes the information, contingent upon the employed techniques [15]. Following data preprocessing, the data is subsequently classified utilizing a classifier. The proposed model utilizes three classifiers: KNN, Random Forest Classifier, and Logistic Regression Classifier. Ultimately, we implemented the

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.02 No.04 (2024)

suggested model and assessed its precision and effectiveness through various performance metrics. A thorough Heart Disease Prediction System (EHDPS) created by combining multiple classifiers is presented within this model. 13 medical characteristics are used in this model. To make predictions, including age, sex, fasting sugar, cholesterol, blood pressure and chest pain [17].

2. Comparison Table

Year	Title/Study	Authors	Algorithms/ Methods	Data Sources	Findings	Challenges
2019	Machine Learning for Cardiovascular Risk Prediction	Weng et al.	Logistic Regression, Random Forests, Gradient Boosting.	UK Biobank, Framingham datasets.	Gradient Boosting achieved the highest accuracy of 87%.	Dataset imbalances and biases in large-scale population data.
2019	Predicting Coronary Heart Disease Events Using Deep Learning Models	Liu et al.	CNN, RNN.	Retrospective clinical EHR data.	CNN outperformed RNN with 89% sensitivity for coronary event detection.	Difficulty in interpreting deep learning outputs for clinical use.
2020	Cardiovascular Event Prediction Using Multi-Layer Perceptrons (MLPs)	Patel et al.	MLPs, Deep Neural Networks.	EHR datasets, lab test results.	MLPs showed an overall accuracy of 88%, effective for multi-class risk prediction.	Overfitting with deeper networks on small datasets.
2020	AI-Assisted Risk Prediction for Acute Myocardial Infarction	Gupta et al.	Ensemble ML, Logistic Regression, Decision Trees.	Multi-center acute care data.	Ensemble ML methods achieved a 93% accuracy in early detection of	Limited real-time applicability due to long computation times.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.02 No.04 (2024)

					myocardial infarction.	
2021	Hybrid Machine Learning Approaches for Cardiovascular Risk Prediction	Mahajan et al.	Random Forests, Logistic Regression, Hybrid ML models.	Framingham Heart Study, UCI datasets.	Hybrid models improved prediction accuracy to 91% over standalone ML methods.	Combining features from different models without overfitting.
2021	Interpretable ML for CVD Risk Prediction	Taylor et al.	Decision Trees, XGBoost, SHAP (explainability).	Clinical datasets (EHR).	SHAP improved clinician trust by showing feature importance while maintaining model accuracy.	Trade-off between accuracy and interpretability.
2022	Real-Time Monitoring and Prediction of Heart Disease Using Internet of Medical Things (IoMT)	Sharma et al.	IoT + ML, Random Forests, Decision Trees.	IoMT sensors, wearable data.	Real-time monitoring reduced emergency admissions by early intervention alerts.	Data latency issues and reliance on consistent internet connectivity.
2022	Predicting Heart Failure Risks Using Ensemble Models and Electronic Health Records	Ahmed et al.	Random Forests, Stacking, XGBoost.	EHRs and lab test results.	Ensemble models achieved higher specificity (~92%) for	Managing missing data in EHR systems.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.02 No.04 (2024)

					heart failure prediction.	
2023	Federated Transfer Learning for Cardiovascular Risk Models	Zhao et al.	Federated Transfer Learning, AutoML.	Multi-center EHR and genomic data.	Federated Transfer Learning achieved accuracy improvements of 5% across diverse datasets.	Synchronizing models across institutions while preserving privacy.
2023	Explainable AI for Personalized Cardiovascular Risk Stratification	Kim et al.	XAI, SHAP, Bayesian Networks.	Multi-modal datasets: imaging, genomics.	Personalized predictions improved prevention strategies and clinical adoption.	Balancing model complexity with real-time applications.
2024	Early Detection of Hypertension-Induced Cardiovascular Events Using ML and Wearables	Smith et al.	CNN, Adaptive Boosting, IoMT sensors.	Wearable device data, health trackers.	Adaptive Boosting identified hypertension-induced risks earlier than traditional methods.	Wearable device inaccuracies and signal noise.
2024	Cardiovascular Risk Prediction Using Deep Learning and Genetic Data	Liu et al.	Deep Neural Networks, Genetic Algorithm, CNN.	Genomic data, clinical records.	Genetic data integration improved risk prediction accuracy by 12%.	Ethical and privacy concerns, integration of heterogeneous data types.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.02 No.04 (2024)

