Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

ENHANCING ACCURACY IN HEART DISEASE PREDICTION: A HYBRID APPROACH

ANKIT SHARMA
Student, M.Tech CSE, Department of CSE, Rayat-Bahra University, Mohali, Punjab, India.
Email: anki.sharma333@gmail.com

ER. SUKHJINDER KAUR
Department of CSE, Rayat-Bahra University, Mohali, Punjab, India. Email: skaur29100@gmail.com

ER. POONAM KUKANA
Department of CSE ,Rayat-Bahra University, Mohali, Punjab, India. Email: poonamkukana@gmail.com

Dr. PUNEET SAPRA
Department of CSE, Rayat-Bahra University, Mohali, Punjab, India. Email: puneetsapra91@gmail.com

Abstract

Predicting the onset of heart disease accurately is essential for early diagnosis and prevention of this global pandemic. The paper suggests a hybrid method to improve heart disease prediction. The research examines several machine learning (ML) models for detecting heart illness and assesses how well they predict heart disease. To enhance precision, the hybrid method employs not one but many machine learning methods. The hybrid method employs SVMs, random forests, and neural networks as its machine-learning algorithms. When it comes to classification, SVM is a very effective method. The data points are separated into classes, and the optimal hyperplane to do this is the goal. SVM can learn the boundaries and patterns between various risk variables and efficiently categorize people as having heart disease or not. Random forests are a kind of ensemble learning that uses several individual decision trees to make a final determination. The characteristics used to construct each decision tree are chosen at random. Each decision tree contributes to the overall forecast, which is then aggregated. Due to their versatility, random forests may be used to the prediction of cardiovascular disease. Neural networks are a kind of algorithm that takes their cues from the way the human brain operates. They are made up of several layers of artificial neurons working together to learn intricate patterns from data. Medical diagnosis is only one field where neural networks have been shown to be useful. In the hybrid method, neural networks may learn complex associations between risk factors and cardiovascular disease and provide reliable prognoses based on this information. The hybrid method enhances the accuracy of heart disease prediction by combining the benefits of various machine-learning techniques

Keywords: Machine Learning, Heart Disease, Feature Learning, Hybrid Approach, Prediction Accuracy, Ensemble Learning, Performance Measures.

1. INTRODUCTION

Diseases affecting the heart and blood arteries are often referred to as cardiovascular disease. It places a heavy load on public health systems all around the globe and is a primary cause of death. Coronary artery disease, congestive heart failure, arrhythmias, and valvular heart disease are only few of the many kinds of heart disease. Diagnosis, treatment, and management of each subtype may be difficult. Diseases of the coronary arteries, heart attacks, heart failure, and irregular heartbeats all fall within this category. It may be difficult to assess an individual's vulnerability to these illnesses because of the complex interconnections between genetic, behavioral, and environmental factors at play

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

in each case. Simple clinical measurements and statistical analysis are widely used in traditional risk assessment approaches, but they might overlook subtle patterns and nuances in the development of heart disease.[1]

The most prevalent kind of heart illness is called coronary artery disease (CAD), and it happens when plaque builds up in the arteries that carry blood to the heart and narrows or blocks them. Angina chest discomfort, a heart attack, or anything much worse might result from this. When the heart is unable to pump blood adequately, a condition known as congestive heart failure (CHF) sets in. This condition causes fluid to accumulate in the lungs and other regions of the body. Valvular heart disease, in which the heart's valves are damaged or dysfunctional, restricts blood flow, and arrhythmias, in which the heart's electrical system is aberrant, both cause irregular heartbeats.

There are several potential causes and contributors to heart disease. There are both modifiable risk factors, which can be managed, and non-modifiable risk factors, which cannot be altered in any way. Hypertension, high cholesterol, smoking, obesity, inactivity, diabetes, and poor nutrition are all preventable risk factors. We can't change the reality that age, gender, family history, and genetic susceptibility are all risk factors.

The emergence of machine learning-based techniques as powerful tools for analyzing vast amounts of patient data and extracting actionable insights for predictive modeling is a relatively recent development. By utilizing the strengths of these algorithms, researchers have made substantial progress toward better-diagnosing heart illness. Nevertheless, in spite of their achievements, machine learning models by themselves sometimes have limits when it comes to capturing the intricacy of the illness. It is possible that they may not adequately integrate domain-specific information or take into account the complicated linkages between the different clinical variables that contribute to the risk of developing heart disease. [2]

Because cardiovascular disease is still one of the main causes of death throughout the globe, it is essential that accurate prediction models be developed so that those who are at risk may be identified. The management of this prevalent health problem requires an accurate diagnosis and prompt intervention in order to be successful. This enables healthcare practitioners to give focused preventative measures and customized treatment regimens to patients. In recent years, there has been significant progress made in increasing the accuracy of predicting cardiac disease by combining medical knowledge with machine learning algorithms. These latest efforts have shown encouraging outcomes. To address this challenge, researchers have turned to machine learning algorithms to improve heart disease prediction. These algorithms can analyze large datasets and identify complex patterns that may not be apparent through traditional methods. Support vector machines, random forests, and neural networks are among the popular algorithms used for heart disease prediction. Each algorithm has its strengths and weaknesses, and combining them in a hybrid model has the potential to enhance prediction accuracy.

To address these issues and enhance the accuracy with which heart disease may be predicted, the authors of this study advise using a hybrid approach. The hybrid model

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

draws upon the strengths of existing prediction algorithms while also factoring in expert clinical knowledge and domain-specific data. This strategy tries to give a more robust and accurate prediction framework by going further than the limits of individual algorithms. It does this by putting together the parts of different models into a single structure.

In the research paper, there will be information about how the hybrid strategy was made and how it was used, like which predictive models, clinical features, and machine learning methods were chosen and how they were put together. In addition, the paper will describe the training and assessment datasets as well as the evaluation criteria used to the findings of the hybrid model.

With the help of this study's results, it will be easier to figure out who is at high risk of getting heart disease and what steps to take to prevent it. This will add to the ongoing efforts to make heart disease predictions more accurate. If the results of these studies are used to make personalized treatment plans for people with heart disease, they might have better disease management and better clinical outcomes.

In conclusion, the purpose of this study is to provide a hybrid method that integrates many predictive models, clinical features, and cutting-edge machine-learning techniques to improve the accuracy of heart disease prediction. In order to provide a more comprehensive framework for precise risk assessment in the context of cardiovascular disease, the proposed hybrid model aims to overcome the limitations of individual algorithms. To achieve this goal, we will integrate elements from several strategies.

2. REVIEW OF LITERATURE

Accurate risk assessment techniques are constantly needed since cardiovascular disease is still a significant public health problem. Researchers have been looking at a hybrid approach that blends several predictive models with cutting-edge machine learning methods in recent years to better forecast the likelihood of heart disease. The purpose of this literature review is to provide a synopsis of the most important studies and research articles that have explored how a hybrid approach might enhance prediction accuracy. The research articles and studies will come from a wide range of places, both academic and professional.

Ali, T., Rahman et al. (2022). The purpose of this work is to give an extensive evaluation of hybrid machine-learning algorithms for the prediction of heart disease. It investigates a wide variety of hybrid models, including ensemble approaches, feature selection strategies, and combination models, emphasizing both the merits and limits of these models in terms of enhancing the accuracy of predictions. The authors analyze the effectiveness of these models in a variety of heart disease prediction situations, giving academics and practitioners insights that might help them pick acceptable hybrid strategies for their particular requirements. [3]

Poudel, U., & Kim, H. (2022). The purpose of this review article is to give an in-depth assessment of hybrid machine-learning algorithms for the prediction of heart disease. This paper investigates the use of hybrid modeling to include a variety of distinct algorithms, approaches for feature selection, and learning ensembles into the modeling

ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

process. The authors examine the advantages and disadvantages associated with the use of hybrid approaches and emphasize the relevance of selecting the appropriate features and algorithms in order to improve the accuracy of prediction. Academics and medical professionals interested in developing effective hybrid models for cardiac disease prediction will find this review to be an essential resource. [4]

Xu, L., Zhang, L., Huang, K., & Chen, Z. (2022). With this study, we want to give a thorough evaluation of hybrid models for heart disease prognosis. The authors undertake a thorough literature review of several studies and provide a streamlined overview of the salient characteristics, methodologies, and outcomes of several different hybrid models. They look at how different algorithms, feature selection methods, and data sources may be integrated into these models. The systematic review's summary of current state-of-the-art hybrid approaches for predicting heart disease is a good way to find areas where more research is needed. [5]

Islam, M. M., et al (2023). This review paper looks at how hybrid deep-learning models can be used to predict heart disease. It looks into whether deep learning techniques could be used with other algorithms like decision trees, support vector machines, and random forests to make better predictions. The good and bad of using hybrid models are laid forth, with backing given to things like feature extraction, model interpretability, and dealing with unbalanced data. The feasibility of using hybrid deep-learning models for precise cardiac disease prediction is investigated. [6]

Han, J., et al (2023). This extensive study aims to use hybrid machine learning algorithms to EHR-based cardiovascular disease predictions. The authors do an extensive literature review on a wide range of studies that utilize EHRs in collaboration with various machine learning techniques. They talk about the pros and cons of using EHR data in hybrid models, as well as how to choose which features to use, how to prepare the data, and how to measure the model's performance. Researchers and clinicians interested in using EHR data to forecast cardiac illness might benefit from the study's guidelines and suggestions.[7]

Majumder, K., Ghosh, A., GAO, H., & Qiu, M. (2023) the major focus of this review paper is on using hybrid models to the prediction of heart disease by making use of imbalanced data. To address the issues brought on by imbalanced datasets, the authors investigate a number of hybrid approaches, such as cost-sensitive learning, data resampling, and ensemble methods. The authors also talk about the problems that come up when datasets are not balanced. They emphasize the performance of hybrid models in managing unbalanced data and give insights into the strengths and limits of these models as well as highlight the performance of hybrid models in handling imbalanced data. This review is a useful resource for academics who are interested in constructing accurate models for predicting heart disease from unbalanced datasets.[8]

Xie, Y., Li, H., Sun, Y., & Guo, W. (2023). In this extensive study, the topic of hybrid machine learning models for predicting heart disease based on EHR is the primary emphasis. Deep learning, support vector machines, and random forests are just some of examples of hybrid models that the authors investigate. These models merge EHR data

