E-ISSN : 2988-585X (Online)

Journal of Elektronik Sistem InformasI (JESII)
Volume 2 No 2 Desember 2024
DOI : 10.31848/jesii.xxxx.xxxx

Prediction and Analysis of Heart Failure using Machine Learning Techniques

Mehr Ali Qasimi[1]
[1]Department of Information System, Badakhshan University, Afghanistan

Article Info

Article history:

Received Oct 20, 24
Revised Dec 30, 24
Accepted Dec 31, 24

Keywords:

Analysis
Classification Algorithms
Heart Failure
Machine Learning
Prediction

ABSTRACT

Heart failure, a leading cause of global mortality with approximately 17.9 million deaths annually, arises from impaired myocardial function that hinders the heart's ability to pump oxygen- and nutrient-rich blood effectively. Current risk prediction methods are limited by their reliance on traditional statistical approaches, which fail to capture complex interactions in large, multidimensional datasets. This study aims to enhance survival analysis and mortality prediction for heart failure patients using machine learning techniques. A dataset of 1,000 patients with 12 key features was analyzed using supervised learning algorithms, including Naïve Bayes, decision trees, K-nearest neighbors (KNN), and random forests. The results revealed that KNN provided the highest predictive accuracy, effectively identifying significant patient characteristics associated with mortality risk. This approach demonstrates the potential of machine learning in improving prognosis accuracy and guiding early interventions. Future work should explore the inclusion of larger datasets, real-time applications, and advanced deep learning models to further refine predictive capabilities and support clinical decision-making. the integration of machine learning into heart failure prognosis enhances predictive accuracy and provides a scalable framework for leveraging multidimensional patient data. Future research should explore the inclusion of larger and more diverse datasets, real-time monitoring capabilities, and the application of deep learning models to further refine prediction efficacy and support clinical decision-making.

Corresponding Author:

Mehr Ali Qasimi,
Information System Department, Faculty of Computer Science, Badakhshan University.
Badakhshan University, Badakhshan, Faizabad, Afghanistan. 3401
Email: q1.mehrali@gmail.com

1. INTRODUCTION

Heart failure is an incurable illness. a dangerous condition when the heart's low cardiac output is the result of its inability to pump blood as effectively as it should. The ability of the heart to pump blood and supply the cells with oxygenated, nutrient-rich blood is crucial to the body's ability to operate properly[1]. Heart failure is divided into three categories based on how much blood is pumped out with each beat. Heart failure caused by left ventricular systolic dysfunction, or HFrEF, is represented by an EF < 40% and is referred to as type 1 heart failure. Type 2 heart failure is defined as having an EF between 40 and 49 percent and a mid-range ejection fraction (HfmrEF). Heart failure type 3 is defined as diastolic heart failure with an EF of at least 50%, or heart failure with preserved ejection fraction, or HFpEF[2]. High blood pressure, smoking, obesity, diabetes, ischemic heart disease, and increased cholesterol are the main risk factors for developing the failing condition[3]. Depending on the age, gender, race, and ethnicity, heart failure can present with different characteristics.

The evaluation of the patient's medical history, physical examination report, and analysis of concerning symptoms by medical professionals form the basis of invasive-based approaches for detecting heart disease. Because human error can lead to delays in diagnosis outcomes, all of these procedures typically result in erroneous diagnoses. Not only that, but it requires extra time and money for examinations and is computationally demanding[4]. According to studies, many HF patients receive the wrong diagnosis for a variety of causes. Early patient diagnosis also presents additional challenges. Patients' therapy is hampered by the low rate of early HF diagnosis. Optimizing treatment therefore requires improving the HF diagnosis, particularly in the early stages[5].

Medical organizations worldwide gather information on a range of health-related matters. Various machine learning algorithms can be applied to these data to obtain valuable insights. However, a lot of noisy data is frequently included in the vast amount of data that has been gathered. Several machine learning approaches can be used to easily examine these datasets, which are too large for human minds to process. As a result, these algorithms have developed into highly valuable tools for precisely predicting the existence or absence of heart-related disorders in recent times[6]. Heart failure (HF) has complicated origins, symptoms, and surgical and medicinal therapy. Therefore, heart failure (HF) still presents a challenge to the health care system even after several studies, recommendations, and innovative medicines have been implemented. This area may serve as an excellent example of how Machine Learning (ML) approaches can be used to find potential correlations that traditional research methods have not been able to find[7]. Among the data mining approaches that have drawn attention from research groups are classification algorithms. Treatments and interventions can be administered effectively and specifically when disease stage, etiology, or subtypes are accurately classified. This also makes it possible to evaluate the patient's progress. A variety of data mining techniques have been used to focus on heart failure (HF) in order to distinguish HF patients from controls, identify various HF subtypes, and calculate the severity of HF (NYHA class) (Fig. 1).