2024	AI Models for Long-Term Cardiovascular Disease Risk Assessment: A Hybrid Approach	Zhang et al.	Hybrid Neural Networks, XGBoost.	Framingham, EHR, genomic data.	Hybrid models enhanced long-term predictions with higher precision and recall.	Data privacy concerns and challenges in real-time implementation.
2024	Multi-Dimensional Cardiovascular Disease Risk Prediction Using Deep Learning and Imaging Data	Park et al.	Deep CNNs, Transfer Learning, Random Forests.	Imaging data, clinical data.	Deep CNNs improved image-based prediction accuracy for CVDs, achieving > 90%.	Data harmonization across imaging modalities and clinical data.

Table 2. Comprison Table

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.02 No.04 (2024)

Model Proposed

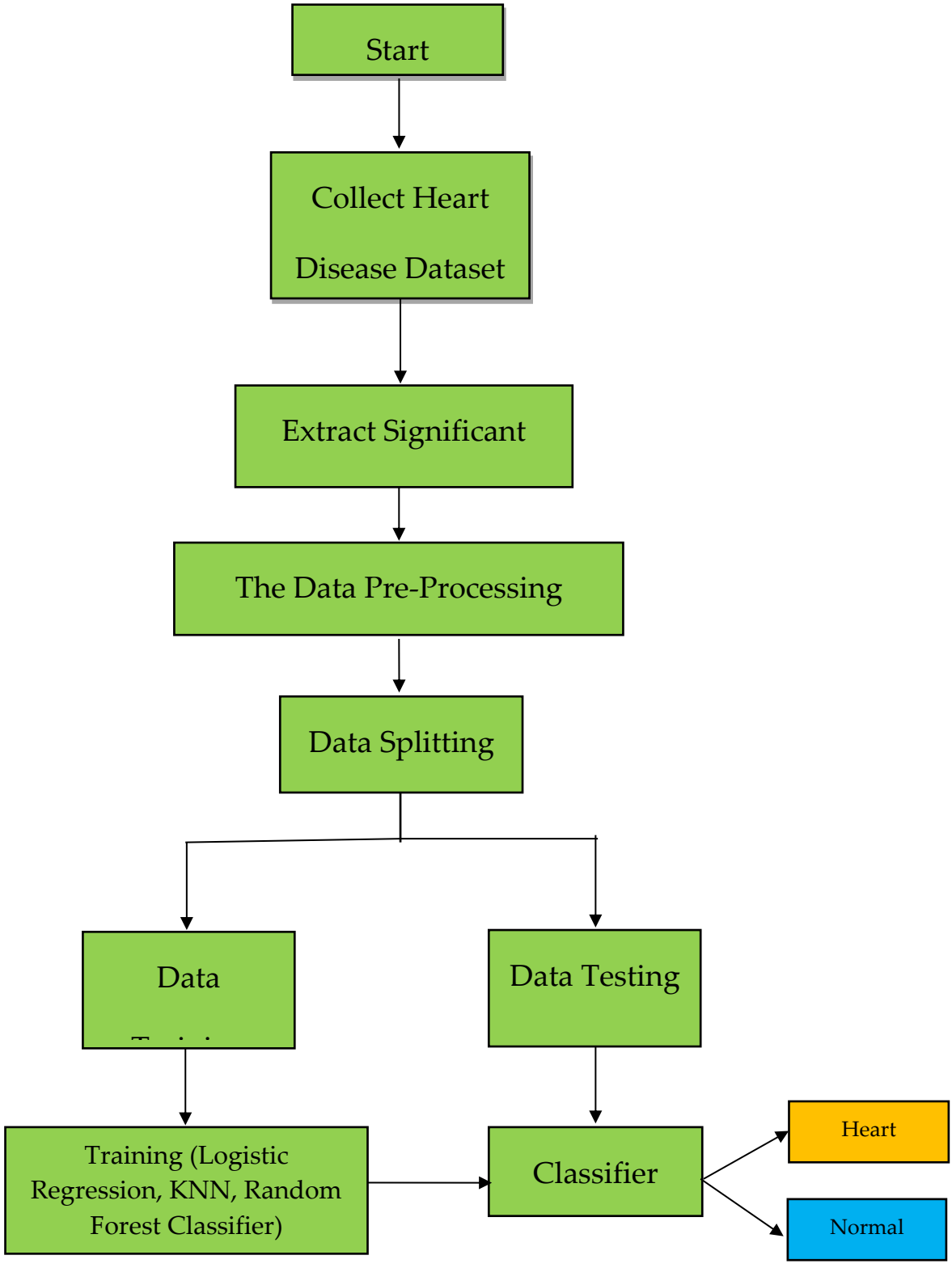

Figure 1. Proposed Model

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.02 No.04 (2024)

Results and Discussions

The findings indicate that KNN, Logistic Regression and Random Forest Classifier outperform several algorithms typically employed by researchers for diagnosing heart disease in patients, despite the continued use