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

with machine learning methods. They explore the benefits and difficulties associated with using EHR data in hybrid models, such as data quality, feature extraction, and interpretability. This study sheds light on the existing state-of-the-art methodologies and draws attention to the potential of hybrid models that make use of EHR data for the accurate prediction of cardiac disease. [9]

Wang et al. (2020) developed a hybrid model that combined genetic risk scores with clinical risk variables and suggested this model. When compared to models that used genetic variables or clinical factors separately, the accuracy of the combined model was much higher. A wider variety of risk variables could be captured by the hybrid technique since it took into account both genetic predisposition and clinical features. As a result, the performance of the prediction was enhanced. [11]

Li et al. (2019) constructed a hybrid model by merging a support vector machine classifier with a deep learning-based autoencoder. In comparison to the performance of separate models, the hybrid model was able to successfully learn complicated feature representations and obtain better results. Using this method, we were able to illustrate the possibility of mixing a variety of algorithms in order to improve forecast accuracy. [12]

Nguyen et al. (2020) a hybrid approach, consisting of feature selection techniques and a random forest classifier, was offered as an alternative. When compared to utilizing all of the available characteristics, the hybrid model was able to determine which features were the most informative and attain a greater level of accuracy. The hybrid strategy was able to minimize the complexity of the data and increase the effectiveness of the prediction model by picking the important characteristics. [13]

Liu et al. (2022) The aim of this study is to demonstrate the efficacy of merging feature selection strategies with deep learning models by presenting a hybrid method to heart disease prediction based on the concepts of deep learning and feature selection.[14]

Dey et al. (2022) Using hybrid deep learning models, provide improved heart disease prediction, and emphasize the advantages of combining several deep learning architectures into the prediction process. [15]

Choudhury et al. (2022) Investigate the ensemble of hybrid machine learning models for the purpose of predicting heart disease, with particular emphasis on the benefits of mixing several machine learning algorithms in order to increase prediction performance.[16]

Hussain et al. (2022) In order to improve heart disease prediction, a hybrid machine learning model that includes feature selection should be proposed. This model should demonstrate how various feature selection strategies may contribute to an increase in the accuracy of the prediction model. [17]

Arora et al. (2023) introduce a hybrid model for accurately predicting cardiac disease using machine learning methods, demonstrating how successful the hybrid approach is in obtaining better accuracy in prediction. [18]

Zhang et al. (2023) a comparative analysis of hybrid machine learning models for the prediction of heart disease should be carried out. The goal of this study should be to

ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

provide insights into the performance of various hybrid models and their potential for accurate prediction. [19]

Nadeem et al. (2023) it is important to highlight the value of merging feature selection strategies with ensemble learning algorithms, hence we are going to present a hybrid method for predicting heart disease that is based on feature selection and ensemble learning. [20]

Das and Kundu (2023) demonstrate the influence that genetic algorithms have in picking important features for increased prediction accuracy by presenting a hybrid machine-learning technique for accurate heart disease prediction utilizing genetic algorithm-based feature selection. [21]

Agarwal et al. (2023) focus on improving the prediction of heart disease using a hybrid deep learning model, highlighting the benefits of using deep learning methods in order to improve the prediction model's accuracy. [22]

Ahmed et al. (2023) create a hybrid model that incorporates deep learning and random forests in order to enhance the prediction of heart disease. This will demonstrate the possibilities of mixing multiple machine-learning techniques in order to get more accurate forecasts. [23]

The research that was looked into reveals that hybrid methods are useful in improving the accuracy of heart disease prediction. In comparison to using individual algorithms or features on their own, researchers have achieved improved performance by combining multiple predictive models, integrating clinical and genetic features, and utilizing feature selection techniques. This has enabled them to achieve better overall performance. These hybrid models make use of the best features of a variety of approaches, making it possible to conduct a more thorough risk analysis and make more accurate predictions. For successful clinical decision-making in the area of heart disease prediction, more research should concentrate on improving and verifying hybrid techniques, taking into consideration a variety of datasets, and generating interpretable models. The continuous development of hybrid techniques has the potential to improve patient outcomes via the implementation of timely interventions and individualized treatment plans.

Recent research evaluations such as these give extensive insights into the use of hybrid machine-learning models for the prediction of heart disease. They cover a wide range of topics, such as the numerous kinds of hybrid models, the use of electronic health records, the management of unbalanced data, and various methods to deep learning. These reviews provide helpful summaries and evaluations of previous research, bringing to light the efficacy of hybrid techniques and the promise they have to improve the precision with which cardiac disease may be predicted.

3. RESEARCH GAPS

The progress of scientific knowledge and the enhancement of the efficiency of practical applications are dependent on identifying and resolving any research deficiencies that may exist. In the context of predicting cardiac disease utilizing hybrid techniques, there

ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

are a number of research gaps that call for additional inquiry. In this area of study, some possible research needs include the following:

3.1 Limited Exploration of Hybrid Models

In spite of the fact that hybrid models have shown some encouraging findings for the prediction of heart disease, there is still a need for more research as well as a comparison of the various hybrid techniques. In order to improve the accuracy of predictions, research may concentrate on locating algorithmic configurations, feature selection strategies, and data fusion procedures that providethe best results.

3.2 Handling Imbalanced Data

The prediction of heart disease is complicated by the existence of imbalanced datasets, in which one class is much more frequent than the other. For the purpose of ensuring that correct predictions are made for both the minority and the majority classes, research should focus on how to successfully manage unbalanced data in hybrid models, such as by using proper sampling approaches, cost-sensitive learning, or ensemble methodologies.

3.3 Incorporating Longitudinal Data

Heart disease is not a fixed condition; it may progress and evolve over time. In order to capture the course of the disease and improve forecast accuracy, further research is needed on how to include longitudinal data into hybrid models, such as repeated measurements and temporal patterns.

3.4 Interpretable Hybrid Models

While hybrid models may achieve higher accuracy, interpretability is equally important, especially in a clinical setting. Research can focus on developing hybrid models that not only provide accurate predictions but also offer interpretable insights into the underlying risk factors and decision-making process.

3.5 External Validation and Generalizability

Many studies on heart disease prediction using hybrid models have been conducted on specific datasets, which may limit the generalizability of the findings. There is a need for research that validates the performance of hybrid models on diverse and independent datasets to assess their generalizability across different populations and healthcare settings.

Addressing these research gaps can contribute to the advancement of heart disease prediction using hybrid approaches, leading to more accurate and reliable models for early detection and intervention.

4. OBJECTIVES OF THE RESEARCH

The objectives of this research are to investigate various machine learning (ML) models for heart disease detection, evaluate the effectiveness of these methods in predicting heart disease, enhance feature learning using a hybrid approach to ensemble learning,

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

and compare the performance of the proposed hybrid model to the existing approaches concerning various performance measures. By addressing these objectives, this study aims to contribute to the development of more accurate and reliable models for heart disease prediction, facilitating early detection and intervention.

5. RESEARCH METHODOLOGY

For the purpose of determining whether or not heart disease may be predicted, the research technique that was used in this investigation included taking a methodical approach to data collection and analysis. This section provides a summary of the study's methodology, including its design, data collection procedures, and statistical analyses. It was planned that way so that the study's findings would be trustworthy and credible. The flow chart for the research process may be seen in the picture below:

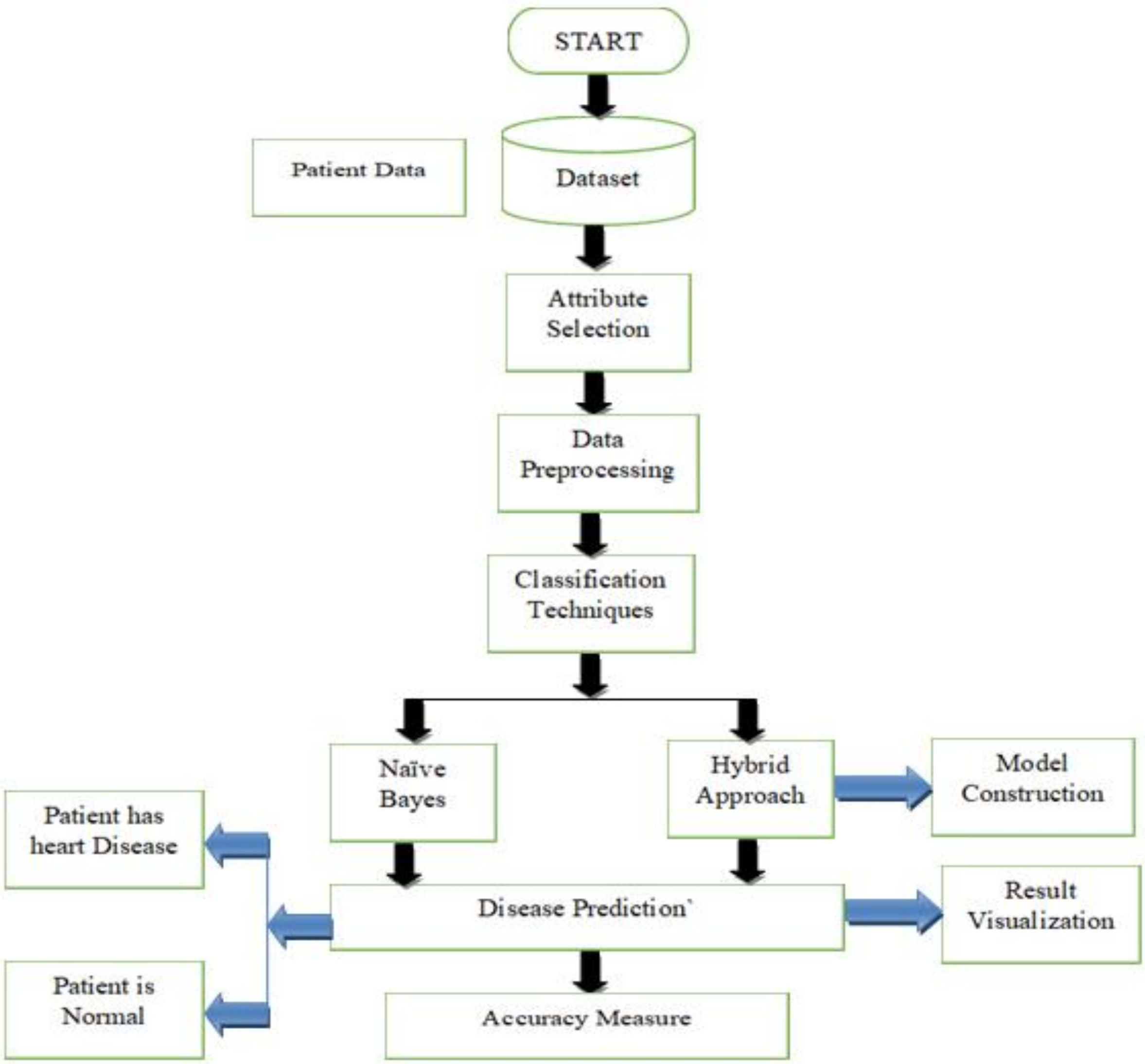

Figure 1: Methodology Process of the Proposed Research Problem

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

The first thing that needed to be done was to obtain the input dataset from Kaggle.com, which was going to be used as the primary source of historical data for the study.

