Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 11, No. 2, November : 2023

HEART DISEASE PREDICTION SYSTEM USING MACHINE LEARNING

Prof. Sai Takawale[1] prof. Ruchira Tare[2] Assistant Professor, Dept. of Computer Engineering, Trinity College of Engineering and Research, Pune, India
Sanket Ugale[3], Bhagyashri Kasle[3], Akanksha Gitte[3], Saurabh Shrivastava[3]
[3]BE Student, Dept. of Computer Engineering, Trinity College of Engineering and Research, Pune, India

Abstract
Machine learning techniques are gaining popularity because of the proper prediction and efficient results. Machine learning classification can help predict the heart disease in efficient way. The research leverages a comprehensive dataset of patient attributes, including clinical and demographic factors, to develop an accurate predictive tool for early detection of heart disease. By training and evaluating the SVM model on this dataset, we aim to enhance the accuracy and reliability of heart disease diagnosis. The results show promising outcomes in terms of classification accuracy, sensitivity, and specificity, suggesting the potential for real-world application in healthcare settings. The propose system contributes to improve cardiovascular disease risk assessment and early intervention, ultimately leading to better patient outcomes and reduced healthcare costs.
Keywords: ***Machine Learning, SVM, SHDP, KNN, Naive Bayes.***

I. Introduction

According to WHO report, because of fast lifestyle and food habits, the heart disease rates are increasing day by day, which needs proper analysis and prediction of such diseases so that it can recover timely. Along with manual diagnosis, automation in prediction can yield better result. The machine learning technology has proven its efficiency in every field that too medical filed Heart disease is the leading cause of death worldwide, accounting for an estimated 17.9 million deaths in 2021. Early detection and treatment are essential for reducing the mortality and morbidity associated with heart disease. SVMs have been shown to be effective in predicting heart disease risk using a variety of features, such as age, sex, blood pressure, cholesterol levels, and smoking status. SVMs work by finding a hyperplane in the feature space that separates the data into two classes with the maximum margin. In the context of heart disease prediction, the two classes would be patients with heart disease and patients without heart disease. SVMs have several advantages over other machine learning algorithms for heart disease prediction. First, they are able to handle complex datasets with a large number of features. Second, they are robust to noise and outliers in the data. Third, they are able to learn complex relationships between the features. SVMs have been shown to achieve high accuracy in predicting heart disease risk in a variety of studies. In one study, an SVM model achieved an accuracy of 92.3% in predicting heart disease risk in a dataset of over 10,000 patients. SVM- based heart disease prediction models have the potential to improve the early detection and treatment of heart disease. They can be used to develop screening programs that can identify individuals at high risk of heart disease. SVM models can also be used to help clinicians make decisions about treatment options. Here is an example of how SVM-based heart disease prediction models could be used in practice: A primary care physician could use an SVM model to predict the heart disease risk of their patients. The model would take into account the patient's features, such as age, sex, blood pressure, cholesterol levels, and smoking status. The model would then output a probability score, which would indicate the patient's risk of developing heart disease. If the patient's risk score is high, the physician could order additional tests or refer the patient to a specialist. The physician could also provide the patient with education and counseling on how to reduce their risk of heart disease. SVM-based heart disease prediction models are a promising new tool for improving the early detection and treatment of heart disease. They have the potential to save lives and improve the quality of life for millions of people around the world.

II. Literature

The literature survey delves into the field of machine learning, an evolving technique that involves extracting valuable information from vast data sets for informed business decisions.

Machine learning is one of the application areas to achieve artificial intelligence. In order to detect the diseases in timely manner various machine learning techniques are used. The classification in machine learning is used to predict the diseases. According the author in [1]. The datasets used to train the machine learning model are very large datasets and need to train through various algorithm like SVM, KNN, Naïve Bayes etc.

The author in [2] have developed the classification model to predict the mental health of the patient. Here it shows through results how the classification are useful to enhance the prediction method.

Its application in the medical domain aids in identifying significant patterns and data crucial for clinical diagnoses. In this context, the study focuses on heart disease diagnosis and introduces SHDP (Smart Heart Disease Prediction), a predictive model based on Navies Bayesian [3]. The rapid advancement in mobile health technology, as a web application, has facilitated the standardized collection of necessary data. For the prediction of heart disease, attributes like age, blood pressure, cholesterol level, gender, and blood sugar are obtained from medical profiles. These attributes serve as inputs for the Navies Bayesian classification, which enables accurate heart disease prediction. The dataset is partitioned into 80% for training and 20% for testing purposes. The proposed methodology encompasses stages such as dataset collection, user registration, and login through an application-based system, classification via Navies Bayesian, prediction, and secure data transfer using AES (Advanced Encryption Standard). The study showcases various knowledge abstraction techniques through data mining, specifically tailored for heart disease prediction. The findings highlight the effective role of the diagnostic system in predicting risk factors associated with heart diseases, thereby emphasizing its potential significance in the medical field.