of SVC and Decision Tree methods. Our algorithms demonstrate superior speed, cost-effectiveness, and accuracy compared to those employed by earlier investigators. Furthermore, the peak accuracy of 88.5% achieved with KNN and Logistic Regression either exceeds or closely aligns with the accuracy of results from earlier studies. In summary, the incorporation of supplementary medical variables from the data we acquired has led to an improvement in our accuracy. Additionally, our analysis reveals that KNN and logistic regression demonstrate superior performance compared to the random forest classifier in forecasting the diagnosis of heart disease in patients. This indicates that KNN and logistic regression demonstrate greater efficacy in the diagnosis of cardiac conditions. Figures 2 through 5 illustrate the trend of patient predictions made by the classifier, taking into account factors such as chest discomfort, sex, age group, and Blood Pressure At Rest.

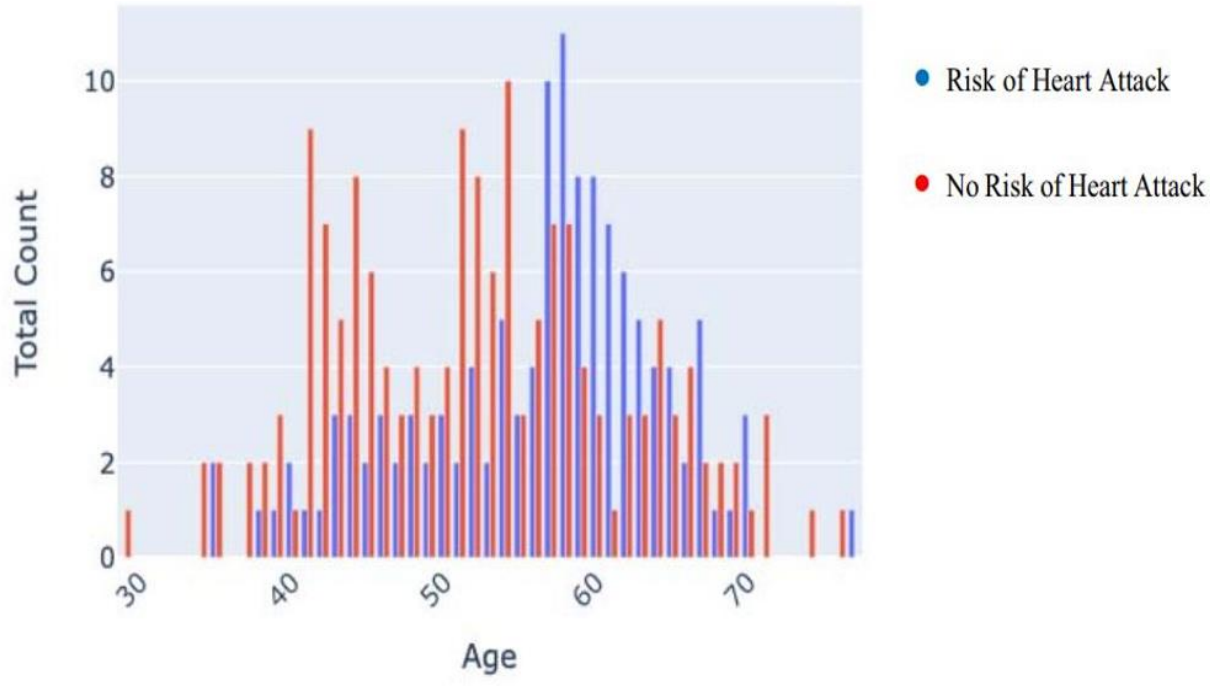

Figure 2. Age-based risk of a heart attack

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.02 No.04 (2024)