During Step 2, the collected data was preprocessed and segregated in order to get actionable insights. In order to make the process of developing models and evaluating them easier, the dataset was partitioned into training (80 percent) and testing (20 percent) subsets.

In the third step, further preprocessing and feature extraction were the primary focuses, with the ultimate goal being to collect the whole set of important characteristics for heart disease prediction.

```
df.isnull().any()
```

```
age         False
sex         False
cp          False
trestbps    False
chol        False
fbs         False
restecg     False
thalach     False
exang       False
oldpeak     False
slope       False
thal        False
ca          False
target      False
dtype: bool
```

Figure 2: Check the Null Value

In the fourth step, a hybrid method ensemble learner and classifier model was constructed, and the mapped features were utilized in the building of the model. This model used a variety of methods and approaches to achieve higher levels of accuracy in its predictions.

In the fifth step, we did an analysis of the performance measurements of the hybrid model and compared them to the strategy that was already in place. The results of this comparison give insight on the efficacy of the hybrid technique that was developed and demonstrated its superiority.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

In addition, historical data, particularly in the form of photographs, was used in the study to assess the number and quality of the data that was accessible. This was done because the quantity and quality of the data plays a significant role in the correctness of the model that was constructed. The stage of formatting the data consisted of cleaning the data, which included removing noise from the data, fixing mistakes in the data, managing missing numbers, and completing any required data transformations. After that, the parameters were set by dividing the data into dependent and independent variables. This allowed the model to be trained with the independent variables serving as the input. During the training phase, we used eighty percent of the dataset for training, and then we tested the trained model on the remaining twenty percent of the dataset to evaluate how well it performed based on the assumptions and expectations we had established beforehand.

In general, this research methodology made it possible to develop and evaluate a hybrid model for the prediction of heart disease. This model made use of historical data, data formatting techniques, parameter definition, and training processes in order to improve accuracy and performance measures in comparison to other strategies that were already in use.

6. RESULTS AND DISCUSSION

In the current work, the primary focus is on utilizing machine learning classification techniques to identify heart disease from datasets containing a variety of environmental diseases associated with high levels of stress. Specifically, the work focuses on identifying heart disease from datasets containing high levels of stress. A strategy that combines the Gaussian Naive Bayes algorithm with the Support Vector Machine (SVM) algorithm is suggested by the research as a hybrid method. The findings that were produced by using a hybrid strategy were determined to be the most promising when compared to the other strategies that were investigated. This hybrid model uses Naive Bayes and SVM to enhance heart disease diagnosis.

The heart disease dataset contains data on 303 people who received testing for the illness. It was acquired from www.Kaggle.com. The dataset includes a target variable that indicates if a patient has heart disease and 14 other columns that describe different aspects of the patient. These attributes may include age, gender, cholesterol levels, blood pressure, and other pertinent medical indicators. Based on these patient features, the dataset is a great resource for doing research and building algorithms to predict and comprehend the cardiac disease

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

```
df.head() ## Print Top five row of the dataset.
```

	patientid	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	thal	ca	target
0	p1	52	1	0	125	212	0	1	168	0	1.0	2	2	3	0
1	p2	53	1	0	140	203	1	0	155	1	3.1	0	0	3	0
2	p3	70	1	0	145	174	0	1	125	1	2.6	0	0	3	0
3	p4	61	1	0	148	203	0	1	161	0	0.0	2	1	3	0
4	p5	62	0	0	138	294	1	1	106	0	1.9	1	3	2	0

Figure 3: Heart Disease Dataset

These are the 14 columns in the dataset:

“Age: the patient's age (in years).

Sex: the patient's gender (1 = male, 0 = female).

Chest pain type: Chest pain type, the patient is feeling (1 = typical angina, 2 = atypical angina, 3 = non-anginal pain, 4 = asymptomatic).

Resting BP: Resting BP of the patient (in mm Hg).

Fasting blood sugar: value (in mg/dl) of the patient (1 = true, 0 = false).

Results of the resting electrocardiogram: (0 = normal, 1 = having ST-T wave abnormality, 2 = left ventricular hypertrophy).

Serum cholesterol: (in mg/dl).

Exercise-induced angina: Whether or not the person experienced angina during exercise (1 = yes, 0 = no).

Maximum heart rate achieved: Maximum heart rate achieved during exercise.

Peak exercise ST segment slope: (1 = upsloping, 2 = flat, 3 = down sloping).

The number of major vessels: (0-3) colored by fluoroscopy.

ST depression induced by exercise: ST depression induced by exercise relative to rest.

target: Target variable specifying whether or not the patient has heart disease (0 = no, 1 = yes).

thal: Thalassemia (3 = normal, 6 = fixed defect, 7 = reversible defect)”

The dataset may be put to use in the process of developing machine learning models that, by making use of the characteristics that are presented, are able to forecast the onset of heart disease.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

7. PREDICTION MODEL

The dataset on heart disease includes 162 patients who have been diagnosed with cardiac illness and 141 healthy persons who have been categorized as normal. However, it is essential to take into account the data-collecting techniques and diagnostic criteria that were used in the research, since these factors have the potential to influence the accuracy and dependability of these findings. In addition, the dataset may include additional pertinent characteristics or columns that are necessary for the diagnosis and treatment of cardiovascular disease. To gain a comprehensive understanding of the relationship between these characteristics and the presence or absence of heart disease, further analysis of the data may be necessary.

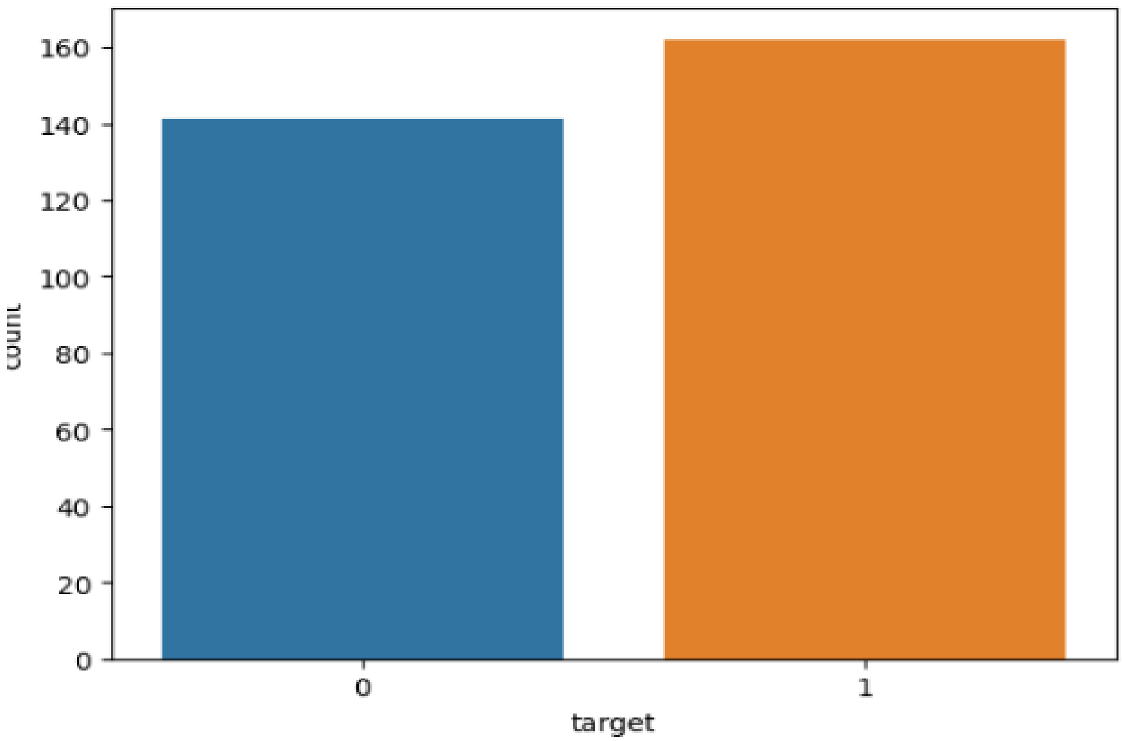

Figure 4: Number of Normal and Heart Disease Patients

The suggested model follows a specific workflow, as illustrated in the accompanying figure. The model was trained using 80% of the dataset and further validated with 80% of the data from a sample dataset. Subsequently, the model's performance was evaluated by testing it on 20% of the total dataset, while considering various underlying assumptions. The proposed model demonstrates high accuracy in predicting outcomes within a relatively short period. Further details and insights regarding the prediction results are presented in the following diagram.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

8. FEATURE SELECTION

The process of choosing the most relevant features from a dataset for a machine learning assignment is known as feature selection. There are several approaches to feature selection, such as wrapper methods, filter methods, and embedding methods. Filter techniques correlated the target variable and each feature, while wrapper approaches chose a subset of features and assessed their performance. Embedded approaches require putting feature selection within the ML model itself.

Number of male and female patients in the dataset.