Furthermore, data mining techniques can be helpful even in cases of late-stage heart failure (HF), when the chances of survival and treatment advantages are reduced. This is because they enable prompt prediction of readmission risk, morbidity, and death[8]. With the use of different machine learning models and a range of inputs, machine learning techniques may be able to predict and classify individuals who have cardiac disease. But it's been claimed that sophisticated systems that build databases of these patients and help doctors predict how serious a patient's condition would be are needed in hospitals all around the world. The goal of the current project is to increase the examining physician's confidence by assessing the illness using an intelligent system that offers an evaluation based on multiple independent tools running on an embedded system. The next paragraph provides a comprehensive overview of the literature to demonstrate the current level of interest in this area[9].

Our goal in this study is to use the patient's data to forecast whether or not he is suffering from heart disease. The data contains the patient's personal information, including age, sex, cholesterol, resting electrocardiogram, etc. Because illness prediction can, in the worst situation, potentially result in death, it is an extremely delicate undertaking. As a result, we will analyze the algorithms' accuracy alongside the forecasts and then provide the optimal algorithm for the prediction. To find out if they are predisposed to heart disease or not, people can keep these prediction APIs on their home desktop computers or the computers in the hospital reception area[10].

2. METHOD

Researchers have suggested a number of machine learning-based diagnosis methods for HD in the literature. In order to highlight the significance of the suggested work, this research study presents a few Machine learning-based diagnosis strategies that are now in use [11]. Classification is a crucial technique for making decisions in the medical sciences. This topic has been the focus of numerous studies. Numerous studies have been conducted on the Pima Indian diabetes database, according to a survey of related literature. Michie et al. [12] study is the most thorough of them all. For classification and risk prediction of heart failure survival, numerous studies have used machine learning algorithms such as ensemble learning approaches, boosting algorithms, Random Forest, Artificial Neural Networks, Support Vector Machines, Logistic Regression, Multi-Layer Perceptrons,and Random Forests. In order to forecast hospital readmissions, numerous studies have also been carried out. We talk about comparable research that have recently conducted in this area. Heart failure is a critical global health challenge, responsible for approximately 17.9 million deaths annually, making it one of the leading causes of mortality worldwide. The condition arises when the heart cannot pump blood effectively, depriving the body's tissues and organs of essential oxygen and nutrients, which are crucial for maintaining physiological balance. Despite advancements in medical diagnostics and treatment strategies, accurately predicting and analyzing heart failure outcomes remains a significant challenge. Existing risk prediction models predominantly rely on conventional statistical techniques, which are often limited in their ability to handle the complex, multidimensional nature of healthcare datasets. These methods typically fail to capture intricate relationships between diverse patient attributes, leading to moderate prediction accuracy and suboptimal decision support for clinicians.

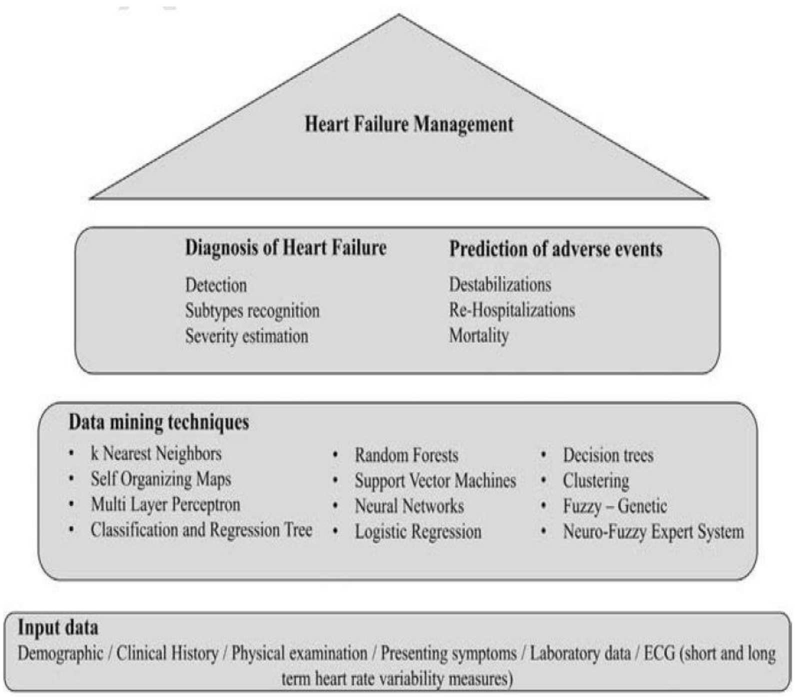

Figure. 1. Overview of studies on heart failure management[8].