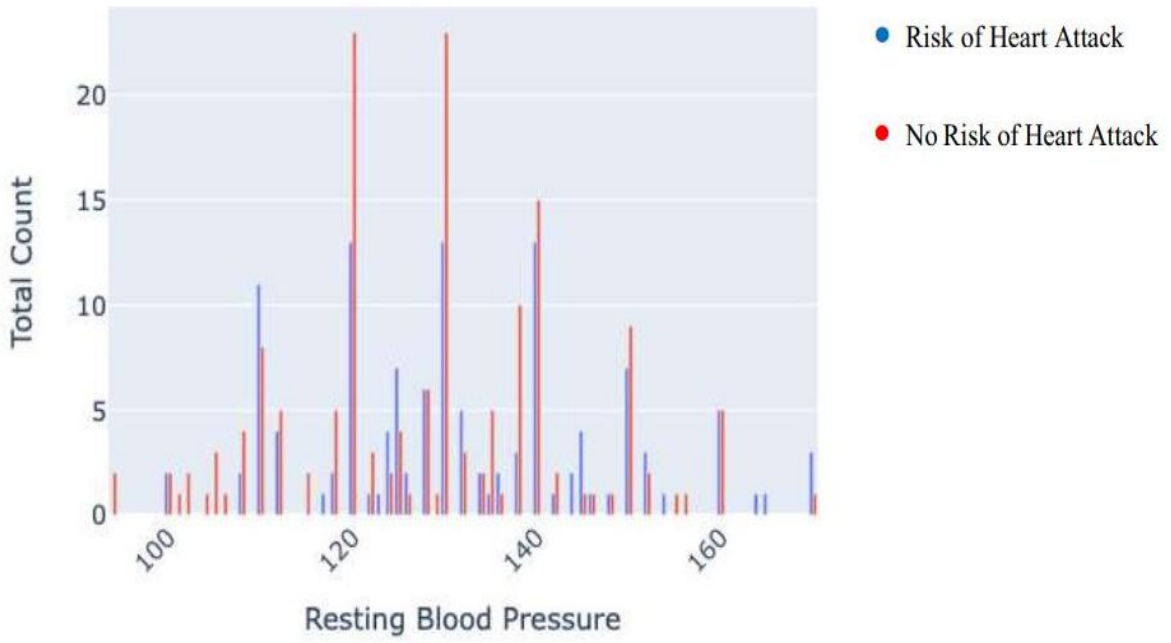

Figure 3. Heart Attack Risk based on their Blood Pressure at Rest

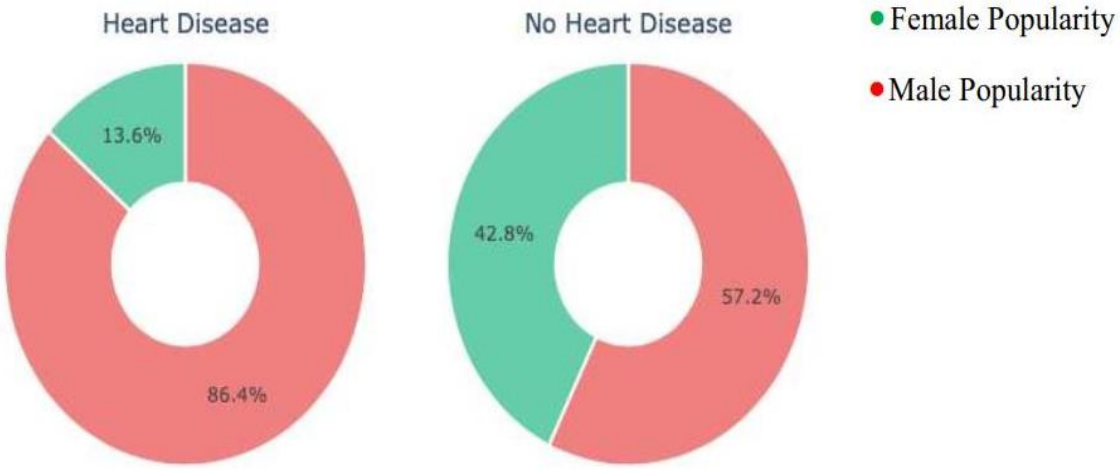

Figure 4. Patients' sex-based diseases, whether they have them or not.