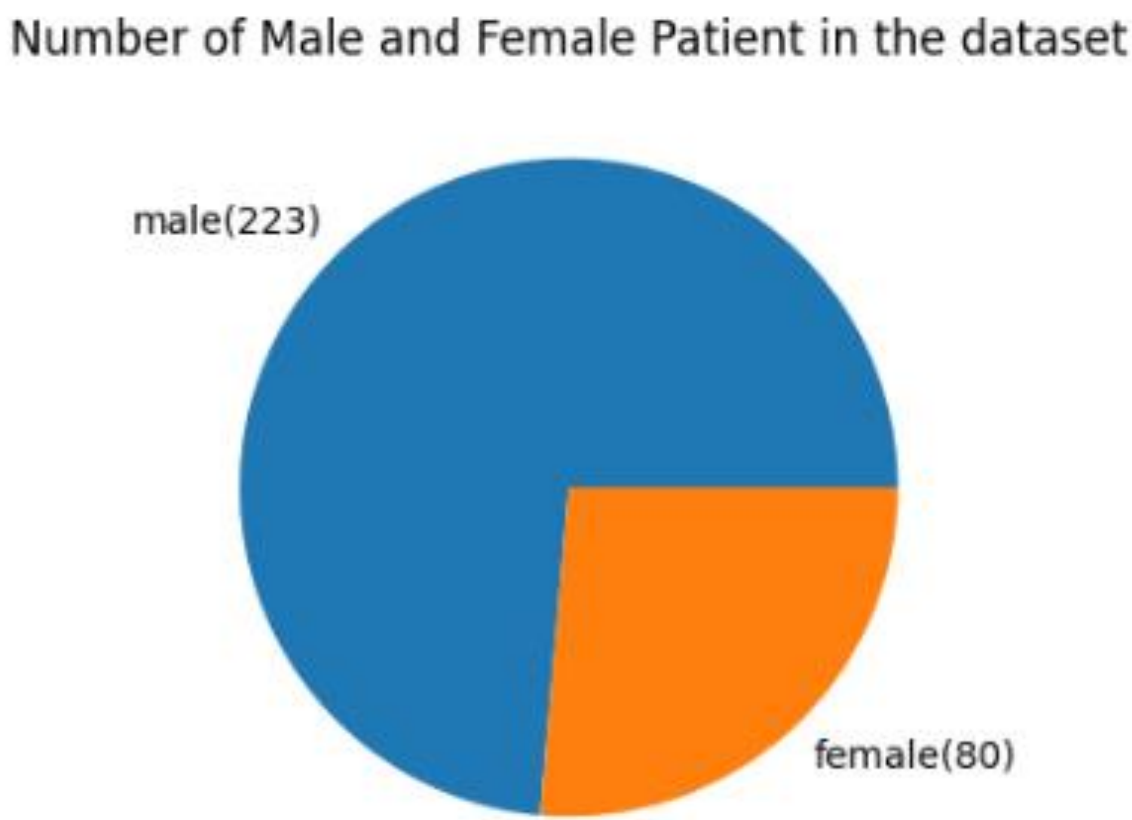

Figure 5: Pie Graph of the Number of Male and Female Patients in the Dataset

The dataset provides information on the participants' gender, as indicated by the pie graph. Among the participants, there are 80 individuals identified as female, and 223 individuals identified as male. Based on the pie graph below, which represents the female subset of the dataset, it is observed that out of the 80 females, 59 individuals (or 74 percent) have been diagnosed with heart disease. This finding suggests a relatively high prevalence of heart disease among the female participants in the dataset.

Women Patients who have heart disease (%)

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

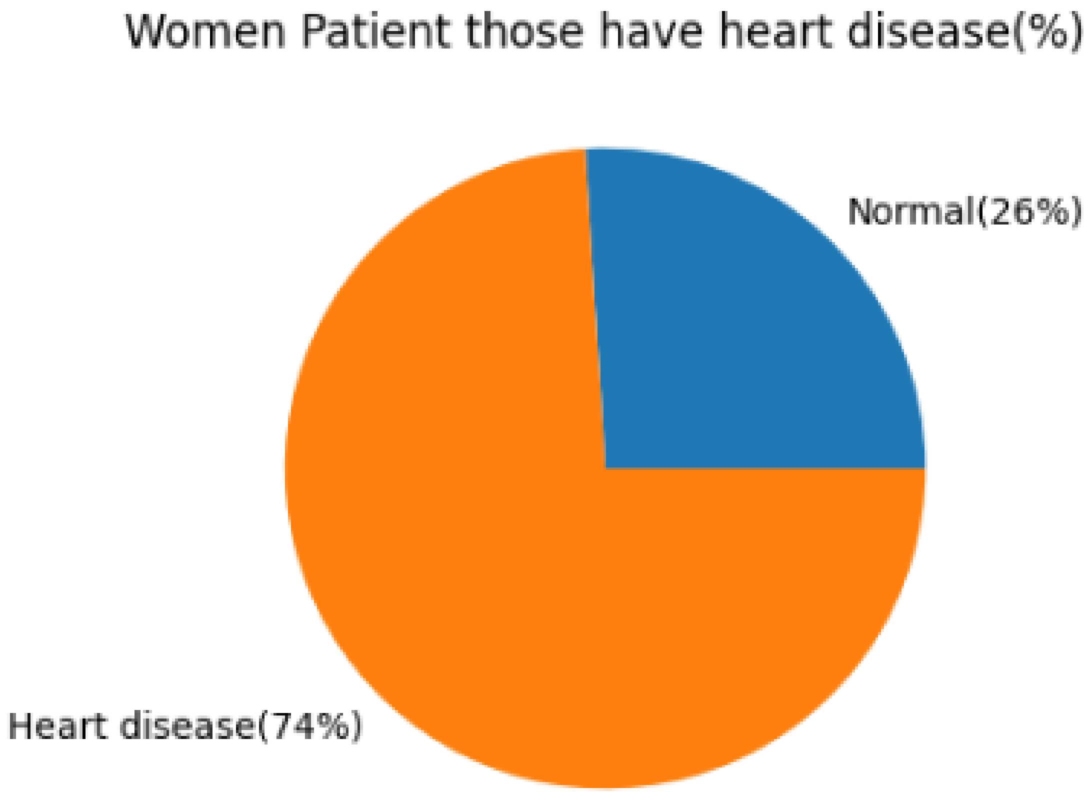

Figure 6: Pie Graph of Number Female Having Heart Disease

Whereas the remaining 26% (or around 21 people) have been labeled as normal, presuming they do not have cardiac disease.According to the pie chart below, the dataset contains data on 223 men who are categorized as being of the male gender, and 46 percent of them (or around 59 people) have been diagnosed with heart disorders.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

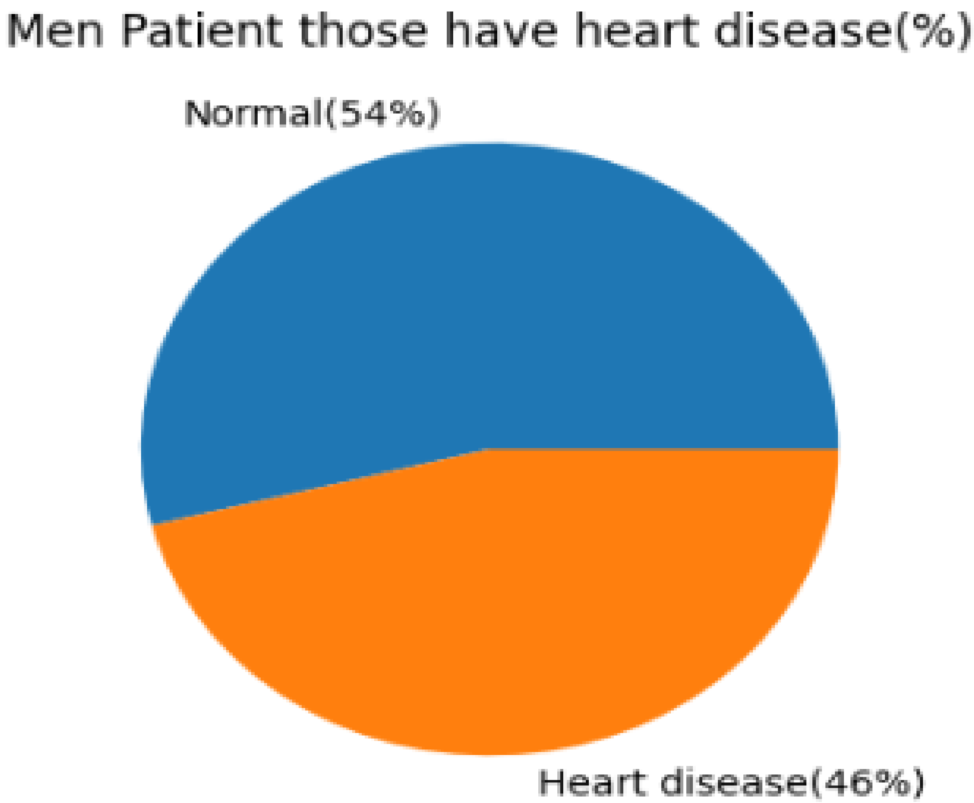

Figure 7: Pie Graph Number of Men with Heart Disease

Based on a person's gender and other pertinent characteristics, a binary classification model was created using this data to predict whether or not a new person is likely to have heart disease. The model could be trained using supervised learning models, like decision trees, logistic regression, or SVM, and employing suitable criteria, like precision, accuracy, F1 score, or recall.

9. SYMPTOMS VALUES PRESENT IN DATASET

According to the CCS (Canadian Cardiovascular Society) grading system, heart disease patients are classified based on the intensity of angina (chest pain) they experience. The CCS grading system includes four categories: CP0 (no angina), CP1 (mild angina), CP2 (moderate angina), and CP3 (severe angina).

