WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Diagnosis of Cardiovascular Disease using Machine Learning Techniques

Pooja Shrestha [a], Rashmi R [b]

[a] *Undergraduate, Dept of Information Science, BMS College of Engineering, India, pooja.is20@bmsce.ac.in*
[b] *Assistant Professor, Dept of Information Science, BMS College of Engineering, India, rashmir.ise@bmsce.ac.in*

A B S T R A C T

Cardiovascular diseases (CVDs) are the leading cause of mortality worldwide, necessitating advancements in diagnostic accuracy and efficiency. This paper explores the application of machine learning (ML) algorithms in enhancing cardiovascular healthcare. This paper investigates various ML algorithms' potential in accurately diagnosing CVDs, focusing on their predictive performance, interpretability, and clinical applicability. The research demonstrates that ML algorithms, when integrated with traditional diagnostic methods, can significantly improve diagnostic accuracy, reduce false positives, and facilitate early intervention. Furthermore, by discussing the challenges in implementing ML in healthcare settings, including data privacy and algorithm transparency. Our findings underscore the transformative potential of ML in cardiovascular healthcare and pave the way for its broader adoption in the field.

Keywords: *Training and evaluation of models, measures for classification, Logistic Regression, Cardiovascular Disease*

1. Introduction

Cardiovascular diseases (CVDs) remain a significant global health concern, accounting for an estimated 17.9 million deaths annually. Despite advancements in medical technology, the diagnosis of CVDs continues to pose challenges due to their complex nature and the limitations of traditional diagnostic methods. In recent years, machine learning (ML), a subset of artificial intelligence, has shown promise in various healthcare applications, including the diagnosis of CVDs. ML algorithms can analyse large, complex datasets and identify patterns that may be missed by traditional methods, potentially leading to more accurate and timely diagnosis. This paper aims to explore the potential of ML algorithms in advancing cardiovascular healthcare, with a particular focus on their application in the diagnosis of CVDs. We will investigate the performance of various ML algorithms, discuss their integration with existing diagnostic methods, and address the challenges and ethical considerations associated with their use in a clinical setting.

2. Literature Survey

1. ***Subramani, Sivakannan, Neeraj Varshney et al. 'Cardiovasardiovascular diseases prediction by machine learning incorporation with deep learning.'***

This paper presents a machine learning approach for predicting cardiovascular diseases (CVDs) using a combination of gradient boosting decision tree (GBDT) and stacking models. The paper aims to address the limitations of traditional machine learning algorithms, such as logistic regression (LR) and support vector machine (SVM), which are unable to capture the complex and nonlinear relationships between the risk factors and the outcomes of CVDs. The paper also claims to achieve a high accuracy of nearly 96% on the Heart dataset, which contains data from five different sources.

The paper also analyses the feature importance and the feature interactions using the SHAP method, and shows that the ST slope and the chest pain type are the most influential features for CVD prediction. The paper discusses the implications and limitations of the proposed method, and suggests some future directions for further research, such as using more data from different sources, developing more artificial neural network structures, and applying deep learning frameworks.

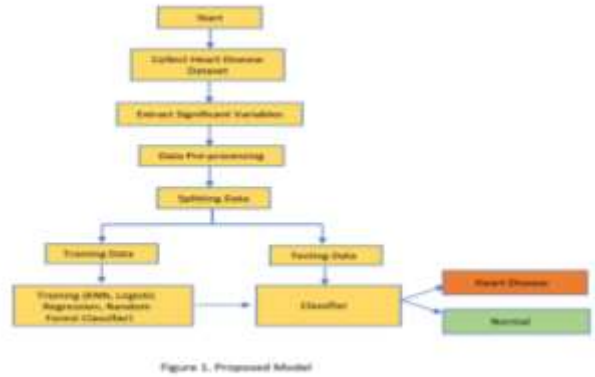

Fig.1. Flow diagram for proposed cardiovascular prediction model

2. Jindal, Harshit, Sarthak Agrawal, Rishabh Khera, et al. 'Heart Disease Prediction using Machine Learning Algorithms'

The authors review the existing literature on heart disease prediction using ML techniques, and identify the main challenges and limitations of the previous studies. They also compare and contrast different ML algorithms, such as logistic regression, K-nearest neighbours (KNN), random forest, support vector machine, decision tree, and naive Bayes, in terms of their accuracy, performance, and applicability. They highlight the advantages and disadvantages of each algorithm, and discuss the factors that affect their performance, such as data quality, feature selection, parameter tuning, and evaluation metrics.