An expert medical diagnosis approach for HF identification was proposed by Palaniappan et al. [11]. Artificial Neural Networks (ANN), Decision Trees (DT), and Navies Bays (NB) are a few examples of the predictive machine learning models that were employed in the system's development. NB achieved 86.12% accuracy, ANN achieved 88.12% accuracy, and the DT classifier achieved 80.4% accuracy.In order to diagnose heart failure with maintained ejection fraction, Tabassian et al. [13] developed a framework by examining the spatiotemporal patterns of echocardiograph curves using statistical modeling and machine learning techniques. The ML model with an area under the curve of 0.89 was based on the distance-weighted k-nearest neighbors. In Vosough et al.'s study [14], six different algorithms were compared for their ability to predict hospital readmission for patients with heart failure: SVM, least-square SVM, random forest, bagging classifiers, ada-boost, and naïve bayes. Of these, random forest demonstrated the best performance. Tang et al. [15] created granular support vector machines, a novel learning model for data classification issues. The idea behind granular support vector machines is to formally and methodically integrate the ideas of granular computing theory and statistical learning theory. SVM and GSVM have accuracy rates of 83.04% and 84.04% [15]. suggest a machine learning classification method for diagnosing cardiac disease in the context of e-healthcare. A variety of categorization strategies are employed in the study to improve identification accuracy. With an accuracy rate of 90.47%, precision value of 0.909, and recall value of 0.912, the Random Forest classifier performs admirably. This approach focuses a lot of emphasis on applying machine learning methods to the field of e-healthcare, particularly to correctly diagnose heart disease patients. A complex machine-learning approach was presented by Alsafi et al. to identify coronary heart disease [11]. This method efficiently uses patient features to provide accurate diagnoses by analyzing patient data using machine learning techniques. A remarkable degree of accuracy 97.4% is shown by the simulation findings. This indicates how their methods can help in the diagnosis of coronary heart disease. Ren Y. created a hybrid neural network that combines self-encoding network with two-way long-term memory. The model was used to predict renal illness in hypertensive patients, with an accuracy of 89.7% using data from 35,332 individuals [16].

Heart disease can be predicted using certain characteristics that are obtained from the patient; in order to do so, medical records and prior patient histories are also gathered [17]. They made advantage of pictures from electronic medical records. Hospital, doctor, and diagnostic patient records are all included in the electronic health record (EHR). They obtained some output in the form of unstructured visual data based on the EHR records, and by locating and examining the connections between the events, they are eventually able to forecast the date of a patient's diagnosis. Data from electronic health records are few, thus we are unable to forecast or analyze them. Unstructured and non-standardized data are extracted from electronic health records. Because the available data are scarce, using them directly will be challenging. In order to identify cardiac disease, Gudadhe et al. presented a system that combines multilayer perceptron neural network architecture and Support Vector Machine. Using the Support Vector Machine to indicate if cardiac disease was present or absent, they separated the database into two groups. While the artificial neural network classifies the heart disease data into 5 with an accuracy of 97.5%, they only achieved 80.41% accuracy [18].

A hybrid strategy was presented by Kanika Pahwa and Ravinder Kumar to choose features from a dataset on heart disease for predictive purposes. The SVM-RFE and gain ratio were used by the author in their feature selection technique to eliminate superfluous and unnecessary features. Determining the characteristics is crucial for forecasting. To categorize the data set as having heart disease or not, they applied Random Forest and Naïve Bayes to a subset of the features. When they used specific traits, they were able to obtain findings with more precision. Naive Bayes reached 84.15% accuracy, whereas Random Forest achieved 84.16%. Both approaches produced results that were comparable in terms of accuracy.When predicting short-term time series data, the auto regressive integral moving average model performs better and is appropriate for numerical sequences. but a neural network can be used to solve problems for non-numerical time series, this approach is not very efficient, but it can produce correct results [19]. Various results obtained from these studies suggest that machine learning methods could be well equipped to provide meaningful, accurate, and explainable risk prediction techniques which could not only save time but also increase the patient outcomes and satisfaction levels. We require a system which can predict the possibility more accurately.

2.1. DATASET

The dataset consists of 918 actual examples of data with 12 different attributes (11 predictors; 1 class) such as Age, Sex, Chest Pain Type, RestingBP, Cholesterol, FastingBS, RestingECG and so on (Table 1). This project helps to predict whether the patient has heart disease or not. This prediction is made using the clinical data of patients[20]. The main research process of this study included Six parts: data collection, EDA (Exploratory Data Analysis), Data Preprocessing, construction of classification model, transfer study and visualization. To obtain the causes of heart disease and build a model with the highest level of accuracy, we have employed Five Classification algorithms in this study.