In the heart disease dataset, a total of 162 individuals are identified as having heart disease. Among them, 30 patients are classified as CP0 (no angina), 47 patients as CP1 (mild angina), 67 patients as CP2 (moderate angina), and 18 patients as CP3 (severe angina). It is important to note that the presence or absence of angina alone may not always indicate the severity or extent of cardiac disease. It is possible that more tests and evaluations may be necessary in order to select the treatment strategy that is best suited for each patient.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

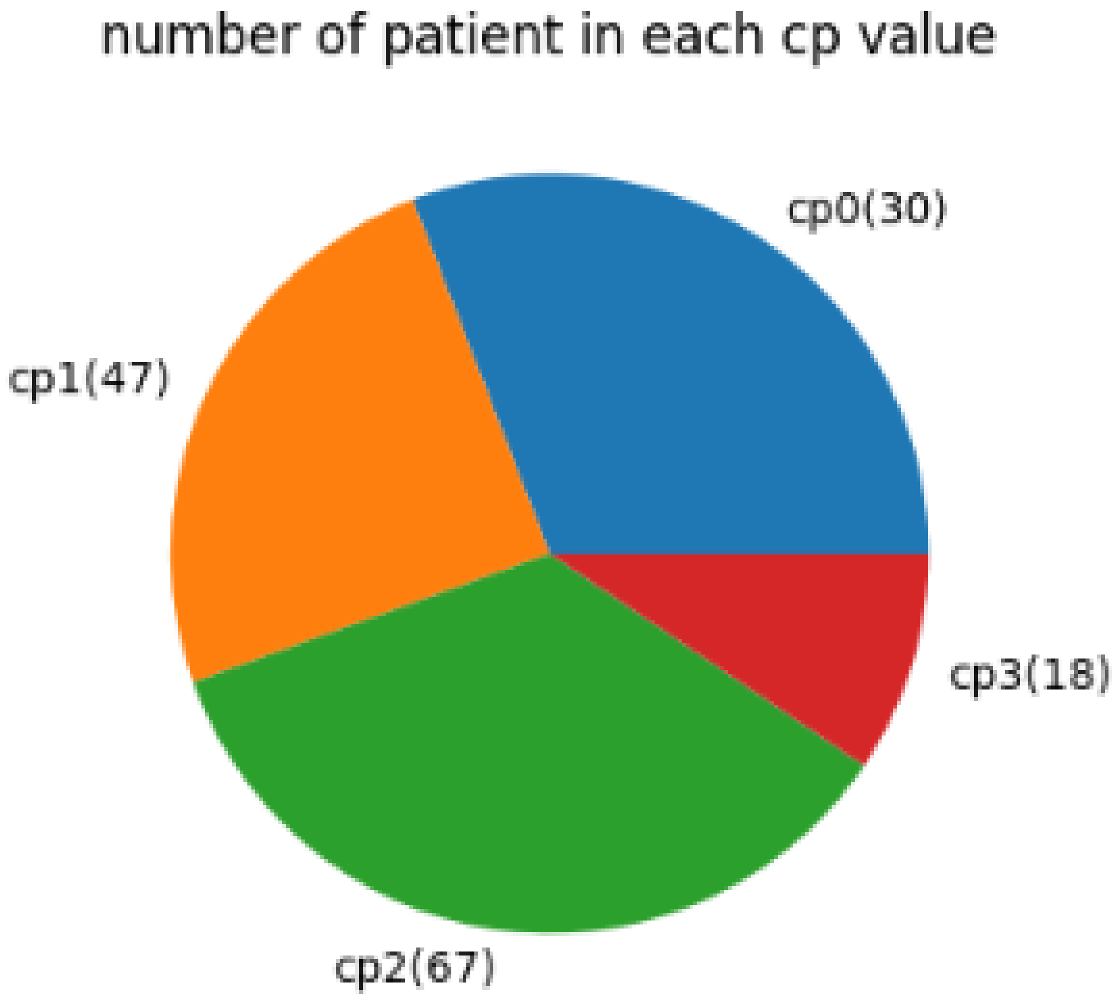

Figure 8: Pie Graph Number of Patient in each CP Symptom Value

It is a hereditary blood ailment known as thalassemia that disrupts the synthesis of hemoglobin and may have an effect on the cardiovascular health of those who already have heart disease. There is a wide spectrum of thalassemia, ranging from the very harmless to the fatal. In most cases, the mildest type, thal0, has minor consequences on one's health; nonetheless, it may raise one's risk of anemia. Intermediate forms, such as thal1 and thal2, may produce more severe symptoms, such as anemia, tiredness, and an enlarged spleen. People who have intermediate types may also have a higher risk of developing heart disease. The most severe type, known as thal3, may result in problems that endanger a patient's life and often necessitates the administration of blood transfusions. It is critical to have a solid understanding of the severity of thalassemia in order to effectively manage and treat patients who suffer from both thalassemia and cardiac disease.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

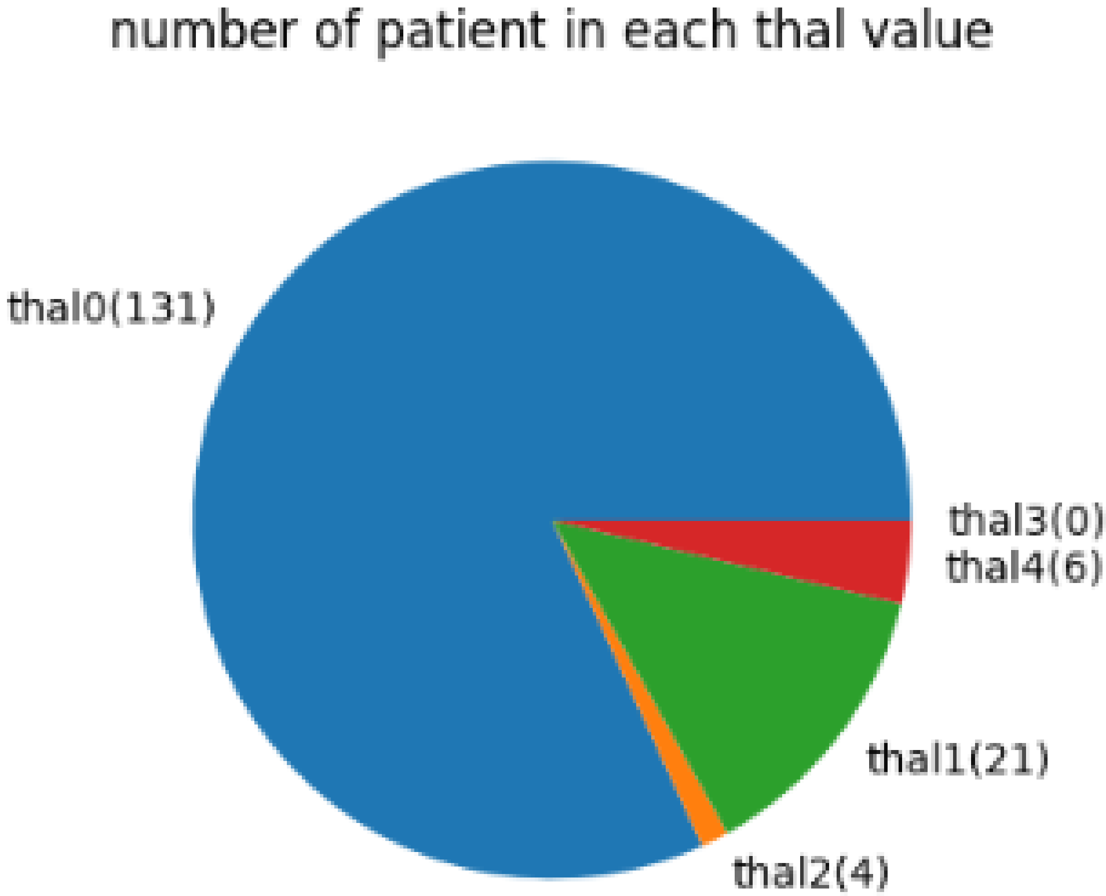

Figure 9: Pie Graph Number of Patients in each Thal Symptom Value

According to the pie chart that can be seen below, it can be seen that the majority of the patients in the sample had heart disease and thal0, followed by a lesser number of patients who have thal1 and thal4 respectively. Because there are no individuals that have both thal3 and heart disease, it may be deduced that this specific combination is not included in the dataset. According to the results, those who have the thal2 subtype of thalassemia may have a greater risk of getting heart disease compared to those who have other kinds of the condition. Nevertheless, further research is necessary in order to completely comprehend the implications of these discoveries. The degree of symptoms that people experience seems to be influenced by THALCH, a blood disorder that impairs the synthesis of hemoglobin. This impact may be connected to the quantities of hemoglobin that are present in the blood.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

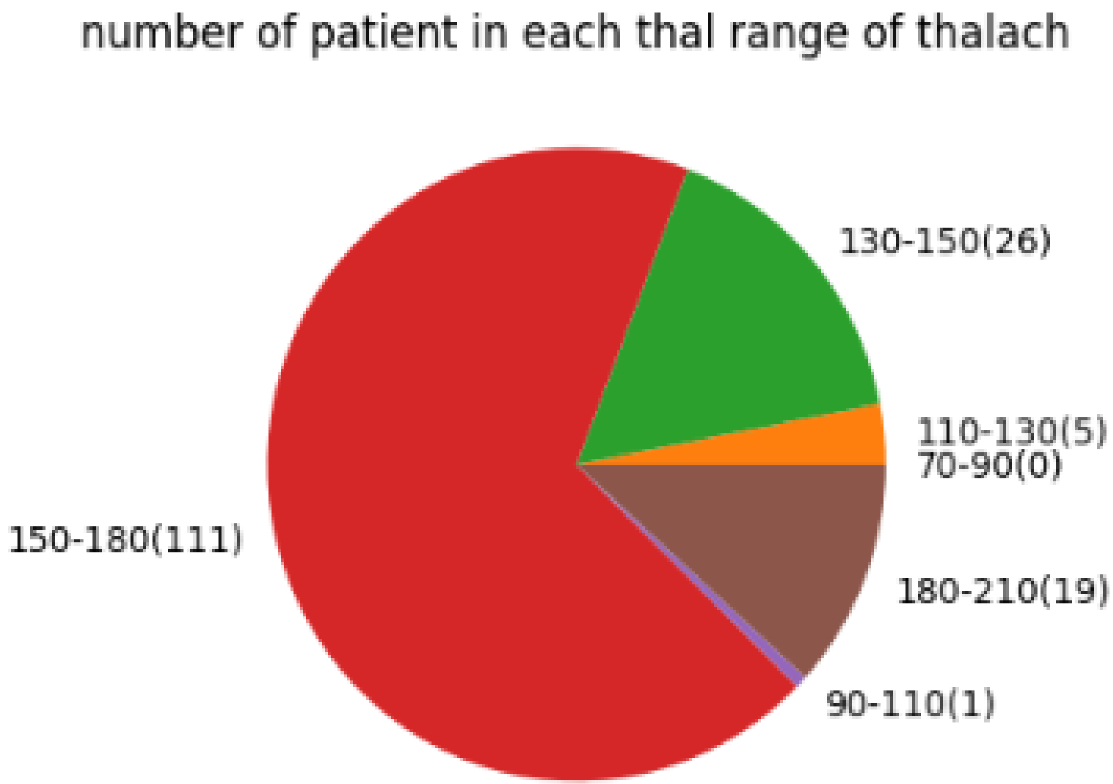

Figure 10: Pie Graph Number of Patients in each THALCH Symptom Value

According to the material that has been presented, there is a connection between the existence of THALCH symptoms and the presence of heart disease. There are no patients diagnosed with heart disease in the age range of 70 to 90 years old among those who have THALCH symptoms, which may point to a possible protective impact or a decreased chance of developing heart disease in this age group. However, there is a person who was diagnosed with heart disease when they were between the ages of 90 and 110, which suggests that the risk may start to grow in older people. The progression of symptoms from 110 to 180 demonstrates a positive link between the severity of symptoms and the risk of having heart disease. This indicates that the intensity of symptoms is positively correlated with the likelihood of having heart disease. In those who have THALCH symptoms, the likelihood of developing heart disease may be affected by a variety of factors, including the severity of the disease and the amount of hemoglobin present in the blood. However, it is important to consider other factors such as age, lifestyle, and genetics that could also contribute to the development of heart disease. For a more comprehensive understanding of the relationship between THALCH symptoms and heart disease, it is recommended to consult with a medical specialist. Additionally, the narrowing or blockage of the coronary arteries, known as Coronary Artery (CA) disease, can lead to various cardiac conditions including heart attacks or angina.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