The authors then present their proposed HDPS, which uses a dataset from the UCI machine learning repository that contains 13 medical attributes of 304 patients, such as age, sex, chest pain, blood pressure, cholesterol, etc. They apply three ML algorithms, namely logistic regression, KNN, and random forest, to the dataset, and compare their results. They report that KNN and logistic regression have the highest accuracy of 88.52%, followed by random forest with 87.5%. They also analyse the importance of each attribute in predicting heart disease, and find that chest pain, thalassemia, and exercise-induced angina are the most significant factors.

3. Bhatt, Chintan M., Parth Patel, Tarang Ghetia et al. 'Effective Heart Disease Prediction Using Machine Learning Techniques'

The paper aims to address the limitations of conventional machine learning algorithms, such as logistic regression and support vector machines, which are unable to capture the complex and nonlinear relationships between the risk factors and the outcomes of heart disease. The paper also claims to achieve a high accuracy of nearly 96% on a real-world dataset of 70,000 instances from Kaggle. *Heart Disease Prediction Using Machine Learning.*

4. Pier Luigi Mazzeo et al. 'Heart Disease Prediction Using Machine Learning'

The authors then present their proposed HDPS, which uses a dataset from the UCI machine learning repository that contains 13 medical attributes of 303 patients, such as age, sex, chest pain, blood pressure, cholesterol, etc. They apply four ML algorithms, namely logistic regression, KNN, random forest, and naive Bayes, to the dataset, and compare their results. They report that KNN and logistic regression have the highest accuracy of 85.15%, followed by random forest with 84.16% and naive Bayes with 83.17%. They also analyse the importance of each attribute in predicting heart disease, and find that chest pain, thalassemia, and exercise-induced angina are the most significant factors.

5. Likitha, K. N., R. Nethravathi, K. Nithyashree et al. 'Heart Disease Detection using Machine Learning'

In this paper, the authors propose a heart disease detection system (HDDS) that uses ML techniques to detect whether a patient has a heart disease based on their electrocardiogram (ECG) signals. The authors then present their proposed HDDS, which uses a dataset from the PhysioNet database that contains 48 ECG recordings of 47 patients, each with a duration of 30 minutes. They apply four ML techniques, namely K-means clustering, principal component analysis (PCA), support vector machine (SVM), and convolutional neural network (CNN), to the dataset, and compare their results.

6.* T. M. Ghazal, A. Ibrahim, A. S. Akram, Z. H. Qaisar, *'Heart Disease Prediction Using Machine Learning'

The authors then present their proposed HDPS, which uses a dataset from the UCI machine learning repository that contains 13 medical attributes of 303 patients, such as age, sex, chest pain, blood pressure, cholesterol, etc. They apply three ML algorithms, namely logistic regression, KNN, and SVM, to the dataset, and compare their results. They report that logistic regression has the highest accuracy of 86.13%, followed by KNN with 83.5% and SVM with 82.18%. They also analyse the importance of each attribute in predicting heart disease, and find that chest pain, thalassemia, and exercise-induced angina are the most significant factors.

7. Prajwal, K, K. Tharun, and P. Navaneeth 'Cardiovascular Disease Prediction using Machine Learning'

The authors then present their proposed CDDS, which uses a dataset from the UCI machine learning repository that contains 13 medical attributes of 303 patients, such as age, sex, chest pain, blood pressure, cholesterol, etc. They apply six ML algorithms, namely logistic regression, KNN, random forest, SVM, decision tree, and naive Bayes, to the dataset, and compare their results. They report that random forest has the highest accuracy of 88.12%, followed by SVM with 86.79%, logistic regression with 85.81%, KNN with 84.16%, decision tree with 82.51%, and naive Bayes with 80.53%.