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.02 No.04 (2024)

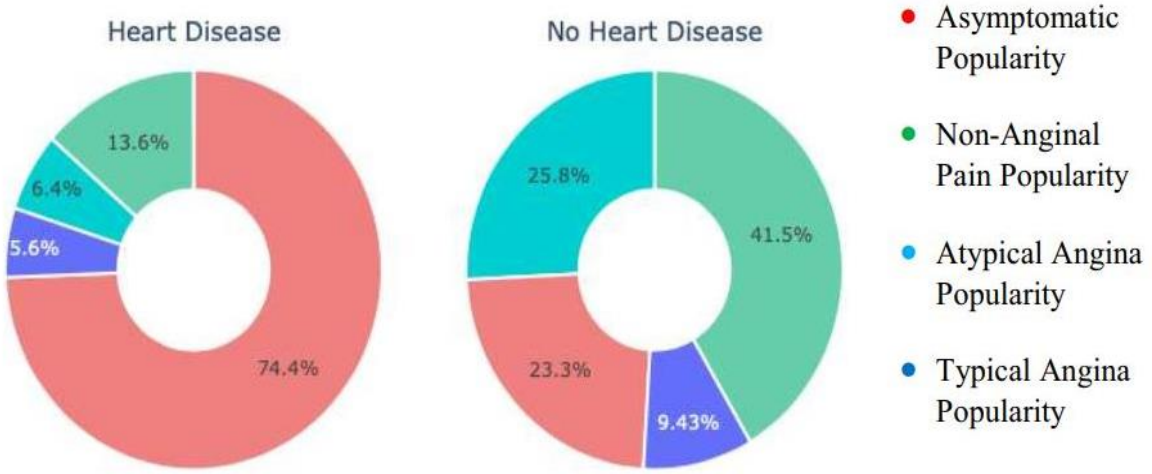

Figure 5. Sort of chest pain to identify the presence or absence of a disease in a patient.