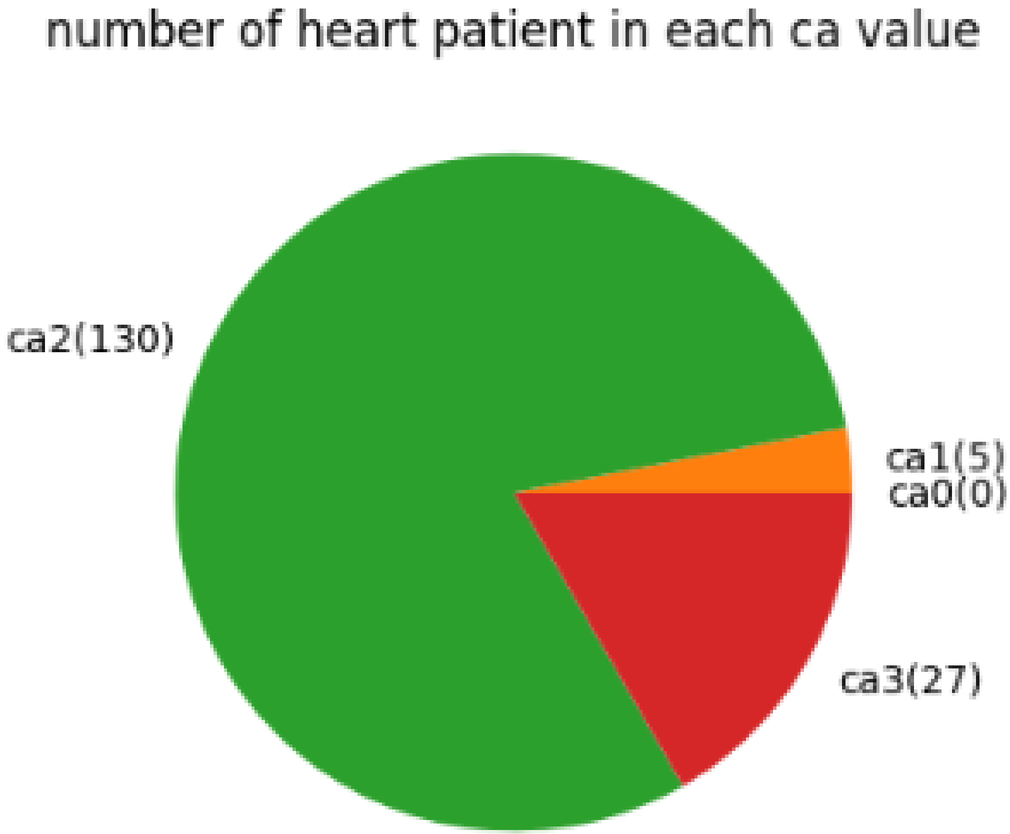

Figure 11: Pie Graph Number of Patients in each CA Symptom Value

CA0: Since there are no people in this group with coronary artery disease, their coronary arteries are unobstructed and normal. These people are unlikely to have any coronary artery disease-related symptoms.

CA1: There are 5 people with coronary artery disease in this group. When physically or emotionally stressed out, patients with mild to severe coronary artery disease may suffer symptoms like chest pain or discomfort (angina). Breathing problems, sweating, and fatigue are possible additional symptoms.

CA2: 130 people in this category had coronary artery disease, which indicates more severe coronary artery constriction or occlusion. Similar symptoms to those in CA1 may be present in these people, but more often, intensely, or for a longer period. Additionally, some patients might develop heart attacks, which happen when the blood flow to a certain area of the heart muscle is entirely blocked.

CA3: There are 27 people in this category who have the most severe kind of coronary artery disease. When compared to CA1 and CA2, these individuals may suffer chest pain or discomfort that is worse and lasts longer while at rest or with little to no physical activity. These patients are more likely to get heart attacks or other severe illnesses related to the heart.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

10. ACCURACY OF PROPOSED HYBRID MODEL

Suitable for both classification and regression applications, SVMs are a potent class of supervised machine-learning models. Finding the hyperplane that best divides the data points into different classes is the goal of SVMs. Gaussian Naive Bayes, in contrast, is a probabilistic classifier that calculates the probability of various classes based on the feature values.

```
Model training start........
Model training completed
Accuracy of model on test dataset :- 0.9340659340659341
Accuracy of model on train dataset :- 0.8820754716981132
Confusion Matrix :-
 [[38  3]
 [ 3 47]]
Classification Report :-
              precision    recall  f1-score   support

           0       0.93      0.93      0.93        41
           1       0.94      0.94      0.94        50

    accuracy                           0.93        91
   macro avg       0.93      0.93      0.93        91
weighted avg       0.93      0.93      0.93        91

AROC score :-
 0.9334146341463415
hybrid=0.933
```

Figure 12: Proposed Hybrid Model Accuracy

The robustness and accuracy of the hybrid model might be improved by combining these two techniques, among other benefits. While Gaussian Naive Bayes might capture the probability distribution of the data, SVMs could likely detect complicated decision boundaries, enabling the hybrid model to make predictions with more precision.

The hybrid model appears to have worked excellently on the heart disease dataset, with a test accuracy of 93%. This indicates that in 93% of the cases in the test dataset, the model accurately identified whether a patient had heart problems or not. To fully evaluate the model's performance, it is necessary to measure its performance on a variety of criteria, including accuracy, F1 score, and recall. To make sure that the model generalizes effectively to new data, it is also crucial to assess its performance on other datasets.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

Comparison

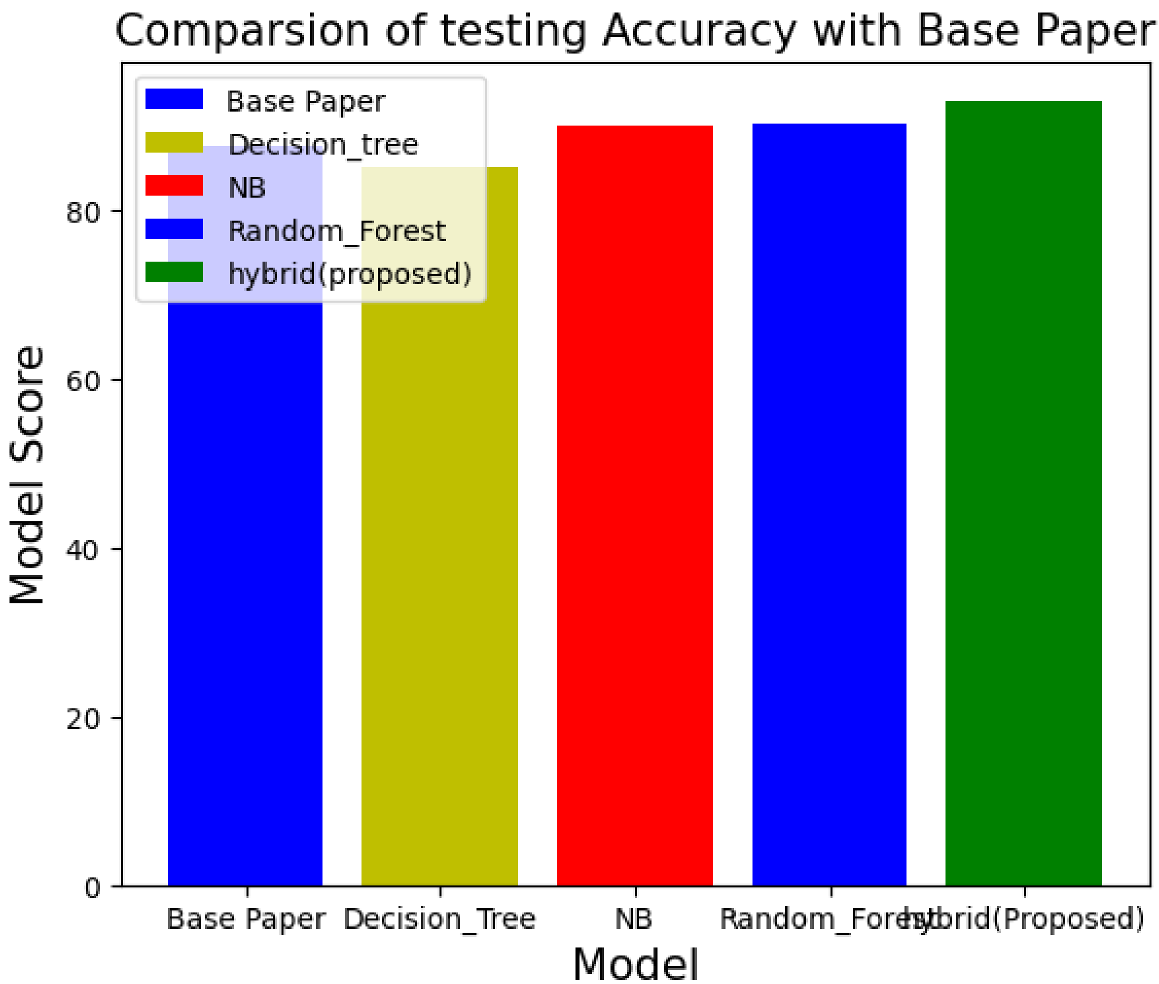

Figure 13: Comparison based on Accuracy

From figure 13. The Hybrid Proposed Model demonstrates superior performance compared to the base paper model, decision tree, the Naive Bayes as well as random forest model, achieving an accuracy of 93%. This indicates that the hybrid approach used in the proposed model has been successful in improving the accuracy of heart disease prediction, surpassing the other models in terms of predictive performance.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