Table 1. Attributes and details of dataset of heart disease

Content	Should be Fulfilled
Age	Age of a patient [years]
Sex	Gender of the patient [M: Male, F: Female]
ChestPain	Chest pain type [TA: Typical Angina, ATA: Atypical Angina, NAP: Non-Anginal Pain, ASY: Asymptomatic]
RestingECG	Resting electrocardiogram results [Normal: Normal, ST: having ST-T wave abnormality (T wave inversions and/or ST elevation or depression of > 0.05 mV), LVH: showing probable or definite left ventricular hypertrophy by Estes' criteria].
MaxHR	Maximum heart rate achieved [Numeric value between 60 and 202]
ExerciseAngina	Exercise-induced angina [Y: Yes, N: No]

Content	Should be Fulfilled
Oldpeak	oldpeak = ST [Numeric value measured in depression]
ST_Slope	The slope of the peak exercise ST segment [Up: upsloping, Flat: flat, Down: downsloping]
HeartDisease	Output class [1: heart disease, 0: Normal]

2.2. DATA PREPROCESSING

Preprocessing of the data is required for data to be represented efficiently and for machine learning classifiers to be trained and tested effectively. The dataset has undergone preprocessing methods like standard scalar, MinMax Scalar, and missing value removal in order to make it suitable for use in classifiers. By ensuring that each feature has a mean of 0 and a variance of 1, the standard scalar brings all features to the same coefficient. Similar to this, MinMax Scalar adjusts the data so that every feature falls between 0 and 1 .Real-world data comprises noisy and high amounts of missing data. These data have been preprocessed to get around these problems and produce accurate forecasts. The sequential chart of our suggested model is described in Figure 2. Noise and missing values are typically present in cleaned data. The data must be cleaned up of noise and any missing values must be filled in in order to get an accurate and useful output. Transformation is the process of converting data from one format to another so that it can be understood better. It involves duties related to aggregation, standardization, and smoothing. Integration: The data must be integrated before processing because it may come from a variety of sources rather than just one. Decrease Acquiring complex data necessitates formatting in order to produce meaningful outcomes. After the data have been categorized, they are divided into training and test sets, which are then subjected to a variety of algorithms to produce accuracy score figures.

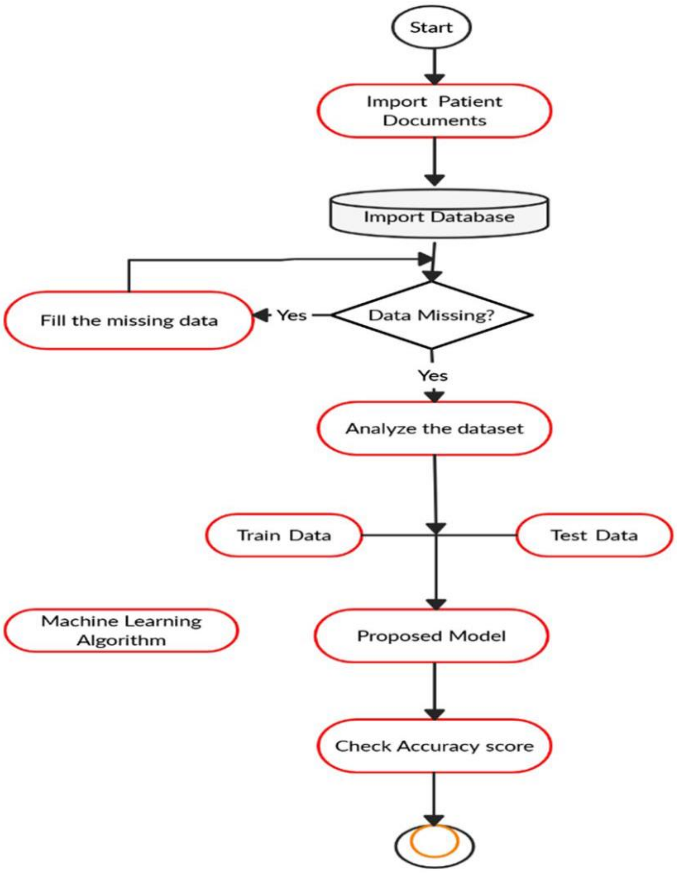

Figure. 2 Sequential charts of proposed model.

2.3. MODEL DEVLOPMENT

The most crucial aspect of machine learning techniques was the model generation process, wherein various approaches took into account the nature of the problem and the properties of the dataset. The following machine learning algorithms were applied, and their performance analyzed on the dataset for a comparative study for the best classifier to predict the existence of Heart Disease based on the various health parameters being considered [1]. These characteristics of the dataset led to the decision that using classification techniques in the study was appropriate. By examining the patterns in the training set of data, classification algorithms trained the model. This allowed it to classify the test data in a highly accurate way that had never been seen before. In this study, the Python 3 programming language was used to model Heart Failure kernels using the following classifiers: Naive Bais (NV), Support Vector Classifier (SVC), Random Forest (RF), Decision Tree (DT) and K Nearest Neighbors (k-NN). These are machine learning techniques that are commonly applied to challenges involving classification.

2.3.1. Support Vector Classifier (SVC): To start with, a basic version of SVC with Radial Basis Function (RBF) kernel was implemented[1].

2.3.2. Naive Bayes: The NB is a supervised learning technique for classification. The conditional probability theorem provides the foundation for figuring out a new feature vector's class. the NB determines the conditional probability value of vectors for a particular class using the training dataset. Each vector's probability conditional value is calculated first, and then the new vector class is determined using the conditionality probability of each vector. For the classification of text-related problems, NB is employed [4].