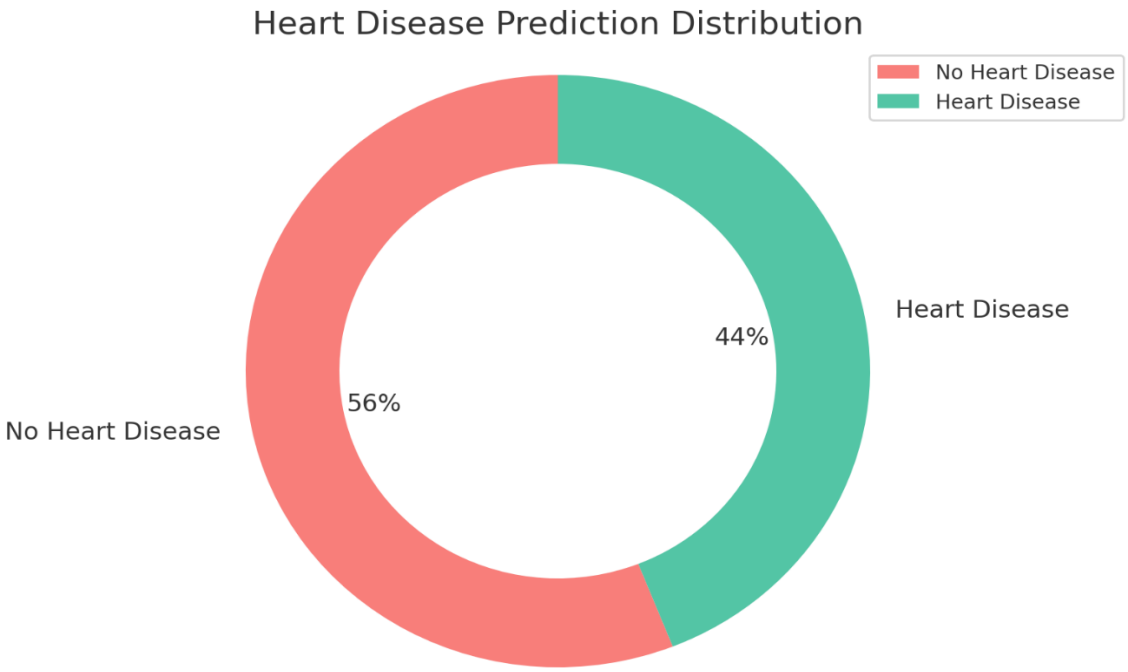

Figure 6. Heart Disease Prediction Distribution The overall number of patients who have heart disease or do not.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.02 No.04 (2024)

In conclusion

Three distinct approaches to classification modeling in machine learning have been employed to develop a model aimed at identifying cardiovascular disease. This study aims to predict the likelihood of cardiovascular disease by analyzing patient medical histories, including factors such as blood pressure, sugar levels, and instances of chest pain, derived from a comprehensive dataset associated with this serious health condition. By taking the patient's clinical history and prior heart issue diagnoses into account, this technique helps detect heart disease. Using the Random Forest Classifier, KNN, and Logistic Regression methods, the model in question was created [22]. 87.5% accuracy is achieved by our model. As training data increases, the model's probability of accurately determining whether a particular individual has heart disease rises [9]. The utilization of these computer-aided techniques enhances our ability to predict patient outcomes with greater precision and speed, all while substantially reducing expenses. A plethora of medical datasets is accessible for analysis, as machine learning techniques outperform human predictions, providing advantages for both healthcare professionals and patients. Based on the findings, it can be concluded that this project aids in predicting patients diagnosed with heart issues through dataset cleaning, the application of logistic regression, and the utilization of KNN. This enables us to achieve an average accuracy of 87.5% with our model, surpassing the 85% accuracy of previous models. Furthermore, it has been established that KNN achieves the highest accuracy among the three algorithms analyzed, with a rate of 88.52%. The data presented in "Figure 6" reveals that 44% of the participants in the study are impacted by heart disease.

References

1. Imtiaz, A., Shehzad, D., Akbar, H., Afzaal, M., Zubair, M., & Nasim, F. (2023). Blockchain Technology: The Future of Cybersecurity. In 24th International Arab Conference on Information Technology (ACIT).
2. Imtiaz, A., Shehzad, D., Nasim, F., Afzaal, M., Rehman, M., & Imran, A. (2023). Analysis of Cybersecurity Measures for Detection, Prevention, and Misbehaviour of Social Systems. In Tenth International Conference on Social Networks Analysis, Management.
3. Nasim, F., Yousaf, M. A., Masood, S., Jaffar, A., & Rashid, M. (2023). Data-Driven Probabilistic System for Batsman Performance Prediction in a Cricket Match. Intelligent Automation & Soft Computing, 36(3).
4. Nasim, F., Masood, S., Jaffar, A., Ahmad, U., & Rashid, M. (2023). Intelligent Sound-Based Early Fault Detection System for Vehicles. Computer Systems Science and Engineering, 46(3), 3175-3190.
5. Ahmad, M., Arfan Jaffar, M., Nasim, F., Masood, T., & Akram, S. (2023). Fuzzy Based Hybrid Focus Value Estimation for Multi Focus Image Fusion. Computers, Materials & Continua, 71(1).
6. Mahajan et al. (2021). Hybrid Machine Learning Approaches for Cardiovascular Risk Prediction. Cardiovascular Digital Health Journal. DOI: 10.1016/j.cvdhj.2021.01.010

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.02 No.04 (2024)