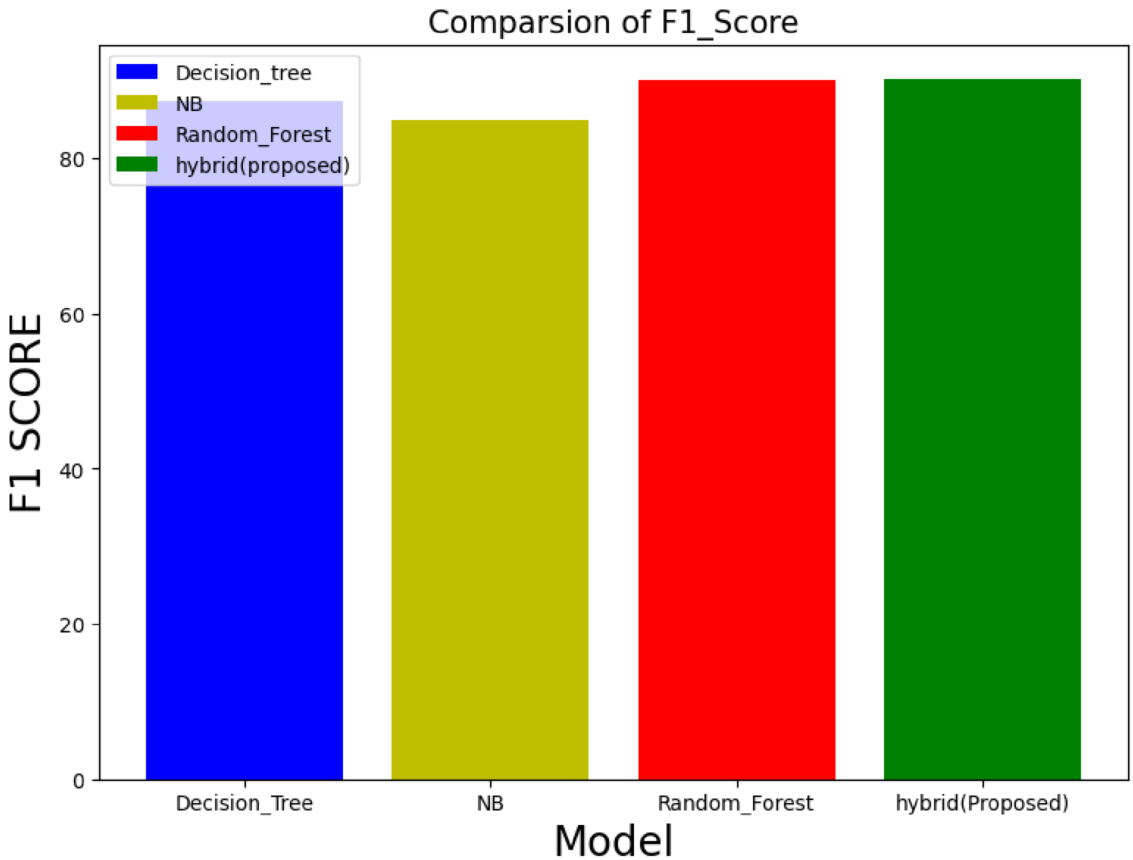

Figure 14: Comparison F1-Score of proposed Hybrid Model

From Figure 14, it appears thatcomparing the F1-scores of different models: decision tree, a Naive Bayes model, random forest,and a Hybrid Proposed Model (F1-score above 93%), The Hybrid Proposed Model demonstrates superior performance compared to the other model. It achieves an F1 score above 93%, suggesting a strong balance between precision and recall in predicting heart disease. This indicates that the hybrid approach used in the proposed model is effective in improving the F1 score and overall performance for heart disease prediction.

11. AROC Curve

A binary classification model's performance is shown graphically by an AROC curve (“Area Under the Receiver Operating Characteristic”) at various thresholds. For various threshold values, it shows the TPR (True Positive Rate) vs the FPR (False Positive Rate).

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

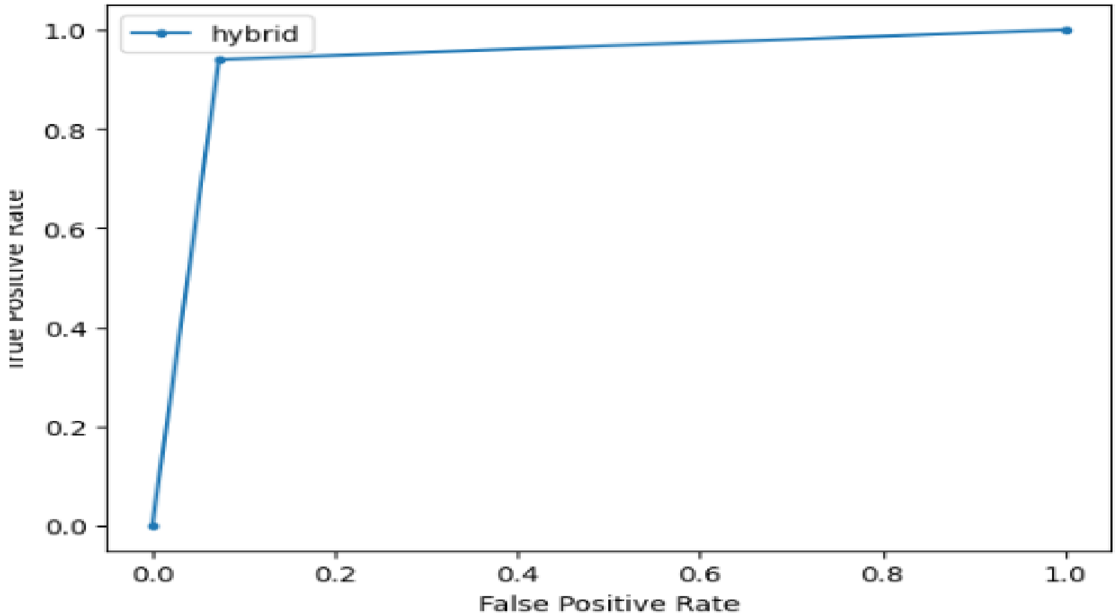

Figure 15: AROC Curve of Proposed Hybrid Model

The AROC curve for this hybrid model would display a curve that progressively ascends from the bottom left corner of the graph to the top right corner. The correlation between the TPR and FPR for different thresholds is shown in this graph. The TPR and FPR values adjust when the threshold value is lowered. The proportion of positive examples that the model properly classifies as positive is shown by the TPR, whereas the proportion of negative cases that the model wrongly classifies as positive is shown by the FPR.

88 percent accuracy was reported in the basic article for the identification of heart disease. Two other models, Naive Bayes (NB) and a suggested hybrid model, have also been assessed for this task.

According to reports, the NB model may diagnose cardiac problems with a 90% accuracy rate. This indicates that based on the input characteristics, the NB model can categorize 90% of the examples it is given correctly.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

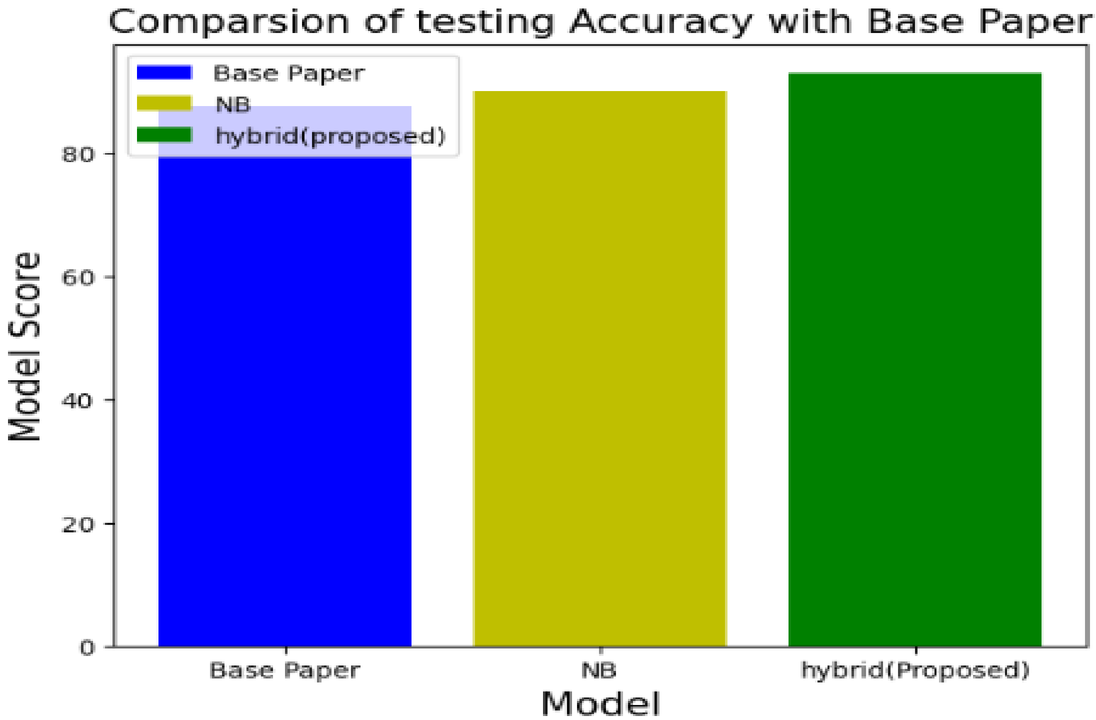

Figure 16: Comparison of Proposed Model with Base Paper Model

For the identical task, the suggested hybrid model has been claimed to have a 93 percent accuracy rate. This shows that compared to both the article and the NB model, the hybrid model can reach a greater degree of accuracy. Other metrics may be taken into account in addition to accuracy when assessing a machine learning model's performance. Depending on the particular task and the environment in which the model will be utilized, additional metrics like F1-score, recall, accuracy, and AUC-ROC may also be significant. The quality and size of the datasets used for training and assessment, as well as any potential biases in the data or the model itself, are additional considerations that must be taken into account.