2.3.3. Decision Tree: Decision Tree is a type of supervised machine learning algorithm. A tree with each node acting as a decision or leaf node is called a decision tree shape. When making decisions, the decision tree's techniques are straightforward and simple to comprehend. Both exterior and internal nodes were connected to one another in a decision tree. The child node that visits the subsequent nodes and the decision-making portion of the system are known as the internal nodes. In contrast, the leaf node is connected to a label and lacks any offspring nodes [21].

2.3.4. Random Forest: A decision tree-based supervised learning approach. Each tree is constructed using a distinct sample and evaluates each one according to the majority vote for classification [1].

2.3.5. K-Nearest Neighbor: K-NN is a classification technique for supervised learning. The K-NN algorithm uses the new input's similarity to its input samples in the training set to predict the class label of a new input. if the training set's samples and the new input are identical. The classification performance of K-NN is subpar. Given an observation x, h(x) can identify the value of y. Let (x, y) be the training observations and h: X⟶Y be the learning function [21].

3. RESULT AND DISCUSSION

This section of the report covered the classification models and results from several models. Initially, we evaluated the effectiveness of several machine learning methods on the Heart Failure dataset using all features. Performance evaluation indicators were used to assess the effectiveness of the classifiers. On an Intel(R) Core(TM) i7-8665U CPU running at 1.90GHz (8 CPUs), approximately 2.1GHz, all calculations were done using Python, with the aim to build a model that will be applied further to make prediction and analysis of a test dataset. At first we have analyzed the data to found the null values (as show in figure 3), duplicate values, unnamed columns, normalization, standardization, correleation between features (as shown in figure 4) and also we have plotted confusion matrix of each algorithm. we have implemented five different classification algorithms on this dataset and every algorithms had different outputs and accuracy, Random forest had the most high accuracy among five classification algorithms. In second step various statistical operations, including mean extraction, standard division, Min-Max scalar, standard scale, and attribute removal with missing values, have been performed on the dataset. The dataset that has been analyzed consists of 918 instances, 11 input attributes, and one output label. The display of data in a graphical style is known as data visualization. It facilitates clear and effective communication of information by condensing and presenting vast amounts of data in an understandable style, hence aiding in people's comprehension of the significance of the data.

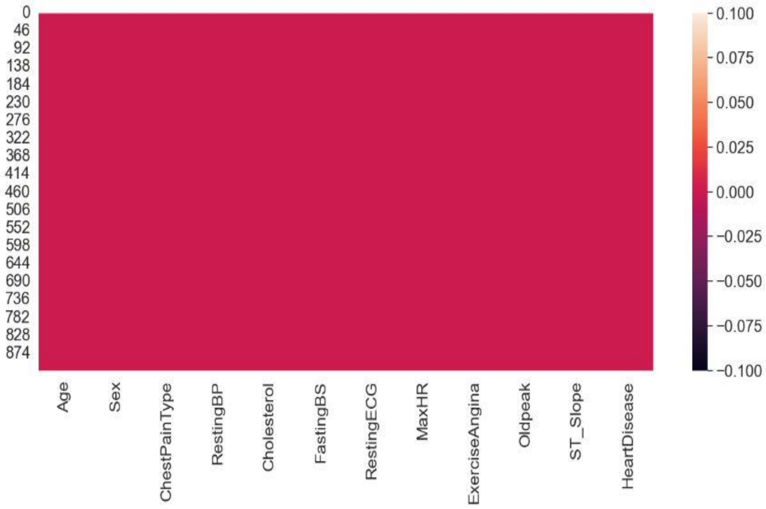

Figure. 3 Counting null values of each feature

This study, drawn from this research offer a novel contribution by demonstrating the superior predictive accuracy of the K-nearest neighbors (KNN) algorithm for heart failure mortality risk, specifically within the context of a dataset comprising 1,000 patient instances and 12 attributes. While previous studies have highlighted the potential of machine learning in healthcare, this research uniquely emphasizes the comparative performance of multiple algorithms, including Naïve Bayes, decision trees, and random forests, identifying KNN as the most effective in this specific application. By leveraging multidimensional data and uncovering significant patient features associated with heart failure, this study extends established findings by providing a more nuanced understanding of predictive capabilities and highlighting the practical applicability of machine learning techniques in clinical settings.