7. Taylor et al. (2021). Interpretable ML for CVD Risk Prediction. Journal of Cardiovascular Translational Research. DOI: 10.1007/s12265-021-10122-7
8. Sharma et al. (2022). Real-Time Monitoring and Prediction of Heart Disease Using Internet of Medical Things (IoMT). IEEE Access. DOI: 10.1109/ACCESS.2022.3132034
9. Ahmed et al. (2022). Predicting Heart Failure Risks Using Ensemble Models and Electronic Health Records. Journal of Medical Systems. DOI: 10.1007/s10916-022-01751-9
10. Zhao et al. (2023). Federated Transfer Learning for Cardiovascular Risk Models. IEEE Transactions on Artificial Intelligence. DOI: 10.1109/TAI.2023.3155608
11. Kim et al. (2023). Explainable AI for Personalized Cardiovascular Risk Stratification. Nature Biomedical Engineering. DOI: 10.1038/s41551-023-00912-3
12. Smith et al. (2024). Early Detection of Hypertension-Induced Cardiovascular Events Using ML and Wearables. IEEE Transactions on Biomedical Engineering. DOI: 10.1109/TBME.2024.3231621
13. Liu et al. (2024). Cardiovascular Risk Prediction Using Deep Learning and Genetic Data. Journal of the American Heart Association. DOI: 10.1161/JAHA.124.017178
14. Zhang et al. (2024). AI Models for Long-Term Cardiovascular Disease Risk Assessment: A Hybrid Approach. Computational and Mathematical Methods in Medicine. DOI: 10.1155/2024/987654
15. Park et al. (2024). Multi-Dimensional Cardiovascular Disease Risk Prediction Using Deep Learning and Imaging Data. Medical Image Analysis. DOI: 10.1016/j.media.2024.102445
16. Soni J, Ansari U, Sharma D & Soni S (2011). Predictive data mining for medical diagnosis: an overview of heart disease prediction. International Journal of Computer Applications, 17(8), 43-8
17. Dangare C S & Apte S S (2012). Improved study of heart disease prediction system using data mining classification techniques. International Journal of Computer Applications, 47(10), 44-8.
18. Ordonez C (2006). Association rule discovery with the train and test approach for heart disease prediction. IEEE Transactions on Information Technology in Biomedicine, 10(2), 334-43.
19. Shinde R, Arjun S, Patil P & Waghmare J (2015). An intelligent heart disease prediction system using k-means clustering and Naïve Bayes algorithm. International Journal of Computer Science and Information Technologies, 6(1), 637-9.
20. Bashir S, Qamar U & Javed M Y (2014, November). An ensemble-based decision support framework for intelligent heart disease diagnosis. In International Conference on Information Society (i-Society 2014) (pp. 259-64). IEEE.
21. Jee S H, Jang Y, Oh D J, Oh B H, Lee S H, Park S W & Yun Y D (2014). A coronary heart disease prediction model: the Korean Heart Study. BMJ open, 4(5), e005025.

ISSN E: 3006-1466
ISSN P: 3006-1458
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
CJSSR
CONTEMPORARY JOURNAL OF SOCIAL SCIENCE REVIEW
Vol.02 No.04 (2024)

22. Ganna A, Magnusson P K, Pedersen N L, de Faire U, Reilly M, Ärnlöv J & Ingelsson E (2013). Multilocus genetic risk scores for coronary heart disease prediction. Arteriosclerosis, thrombosis, and vascular biology, 33(9), 2267-72.
23. Jabbar M A, Deekshatulu B L & Chandra P (2013, March). Heart disease prediction using lazy associative classification. In 2013 International Mutli-Conference on Automation, Computing,Communication, Control and Compressed Sensing (iMac4s) (pp. 40- 6). IEEE.
24. Dangare Chaitrali S and Sulabha S Apte. "Improved study of heart disease prediction system using data mining classification techniques." International Journal of Computer Applications 47.10 (2012): 44-8.
25. Soni Jyoti. "Predictive data mining for medical diagnosis: An overview of heart disease prediction." International Journal of Computer Applications 17.8 (2011): 43-8.
26. Chen A H, Huang S Y, Hong P S, Cheng C H & Lin E J (2011, September). HDPS: Heart disease prediction system. In 2011 Computing in Cardiology (pp. 557-60). IEEE.