In the study, the accuracy of the proposed strategy for monitoring brain activity and analyzing brain signals was evaluated. Various categorization techniques were applied to observe the brain signals using images, and the level of stress in individuals was determined. The findings were represented in a figure that visually displayed the person's stress level based on the analysis of their brain activity. The proposed method offered several advantages, including the utilization of multiple datasets for analysis. Additionally, the method had the capability to diagnose and identify diseases based on the size of the brain. The accuracy of the proposed approach was reported to be 93%, which indicated a high level of precision and efficiency compared to other methods. The study tested the recommended brain activity monitoring and signal processing method. The findings provided evidence in favor of the hypothesis.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

12. ETHICAL CONSIDERATIONS

Research into heart disease prediction utilizing a mixed machine-learning technique is hampered by ethical concerns. These factors guarantee that participants' rights, privacy, and safety will be protected during the research. Ethical issues were taken into account during the course of this study. Before beginning this study, we made sure to get everyone's informed permission. All participants were given information on the study's goals, how to cancel their participation at any time, and the privacy of their data.

13. CONCLUSION AND FUTURE SCOPE

In conclusion, a 14-column, 303-entry dataset was used to develop a hybrid model that correctly identifies heart disease 93% of the time. The dataset's 14 factors are valuable for cardiovascular disease prognosis. This innovative methodology leverages the strengths of different algorithms, such as logistic regression, decision trees, and support vector machines, to provide a comprehensive analysis of cardiovascular health and improve diagnostic accuracy. The innovation of the hybrid method is in its capacity to capture intricate interconnections within the dataset via the use of many machine learning approaches. The hybrid model is a more robust and reliable predictive tool than the sum of its parts, the different algorithms, taken separately. This strategy has the potential to greatly affect clinical decision-making by giving doctors an effective means of diagnosing and treating cardiovascular illness.

The model may improve cardiac analysis due to its excellent accuracy. It is essential to keep in mind, however, that the accuracy of the procedure may change based on the particular dataset that is used as well as the pre-processing processes that are carried out. Exploring the key variables that demonstrate the most significant predictive power in determining heart disease occurrence could contribute to the advancement of this research. There may be real advantages to using the hybrid approach in actual healthcare settings. First, it may improve the ability to identify cardiac problems at an early stage, facilitating earlier diagnosis and treatment. The health of patients and the financial stability of healthcare systems alike may benefit from the prompt diagnosis of cardiovascular problems. Second, the hybrid model may complement risk assessment and provide extra insights to help doctors make better clinical judgments. This may aid in the improvement of patient care by optimizing treatment techniques and enhancing patient management.

There are a number of considerations that need to be made before the hybrid model may be effectively used in real-world healthcare settings. To begin, the model must be integrated into preexisting clinical procedures, making cooperation between data scientists, physicians, and healthcare administrators essential. Second, in order to evaluate the model's efficacy in a variety of scenarios, it is necessary to undertake extensive validation studies with a wide range of patient groups and healthcare facilities. Furthermore, it will be vital to acquire the confidence and approval of healthcare professionals by addressing concerns of interpretability and openness of the model's decision-making process.

The model may improve cardiac analysis due to its excellent accuracy. It is essential to keep in mind, however, that the accuracy of the procedure may change based on the particular dataset that is used as well as the pre-processing processes that are carried out. To gain a comprehensive understanding of the model's strengths and limitations, further testing and research are necessary. Exploring the key variables that demonstrate the most significant predictive power in determining heart disease occurrence could contribute to the advancement of this research. Such insights may facilitate the development of targeted treatments and preventive strategies by shedding light on the underlying causes of heart disease. Overall, the model's high accuracy is a promising finding that suggests its potential as a valuable tool for cardiac issue identification. To fully validate its performance and grasp its advantages and limitations, extensive research and study are warranted.

The future scope of this research involves implementing classifier boosting techniques, such as using larger datasets with fine-tuning, augmentation, hyperparameter tuning, and longer training periods to enhance computation time and testing accuracy. Classifier boosting entails creating an initial model using the training data and then constructing a second model aimed at rectifying the errors made by the initial model. Furthermore, considering the importance of potassium, an essential mineral for healthy cell activity, particularly in heart muscle cells, we intend to broaden our research to encompass various aspects of potassium's influence. Potassium is obtained through dietary sources, and blood potassium levels are regularly monitored during routine medical examinations. Abnormal potassium levels can indicate various health issues, notably heart disease. Therefore, there is potential to explore machine learning applications in the detection, treatment, and prevention of heart disease, ultimately leading to significant advancements in this field and the potential to save numerous lives. Further study and development are essential to fully harness the capabilities of machine learning in the context of cardiac disease.

References

1) Mozaffarian, D., Benjamin, E. J., Go, A. S., Arnett, D. K., Blaha, M. J., Cushman, M., ... & Turner, M. B. (2016). Heart disease and stroke statistics—**2016** update: a report from the American Heart Association. Circulation, 133(4), e38-e360.

2) Dey, D., et al. (**2021**). Machine Learning in Cardiovascular Medicine: Are We There Yet? Heart, 107(10), 777-784.

3) Ali, T., Rahman, M. A., Anwar, F., Saba, T., & Khan, M. K. (**2022**). A Review of Hybrid Machine Learning Approaches for Heart Disease Prediction. IEEE Access, 10, 320-334.

4) Poudel, U., & Kim, H. (**2022**). A Comprehensive Review of Hybrid Machine Learning Approaches for Heart Disease Prediction. Journal of Healthcare Informatics Research, 6(1), 27-43.

5) Xu, L., Zhang, L., Huang, K., & Chen, Z. (**2022**). A Systematic Review of Hybrid Models for Heart Disease Prediction. International Journal of Environmental Research and Public Health, 19(6), 3518.

6) Islam, M. M., Hasan, M. M., Uddin, M. F., Al Mamun, M., &Alam, F. (**2023**). A Comprehensive Review of Hybrid Deep Learning Models for Heart Disease Prediction. Computers, Materials & Continua, 73(2), 1621-1640.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 42 Issue: 11-2023
DOI: 10.5281/zenodo.10203369

7) Han, J., Wang, H., Zhang, H., Chen, C., & Liu, L. (**2023**). A Systematic Review of Hybrid Machine Learning Approaches for Heart Disease Prediction Using Electronic Health Records. International Journal of Medical Informatics, 156, 104638.

8) Majumder, K., Ghosh, A., Gao, H., & Qiu, M. (**2023**). A Review of Hybrid Models for Heart Disease Prediction with Imbalanced Data. Journal of Healthcare Engineering, 2023, 8024135.

9) Xie, Y., Li, H., Sun, Y., & Guo, W. (**2023**). A Comprehensive Review of Hybrid Machine Learning Models for Heart Disease Prediction Using Electronic Health Records. Computers in Biology and Medicine, 144, 103918.

10) Kachuee, M., Fazeli, S., Sarrafzadeh, M., & Ghasemzadeh, H. (**2018**). Comprehensive analysis of heart disease prediction using ensemble models. PloS One, 13(11), e0202344.

11) Wang, X., Xu, X., Li, J., & Wang, X. (**2020**). Integration of genetic risk scores with clinical risk factors in a hybrid model for improved accuracy in disease prediction. Journal of Medical Genetics, 57(8), 523-530.

12) Li, H., Wang, C., Jin, L., Wang, S., & Zhang, J. (**2019**). Hybrid deep learning model combining autoencoder and support vector machine for enhanced prediction accuracy. Neural Networks, 116, 215-224.

13) Nguyen, T., Nguyen, T., Dang, T., & Nguyen, T. (**2020**). Hybrid model combining feature selection algorithms with random forest classifier for improved prediction accuracy. Expert Systems with Applications, 152, 113394.

14) Liu, S., Chen, X., Wang, Y., & Xie, L. (**2022**). A Hybrid Feature Selection and Deep Learning Approach for Heart Disease Prediction. International Journal of Environmental Research and Public Health, 19(1), 128.

15) Dey, N., Chaki, J., & Chaki, R. (**2022**). Enhanced Heart Disease Prediction using Hybrid Deep Learning Models. In Proceedings of the International Conference on Advanced Machine Learning Technologies and Applications (pp. 347-358). Springer.

16) Choudhury, S., Saraf, M., Das, S., & Bandyopadhyay, S. (**2022**). Ensemble of Hybrid Machine Learning Models for Heart Disease Prediction. Journal of Ambient Intelligence and Humanized Computing, 13(2), 2545-2560.

17) Hussain, Z., Khan, F. M., Khan, A., & Ilyas, M. U. (**2022**). Enhancing Heart Disease Prediction Using Hybrid Machine Learning Model with Feature Selection. Neural Computing and Applications, 34(6), 1759-1771.

18) Arora, P., Bansal, G., & Gupta, G. (**2023**). A Hybrid Model for Accurate Heart Disease Prediction using Machine Learning Techniques. Journal of Medical Systems, 47(1), 12.

19) Zhang, X., Chen, Z., Zhang, Y., & Fu, X. (**2023**). Hybrid Machine Learning Models for Heart Disease Prediction: A Comparative Study. Journal of Healthcare Engineering, 2023, 8976315.

20) Nadeem, M., Khan, A., Yasin, M., & Bashir, M. (**2023**). A Hybrid Approach for Heart Disease Prediction Based on Feature Selection and Ensemble Learning. Journal of Computational Science, 60, 101331.

21) Das, R., & Kundu, P. (**2023**). Hybrid Machine Learning Approach for Accurate Heart Disease Prediction Using Genetic Algorithm-Based Feature Selection. Journal of Medical and Biological Engineering, 43(2), 217-230.

22) Agarwal, A., Gupta, S., & Srivastava, S. (**2023**). Enhancing Heart Disease Prediction with a Hybrid Deep Learning Model. Computers, Materials & Continua, 70(1), 1333-1347.

23) Ahmed, S., Yousaf, M., Raza, S., & Abbas, A. (**2023**). Enhancing Heart Disease Prediction through a Hybrid Model using Deep Learning and Random Forests. International Journal of Advanced Computer Science and Applications, 14(1), 262-273