In contemporary society [4], heart disease stands out as a prominent factor contributing to shortened lifespans. Given the reliance of a significant population on the healthcare system, the need for quick and accurate results remains crucial. The continuous generation and collection of a substantial volume of data by healthcare organizations necessitate the utilization of data innovation to extract valuable insights through automated processes. Weighted Association Rule, a type of data mining technique, serves to streamline manual tasks and facilitate direct data extraction from electronic records. This automation not only enhances the efficiency of data extraction but also contributes to cost reduction and potentially life-saving measures. The focus of this paper lies in the derivation of rules that aid in predicting a patient's risk of developing coronary disease. The conducted tests have demonstrated that the majority of these rules significantly contribute to the accurate prediction of coronary illness. This approach holds promising implications for improving the overall efficacy and accessibility of healthcare services, thereby addressing the challenges posed by heart disease in the contemporary healthcare landscape.

Data analysis has emerged as a crucial tool in managing large volumes of healthcare data [5]. While previous medical research has primarily focused on handling and assimilating extensive hospital data, the growing volume of data in the biomedical and healthcare fields emphasizes the importance of accurate data analysis for early disease detection and improved patient care. Challenges arise when dealing with partially missing medical data, which can lead to a decrease in prediction accuracy. To address the issue of missing medical data, the study implements data cleaning and imputation techniques to convert incomplete data into complete datasets. The research focuses on predicting heart disease using the Naïve Bayes and K-Nearest Neighbors (KNN) algorithms based on a specific dataset. To expand the scope of the study, the proposal introduces disease risk prediction using structured data, incorporating a convolutional neural network (CNN) based unimodal disease risk prediction algorithm. The CNN-UDRP algorithm demonstrates a prediction accuracy exceeding 65%, showcasing its potential in predicting disease risks effectively. Moreover, this system serves to provide answers to

queries related to diseases that individuals commonly encounter in their lives, thus contributing to an improved understanding of disease prediction and management in the healthcare domain.

Medical institutions such as hospitals and healthcare centers generate an immense volume of data, yet the effective utilization of this data remains a challenge [6]. The healthcare system is often characterized as "data rich" but "knowledge poor," highlighting the need for efficient analysis methods to uncover meaningful relationships and patterns within healthcare data. Data mining techniques present a viable solution to address this challenge. To this end, various data mining methodologies are being employed. This paper aims to provide an overview of different knowledge abstraction techniques using data mining methods, specifically focusing on their application in contemporary research for the prediction of heart disease. The study evaluates the effectiveness of data mining methods such as Naive Bayes, Neural Networks, and Decision Tree algorithms on medical datasets [7]. By analyzing these data mining techniques on medical datasets, the paper seeks to contribute to the growing body of knowledge in the field of heart disease prediction. The comprehensive review of these methodologies emphasizes the potential of data mining in uncovering valuable insights and patterns within the healthcare domain, thereby improving the overall understanding and management of heart disease.

Heart disease has emerged as the foremost health concern globally, leading to a significant number of fatalities, especially during the initial occurrence of a heart attack. However, it is not just heart attacks that pose health risks, as other conditions such as breast cancer, lung cancer, and ventricular valve issues also contribute to the burden of disease. The need for a robust framework capable of effectively identifying the prevalence of heart disease among numerous samples instantaneously remains crucial. This paper evaluates the potential of nine classification techniques in predicting heart disease, including decision trees, naive Bayesian neural networks, Support Vector Machines (SVM), Artificial Neural Networks (ANN), K- Nearest Neighbors (KNN), Apriori algorithm, and SVM [9][10]. The study utilizes various medical profiles such as age, sex, blood pressure, chest pain type, and fasting blood sugar for prediction. These techniques prove to be effective in accurately predicting the likelihood of patients developing heart disease. The medical community's involvement in the detection and prevention of heart disease is paramount, and the analysis demonstrates that classification-based techniques offer high effectiveness and superior accuracy compared to previous methods. This research emphasizes the significance of implementing advanced classification methodologies in the domain of heart disease prediction, thereby contributing to improved healthcare outcomes and preventive measures.

Proposed System

The machine learning classification method used here to achieve efficiency in predicting heart diseases.