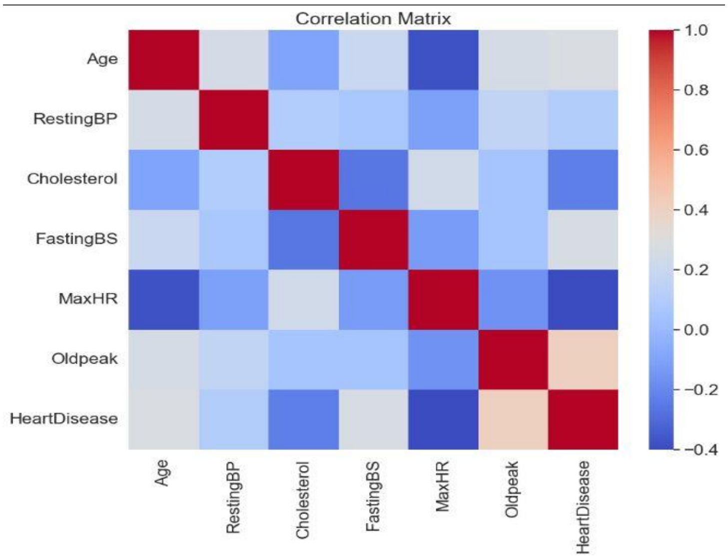

Figure. 4 visualizing the correlation matrix of the numerical columns

There are 918 instances, 11 input attributes, and one output label in the examined dataset. As seen in figure 5, we counted the cases of heart disease among these 918 instances 508 were positive cases and 410 were negative.

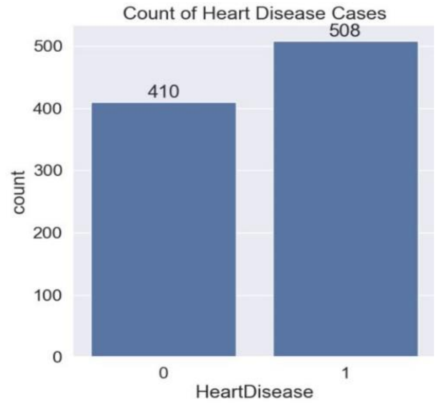

Figure. 5 Counting the positive and negative cases

I also computed the heart failure based on the patients' ages, and as Figure 6 illustrates, the majority of the 918 affected individuals were between the ages of 55 and 70.

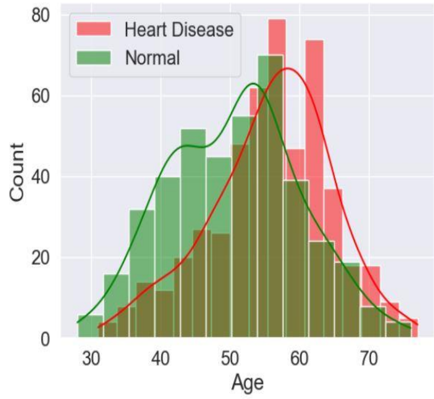

Figure. 6 Distribution plot of Age for Heart Disease .

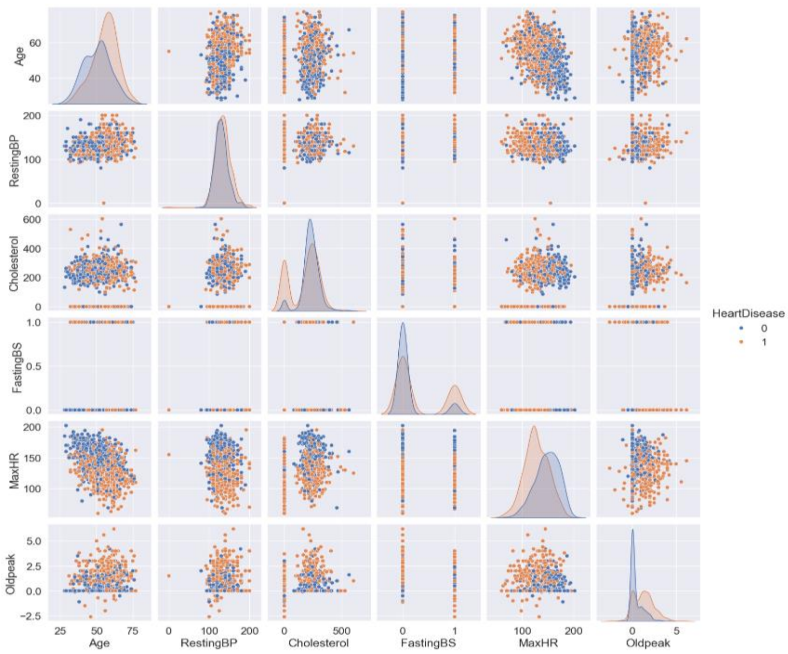

Figure. 7 Histograms of heart disease dataset.

Based on the computation, there are 918 patients total—725 men and 193 women. Figure 8 illustrates that of the 725 men, 458 have positive heart disease cases and 267 have negative instances, while of the 193 female patients, 50 have heart disease and 143 do not.

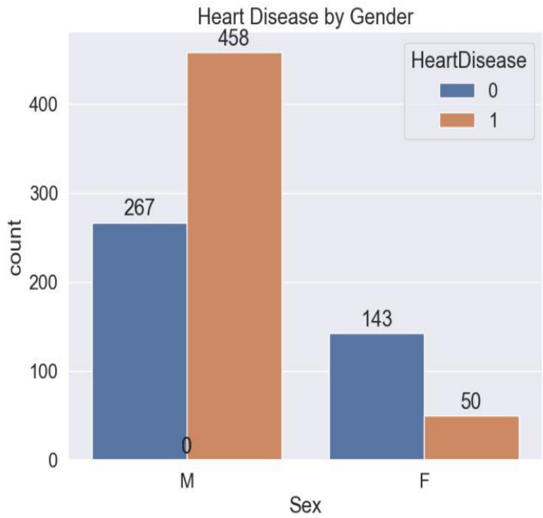

Figure. 8 Heart Disease by gender.