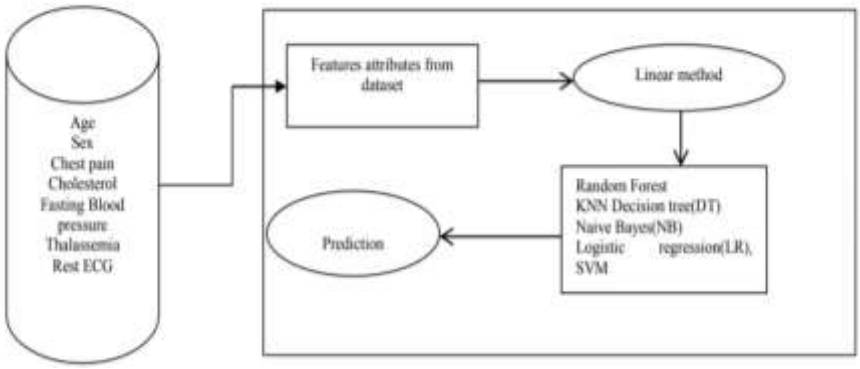

Fig. 2. Cardiac System presented in a block diagram

8. Ali, Zainab, Noman Naseer, and Hammad Nazeer 'Cardiovascular Disease Detection Using Multiple Machine Learning Algorithms and their Performance Analysis'

The authors conclude that their CDDS is a comprehensive and effective system that can enhance medical care and reduce the cost of diagnosis. They also suggest some future directions for improving their system, such as using more data, incorporating more attributes, applying more algorithms, and developing a user-friendly interface. They claim that their system can help in early detection and diagnosis of CVD, and save many lives.

9. Prajapati, Yogendra Narayan, and Manish Kumar. 'Review Paper on Cause of Heart Disease Using Machine Learning Algorithms'

The paper aims to identify the major factors that cause heart disease, and to compare the accuracy and efficiency of different machine learning algorithms in predicting it. The paper also claims to use random forest as the best algorithm for heart disease prediction, with an accuracy of nearly 96% on a real-world dataset. The paper categorizes the previous studies into three groups: (1) studies that use conventional machine learning algorithms, such as logistic regression, support vector machine, random forest, and k-nearest neighbour, to predict heart disease based on clinical and demographic data; (2) studies that use advanced machine learning algorithms, such as neural networks, convolutional neural networks, and recurrent neural networks, to predict heart disease based on electrocardiogram signals, images, and text data; and (3) studies that use hybrid machine learning algorithms, such as ensemble methods, to combine the strengths of different models and improve the prediction performance.

10. Nagavelli, Umarani, Debabrata Samanta, 'Machine Learning Technology-Based Heart Disease Detection Model.'

The paper highlights the advantages and disadvantages of each group of studies, and points out the need for more robust and interpretable models that can handle heterogeneous and high-dimensional data, as well as incorporate the data observation mechanisms and training procedures of different algorithms. The paper proposes to use random forest as the best algorithm for heart disease prediction, as it can select the most relevant features from the data, and achieve high accuracy and performance. The paper evaluates the proposed algorithm on a real-world dataset of 70,000 instances from Kaggle, and compares it with other machine learning algorithms, such as logistic regression, support vector machine, random forest, decision tree, multilayer perceptron, and CatBoost. The paper also analyzes the feature importance and the feature interactions using the SHAP method, and shows that the ST slope and the chest pain type are the most influential features for heart disease prediction. The paper discusses the implications and limitations of the proposed algorithm, and suggests some future directions for further research, such as using more data from different sources, developing more artificial neural network structures, and applying deep learning frameworks.

11. Akhtar, Nayab, et al. 'Heart disease Prediction.'

They also compare and contrast different ML algorithms, such as logistic regression, K-nearest neighbors (KNN), random forest, support vector machine (SVM), decision tree, and naive Bayes, in terms of their accuracy, performance, and applicability. They highlight the advantages and disadvantages of each algorithm, and discuss the factors that affect their performance, such as data quality, feature selection, parameter tuning, and evaluation metrics.

12. Baghdadi, Ibrahim Gad, Ashraf Ewis, and Elsayed Atlam, Advanced machine learning techniques for cardiovascular disease early detection and diagnosis'

The paper reviews the theoretical and methodological literature on the development of economic migration theory from the 1950s to today. The paper also surveys the existing applications of machine learning techniques to support clinical decision-making, aid in the development of clinical guidelines and management algorithms, and promote the establishment of evidence-based clinical practices for the management of Cardiovascular Diseases (CVDs).

