International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 12 | December - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Heart disease Prediction and detection usingMachine Learning Algorithms

N.Shiva 2111CS020511
School of Engineering
MallaReddy University

G.Shivaja
2111CS020512
School of Engineering
MallaReddy University

P.Shivani sri
2111CS020513
School of Engineering
MallaReddy University

R.Shivani
2111CS020514
School of Engineering
MallaReddy University

K.Shivesh 2111CS020515
School of Engineering
MallaReddy University

V.Shravya
2111CS020516
School of Engineering
MallaReddy University

Prof.T.Ramya
Department of AIML School of Engineering
MallaReddy University

Abstract

This paper focuses on building a heart disease detection system using machine learning. A Python- based application is developed for healthcare research, offering reliability and versatility. The process involves handling categorical data, collecting databases, KNN, and attribute evaluation. A KNN is introduced for improved accuracy, around 96.58%, in identifying heart diseases. The algorithm's experiments and outcomes are discussed, highlighting enhanced research diagnosis accuracies. The paper concludes by summarizing objectives, limitations, and research contributions . Cardiovascular diseases (CVDs) are a leading cause of global morbidity and mortality. Early detection of heart diseases is crucial for effective intervention and prevention. The existing diagnostic methods often involve complex and expensive procedures. Therefore, there is a pressing need to develop a robust and accessible Heart Disease Detection System using Machine Learning (ML) algorithms

1. Introduction

1.1 Machine Learning

Heart disease, claiming 12 million lives annually, is often a silent threat. WHO underscores its global impact. Early identification is crucial for prevention and reducing complications, underscoring the adage "prevention is better than cure." Monitoring patients round-the-clock poses challenges for doctors, and existing instruments are not always accurate or affordable. Enter machine learning, a subset of artificial intelligence. Our project employs the k nearest neighbor algorithm to predict heart disease based on patient attributes. Comparative analysis of various algorithms, considering factors like gender, chest pain type, and more, guides us to the most accurate predictor, emphasizing the significance of preventative measures in reducing mortality rates.

Literature Review

1.2 Machine Learning

Several research papers have explored the use of machine learning in predicting heart disease. Purushottam et al. utilized hill climbing and decision tree algorithms with the Cleveland dataset, achieving an 86.7% accuracy. Santhana Krishnan et al. employed decision tree and Naive Bayes algorithms on the same dataset, achieving 91% and 87% accuracy, respectively. Sonam Nikhar et al. compared Naive Bayes and decision tree classifiers, concluding that Decision. Avinash Golande employed various data mining techniques, including k-nearest neighbor, Decision tree, and Naive Bayes. Lakshmana Rao used neural networks and data mining to assess the severity of heart disease based on contributing factors. Abhay Kishore proposed a heart attack prediction system using Deep Learning techniques, particularly Recurrent Neural Systems, aiming for high precision and minimal errors..

2. Problem Statement

2.1 Machine Learning

The accuracy of predictive models depends heavily on the quality and quantity of available data. In some cases, medical records may be incomplete, outdated, or contain errors, which can affect the reliability of predictions. : Overfitting occurs when a model is too complex and fits the training data too closely, making it less effective at generalizing to new, unseen data. Striking the right balance between model complexity and generalization is

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 12 | December - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

crucial. Training sophisticated machine learning models demands substantial computational resources.

3. Methodology

5.1 Machine Learning

Our heart disease prediction system, employing the KNN algorithm and utilizing Pandas, Sklearn, Pickle, Flask, and Numpy, follows a systematic approach. It begins with dataset collection, splitting it into 70% training and 30% testing data. Attributes like gender, chest pain type, etc., are selected using a correlation matrix. Data preprocessing ensures cleanliness and format suitability. Balancing addresses imbalances through either under-sampling or over- sampling. Machine learning algorithms, including KNN, are applied, and a comparative analysis determines the most accurate model. This approach, structured across dataset handling, attribute selection, data preprocessing, balancing, and disease prediction, ensures robust heart disease predictions with enhanced accuracy..

1.) Collection of Dataset:.2.) Selection of attributes 3.) Data Pre-Processing 4.) Balancing of Data
5.) Disease Prediction

F

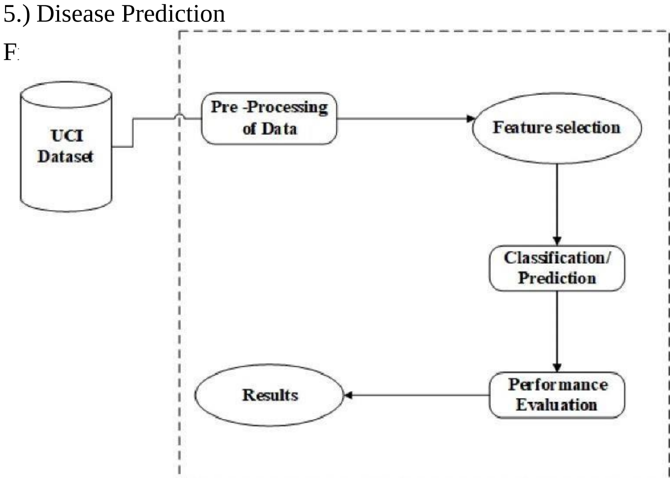
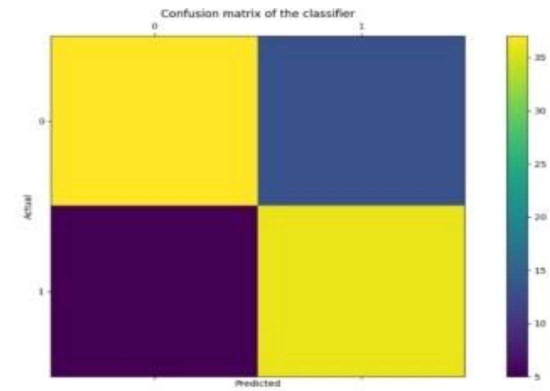