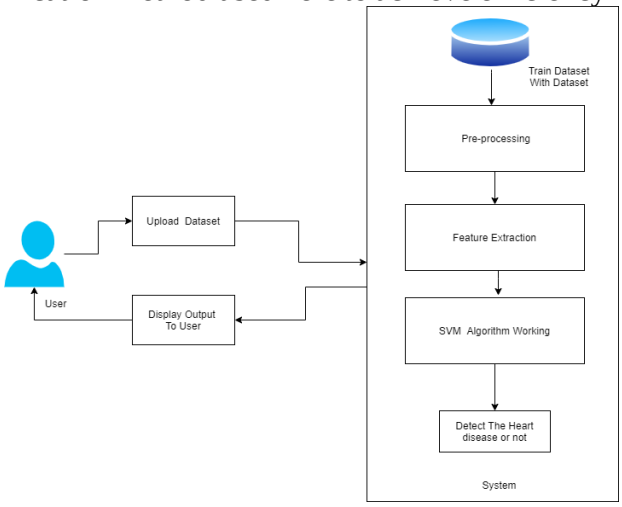

Figure 3.1 Showing Classification Using machine learning for heart disease diagnosis.

Objectives

1. Improve early detection of heart disease.
2. Reduce mortality and morbidity associated with heart disease.
3. Enhance the well-being of individuals suffering from heart disease.
4. Identify risk factors for heart disease.
5. Develop new diagnostic tests for heart disease.
6. Improve accuracy of other machine learning algorithms for heart disease prediction.

Conclusion

Support vector machines (SVMs) are a powerful tool for heart disease prediction. They have been shown to achieve high accuracy in predicting heart disease risk in a variety of studies. SVMs are also robust to noise and outliers in the data, and they are able to learn complex relationships between the features. SVM-based heart disease prediction models have the potential to revolutionize the way that heart disease is diagnosed and treated. They can be used to develop screening programs that can identify individuals at high risk of heart disease, even if they have no symptoms. SVM models can also be used to SVM. The study utilizes various medical profiles such as age, sex, blood pressure, chest pain type, and fasting blood sugar for prediction. These techniques prove to be effective in accurately predicting the likelihood of patients developing heart disease. The medical community's involvement in the detection and prevention of heart disease is paramount, and the analysis demonstrates that classification-based techniques offer high effectiveness and superior accuracy compared to previous methods. This research emphasizes the significance of implementing advanced classification methodologies in the domain of heart disease prediction, thereby contributing to improved healthcare outcomes and preventive measures.

References

[1] Tarwani, Kanchan M., Saleha S. Saudagar, and Harshal D. Misalkar. "Machine learning in big data analytics: an overview." International Journal of Advanced Research in Computer Science and Software Engineering 5.4 (2015): 270-274.

[2] M. A. Abid, Z. Dehghan, T. Shinde and G. Narang, "Machine Learning based approaches for Identification and Prediction of diverse Mental Health Conditions," 2023 IEEE International Conference on Contemporary Computing and Communications (InC4), Bangalore, India, 2023, pp. 1-5, doi: 10.1109/InC457730.2023.10263141.

[3] A. N. Repaka, S. D. Ravikanti and R. G. Franklin, "Design And Implementing Heart Disease Prediction Using Naives Bayesian," 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2019, pp. 292-297, doi: 10.1109/ICOEI.2019.8862604.

[4] A. Chauhan, A. Jain, P. Sharma and V. Deep, "Heart Disease Prediction using Evolutionary Rule Learning," 2018 4th International Conference on Computational Intelligence & Communication Technology (CICT), Ghaziabad, India, 2018, pp. 1-4, doi: 10.1109/CIACT.2018.8480271.

[5] A. Chauhan, A. Jain, P. Sharma and V. Deep, "Heart Disease Prediction using Evolutionary Rule Learning," 2018 4th International Conference on Computational Intelligence & Communication Technology (CICT), Ghaziabad, India, 2018, pp. 1- 4, doi: 10.1109/CIACT.2018.8480271.

[6] S. Ambekar and R. Phalnikar, "Disease Risk Prediction by Using Convolutional Neural Network," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India, 2018, pp. 1- 5, doi: 10.1109/ICCUBEA.2018.8697423.

[7] Purushottam, Kanak Saxena, Richa Sharma, Efficient Heart Disease Prediction System, Procedia Computer Science, Volume 85,2016,Pages 962-969, ISSN 1877-0509.

[8] C. Sowmiya and P. Sumitra, "Analytical study of heart disease diagnosis using classification techniques," 2017 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), Srivilliputtur India, 2017, pp. 1-5, doi:

10.1109/ITCOSP.2017.8303115.
[9]S. Saudagar, M. Kulkarni, I. Raghvani, H. Hirkani,
I. Bassan and P. Hole, "ML-based Java UI for Residence Predictor," *2023 International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)*, Bengaluru, India, 2023, pp. 838-843, doi: 10.1109/IDCIoT56793.2023.10053480.
[10] Saleha Saudagar, Rekha Ranawat, An Amalgamated Novel IDS Model for Misbehaviour Detection using VeReMiNet, Computer Standards &Interfaces, Vo 88, 2024, 103783, ISSN 0920- 5489, https://doi.org/10.1016/j.csi.2023.103783. (https://www.sciencedirect.com/science/article/pii/ S0920548923000648).