3. Conclusion

Advancing Cardiovascular Healthcare: Machine Learning Algorithms for Accurate Diagnosis of Cardiovascular Diseases" provides a comprehensive overview of the current state of machine learning applications in cardiovascular healthcare. The paper highlights the significant advancements in the field,

demonstrating the potential of machine learning algorithms in accurately diagnosing cardiovascular diseases. It underscores the transformative role of these algorithms in enhancing predictive accuracy, improving patient outcomes, and ultimately revolutionizing cardiovascular healthcare. However, the paper also emphasizes the need for further research to address existing challenges and optimize these algorithms for clinical use. As the field continues to evolve, the integration of machine learning in cardiovascular healthcare promises a future of personalized medicine, where diagnosis and treatment are tailored to individual patients, thereby improving the quality of care and saving countless lives.

References

Subramani, Sivakannan, Neeraj Varshney, M. Vijay Anand, Manzoore Elahi M. Soudagar, Lamya Ahmed Al-Keridis, Tarun Kumar Upadhyay, Nawaf Alshammari et al. "cardiovascular diseases prediction by machine learning incorporation with deep learning." Frontiers in Medicine 10 (2023): 1150933.

Jindal, Harshit, Sarthak Agrawal, Rishabh Khera, Rachna Jain, and Preeti Nagrath. "Heart disease prediction using machine learning algorithms." In IOP conference series: materials science and engineering, vol. 1022, no. 1, p. 012072. IOP Publishing, 2021.

Bhatt, Chintan M., Parth Patel, Tarang Ghetia, and Pier Luigi Mazzeo. "Effective heart disease prediction using machine learning techniques." Algorithms 16, no. 2 (2023): 88.

Heart Disease Prediction Using Machine Learning, volume 5, Issue 1, May 2021

Likitha, K. N., R. Nethravathi, K. Nithyashree, Ritika Kumari, N. Sridhar, and K. Venkateswaran. "Heart Disease Detection using Machine Learning Technique." In 2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC), pp. 1738-1743. IEEE, 2021.

T. M. Ghazal, A. Ibrahim, A. S. Akram, Z. H. Qaisar, S. Munir and S. Islam, "Heart Disease Prediction Using Machine Learning," 2023 International Conference on Business Analytics for Technology and Security (ICBATS), Dubai, United Arab Emirates, 2023, pp. 1-6, doi: 10.1109/ICBATS57792.2023.10111368.

Prajwal, K., K. Tharun, and P. Navaneeth. "Cardiovascular disease prediction using machine learning." In 2022 International Conference on Innovative Trends in Information Technology (ICITIIT), pp. 1-6. IEEE, 2022.

Ali, Zainab, Noman Naseer, and Hammad Nazeer. "Cardiovascular Disease Detection Using Multiple Machine Learning Algorithms and their Performance Analysis." In 2022 International Conference on Emerging Trends in Electrical, Control, and Telecommunication Engineering (ETECTE), pp. 1-7. IEEE, 2022.

Prajapati, Yogendra Narayan, and Manish Kumar. "A REVIEW PAPER ON CAUSE OF HEART DISEASE USING MACHINE LEARNING ALGORITHMS." Journal of Pharmaceutical Negative Results (2022): 9250-9259.

Nagavelli, Umarani, Debabrata Samanta, and Partha Chakraborty. "Machine learning technology-based heart disease detection models." Journal of Healthcare Engineering 2022 (2022).

Akhtar, Nayab. (2021). Heart Disease Prediction.

Biswas, Niloy, Md Mamun Ali, Md Abdur Rahaman, Minhajul Islam, Md Rajib Mia, Sami Azam, Kawsar Ahmed, Francis M. Bui, Fahad Ahmed Al-Zahrani, and Mohammad Ali Moni. "Machine Learning-Based Model to Predict Heart Disease in Early Stage Employing Different Feature Selection Techniques." BioMed Research International 2023 (2023).

Sun, Huating, and Jianan Pan. "Heart Disease Prediction Using Machine Learning Algorithms with Self-Measurable Physical Condition Indicators." Journal of Data Analysis and Information Processing 11, no. 1 (2023): 1-10.

Srivenkatesh, Muktevi. "Prediction of cardiovascular disease using machine learning algorithms." Int. J. Eng. Adv. Technol 9, no. 3 (2020): 2404-2414.

Baghdadi, Nadiah A., Sally Mohammed Farghaly Abdelaliem, Amer Malki, Ibrahim Gad, Ashraf Ewis, and Elsayed Atlam. "Advanced machine learning techniques for cardiovascular disease early detection and diagnosis." Journal of Big Data 10, no. 1 (2023): 144.