Using five different algorithms, I have discovered that Random Forest is the most effective method. It has the highest accuracy of any algorithm, as seen in Figure 9.

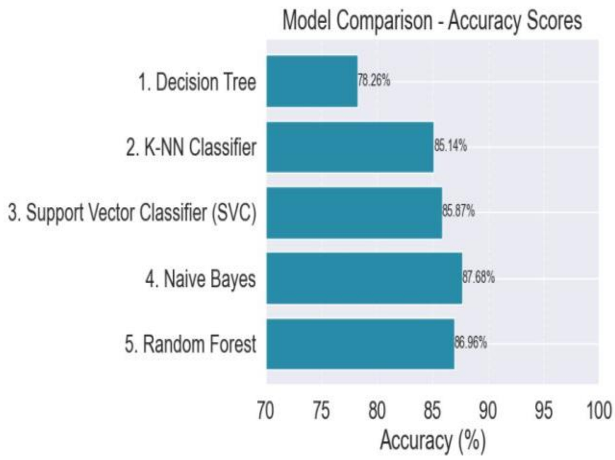

Figure. 9 Model Comprasion – Accuracy Scores.

The study included several experiments that were conducted. To forecast heart disease in patients diagnosed with heart failure, a comprehensive exploratory data analysis was conducted, and then a detailed heart disease analysis was performed. Following separate analysis of each characteristic, the most significant factors influencing survival odds were found to be age, fasting blood sugar, cholesterol, kind of chest pain, and resting blood pressure. We also did some data preprocessing and visualization.

With respect to health factors, a range of machine learning models were tested for survival prediction. Because it might cause target leakage, prediction time for heart disease was removed from the characteristics list used to train the models. ML models such as Naïve Bayes, K-Nearest Neighbor, SVC, decision trees, and random forests were trained using the training set. Random Forest, Naïve Bayes, and SVC were the models that outperformed the others, with respective accuracy values of 88.41, 87.68, and 85.87. The least number of incorrect predictions was produced by KNN and Decision Tress. The feature explanation analysis reveals the patterns in the data that the model considers when making a determination. The most crucial elements of the assessment are covered in this passage.

4. CONCLUSION

In conclusion, heart failure has a high mortality and morbidity rate and is always associated with a higher risk of developing many chronic cardiovascular illnesses. When utilized effectively, the vast amounts of data generated by the medical field—including EHR/EMR records, clinical notes, medical histories, pathological test results, and radiological image data—can reveal numerous hidden patterns that can be used to improve AI solutions for cardiovascular disease diagnosis and treatment proposal development. By offering an accurate prognosis, hospital readmission risks, and survival estimates, as was previously said, such clever solutions can lower the danger of heart failure condition. The integration of technology and data can effectively mitigate treatment inequities, enhance care planning, and lead to better health results. The strategic planning, creation, and implementation of AI-powered smart health solutions would improve medical protocols, enable doctors to think beyond the box, and ultimately help to raise patient satisfaction with healthcare in general as well as heart failure conditions. High-level patient education and awareness campaigns that highlight the advantages of a heart-healthy diet and an active lifestyle may also be beneficial in lowering the death rate from heart failure-related illnesses.

REFERENCES

[1] S. Mishra, "A Comparative Study for Time-to-Event Analysis and Survival Prediction for Heart Failure Condition using Machine Learning Techniques," J. Electron. Electromed. Eng. Med. Informatics, vol. 4, no. 3, pp. 115–134, 2022, doi: 10.35882/jeeemi.v4i3.225.

[2] J. F. Nauta, X. Jin, Y. M. Hummel, and A. A. Voors, "Markers of left ventricular systolic dysfunction when left ventricular ejection fraction is normal," Eur. J. Heart Fail., vol. 20, no. 12, pp. 1636–1638, 2018, doi: 10.1002/ejhf.1326.

[3] A. L. Bui, T. B. Horwich, and G. C. Fonarow, "Epidemiology and risk profile of heart failure," Nat. Rev. Cardiol., vol. 8, no. 1, pp. 30–41, 2011, doi: 10.1038/nrcardio.2010.165.

[4] A. U. Haq, J. P. Li, M. H. Memon, S. Nazir, R. Sun, and I. Garciá-Magarinõ, "A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms," Mob. Inf. Syst., vol. 2018, 2018, doi: 10.1155/2018/3860146.

[5] G. Yang et al., "A heart failure diagnosis model based on support vector machine," Proc. - 2010 3rd Int. Conf. Biomed. Eng. Informatics, BMEI 2010, vol. 3, no. Bmei, pp. 1105–1108, 2010, doi: 10.1109/BMEI.2010.5639619.