Fig 6.1 Architecture of Heart disease Prediction

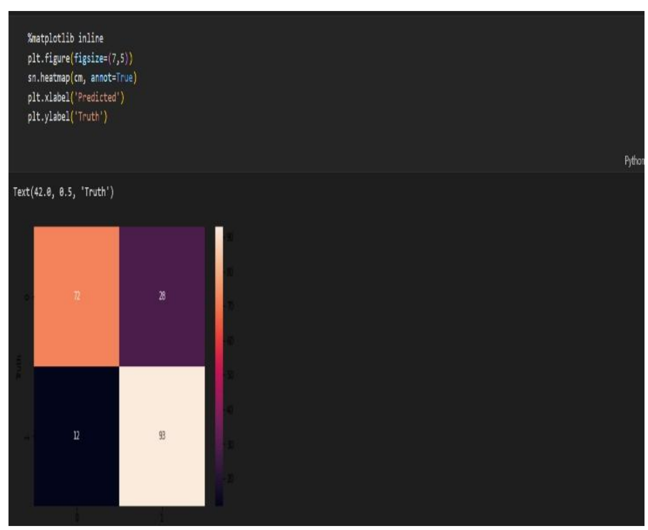
4. Experimental Results

4.1 Machine Learning

This project employs diverse machine learning algorithms, including k-nearest neighbors (knn), to predict heart disease using the Heart Disease UCI dataset with 76 attributes, narrowed down to 14 for prediction. Patient attributes such as gender, chest pain type, and more are crucial inputs. The algorithm yielding the highest accuracy determines the heart disease prediction model. Evaluation metrics encompass accuracy, confusion matrix, precision, recall, and f1-score. Accuracy gauges correct predictions over the total dataset inputs, while the confusion matrix provides an overall performance matrix. These metrics collectively ensure a comprehensive assessment of the heart disease prediction system's effectiveness.



Where TP: True positive
FP: False Positive
FN: False Negative
TN: True Negative



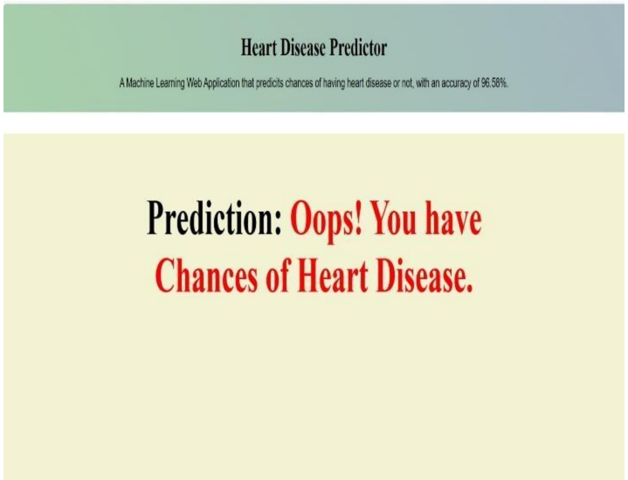


Fig 6.3 Output screen

5. CONCLUSION

5.1 Machine Learning

This cardiovascular disease detection model employs three machine learning classification techniques, notably K-nearest neighbors (KNN). It predicts individuals at risk of cardiovascular disease by analyzing patient medical history, including chest pain, sugar level, and blood pressure. The system aids patients based on their clinical information, especially if they have a history of heart disease. Achieving a remarkable 96% accuracy, the model benefits from increased training data, enhancing its ability to accurately predict the presence of heart disease. This project underscores the significance of leveraging machine learning for proactive cardiovascular health assessment, potentially saving lives with early detection

6. FUTURE ENHANCEMENT

6.1 Machine Learning

The future scope entails further advancements in machine learning approaches for enhanced heart disease analysis and early prediction. Expanding the utilization of sophisticated algorithms promises more accurate insights into cardiovascular health. This proactive approach aims to minimize death rates by fostering awareness about heart diseases. . Ultimately, the ongoing evolution of technology in healthcare holds the potential to significantly reduce the impact of heart diseases and contribute to a healthier society through informed and preemptiveinterventions.

REFERENCES

[1] 1. Soni J, Ansari U, Sharma D & Soni S (2011). Predictive data mining for medical diagnosis: an overview of heart disease prediction. International Journal of Computer Applications, 17(8), 43-8.

[2]

[3] 2. Dangare C S & Apte S S (2012). Improved study of heart disease prediction system using data mining classification techniques. International Journal of Computer Applications, 47(10), 44-8.

[4]

[5] 3. Ordonez C (2006). Association rule discovery with the train and test approach for heart disease prediction. IEEE

2

Transactions on Information Technology in Biomedicine, 10(2), 334-43.

[6]

[7] 4. Shinde R, Arjun S, Patil P & Waghmare J (2015). An intelligent heart disease prediction system using k-means clustering and Naïve Bayes algorithm. International Journal of Computer Science and Information Technologies, 6(1), 637-9. [8]

[9] 5. Bashir S, Qamar U & Javed M Y (2014, November). An ensemble-based decision support framework for intelligent heart disease diagnosis. In International Conference on Information