[6] V. V. Ramalingam, A. Dandapath, and M. Karthik Raja, "Heart disease prediction using machine learning techniques: A survey," Int. J. Eng. Technol., vol. 7, no. 2.8 Special Issue 8, pp. 684–687, 2018, doi: 10.14419/ijet.v7i2.8.10557.

[7] "Prediction of Heart Failure Mortality using Machine Learning Algorithms Rodriguez M , 2021, January.," no. Ml, pp. 1–10, 2021.

[8] E. E. Tripoliti, T. G. Papadopoulos, G. S. Karanasiou, K. K. Naka, and D. I. Fotiadis, "Heart Failure: Diagnosis, Severity Estimation and Prediction of Adverse Events Through Machine Learning Techniques," Comput. Struct. Biotechnol. J., vol. 15, pp. 26–47, 2017, doi: 10.1016/j.csbj.2016.11.001.

[9] A. Domyati and Q. Memon, "Machine Learning Based Improved Heart Disease Detection with Confidence," Int. J. online Biomed. Eng., vol. 19, no. 8, pp. 130–143, 2023, doi: 10.3991/ijoe.v19i08.37417.

[10] A. Chaturvedi, S. Haseeb, and U. Rahman, "Comparitive Analysis for Various Classification Machine Learning Algorithms for Detecting Heart Diseases," www.irjmets.com @International Res. J. Mod. Eng., vol. 533, no. 06, pp. 533–539, 2021, [Online]. Available: www.irjmets.com

[11] J. P. Li, A. U. Haq, S. U. Din, J. Khan, A. Khan, and A. Saboor, "Heart Disease Identification Method Using Machine Learning Classification in E-Healthcare," IEEE Access, vol. 8, no. Ml, pp. 107562–107582, 2020, doi: 10.1109/ACCESS.2020.3001149.

[12] H. Kahramanli and N. Allahverdi, "Design of a hybrid system for the diabetes and heart diseases," Expert Syst. Appl., vol. 35, no. 1–2, pp. 82–89, 2008, doi: 10.1016/j.eswa.2007.06.004.

[13] M. Tabassian et al., "Diagnosis of Heart Failure With Preserved Ejection Fraction: Machine Learning of Spatiotemporal Variations in Left Ventricular Deformation," J. Am. Soc. Echocardiogr., vol. 31, no. 12, pp. 1272-1284.e9, 2018, doi: 10.1016/j.echo.2018.07.013.

[14] R. Najafi-Vosough, J. Faradmal, S. K. Hosseini, A. Moghimbeigi, and H. Mahjub, "Predicting hospital readmission in heart failure patients in Iran: A comparison of various machine learning methods," Healthc. Inform. Res., vol. 27, no. 4, pp. 307–314, 2021, doi: 10.4258/HIR.2021.27.4.307.

[15] Y. Tang, B. Jin, Y. Sun, and Y. Q. Zhang, "Granular support vector machines for medical binary classification problems," Proc. 2004 IEEE Symp. Comput. Intell. Bioinforma. Comput. Biol. CIBCB'04, pp. 73–78, 2004, doi: 10.1109/cibcb.2004.1393935.

[16] W. Chang, Y. Liu, X. Wu, Y. Xiao, S. Zhou, and W. Cao, "A New Hybrid XGBSVM Model: Application for Hypertensive Heart Disease," IEEE Access, vol. 7, pp. 175248–175258, 2019, doi: 10.1109/ACCESS.2019.2957367.

[17] B. Jin, C. Che, Z. Liu, S. Zhang, X. Yin, and X. Wei, "Predicting the Risk of Heart Failure with EHR Sequential Data Modeling," IEEE Access, vol. 6, no. c, pp. 9256–9261, 2018, doi: 10.1109/ACCESS.2017.2789324.

[18] M. Gudadhe, K. Wankhade, and S. Dongre, "Decision support system for heart disease based on support vector machine and artificial neural network," 2010 Int. Conf. Comput. Commun. Technol. ICCCT-2010, pp. 741–745, 2010, doi: 10.1109/ICCCT.2010.5640377.

[19] K. Pahwa and R. Kumar, "Prediction of heart disease using hybrid technique for selecting features," 2017 4th IEEE Uttar Pradesh Sect. Int. Conf. Electr. Comput. Electron. UPCON 2017, vol. 2018-Janua, pp. 500–504, 2017, doi: 10.1109/UPCON.2017.8251100.

[20] D. Shah, S. Patel, and S. K. Bharti, "Heart Disease Prediction using Machine Learning Techniques," SN Comput. Sci., vol. 1, no. 6, 2020, doi: 10.1007/s42979-020-00365-y.

[21] X. Wu et al., Top 10 algorithms in data mining, vol. 14, no. 1. 2008. doi: 10.1007/s10115-007-0114-2.