EAI Endorsed Transactions
on Pervasive Health and Technology
Research Article EAI.EU

A Review: Machine Learning and Data Mining Approaches for Cardiovascular Disease Diagnosis and Prediction

Gorapalli Srinivasa Rao[1*], G Muneeswari[2]

[1,2]School of Computer Science and Engineering, VIT-AP University, Amaravathi, Andhrapradesh, India

Abstract

INTRODUCTION: Cardiovascular disease (CVD) is the most common cause of death worldwide, and its prevalence is rising in low-resource settings and among those with lower incomes.
OBJECTIVES: Machine learning (ML) algorithms are quickly evolving and being implemented in medical procedures for CVD diagnosis and treatment decisions. Every day, the healthcare business creates massive amounts of data. However, the majority of it is inadequately utilized. Efficient techniques for extracting knowledge from these datasets for clinical diagnosis or other uses are scarce.
METHODS: ML is being applied in the healthcare industry all over the world. In the health dataset, ML approaches useful in the prevention of locomotor disorders and heart disease.
RESULTS: The revelation of such vital information allows researchers to acquire significant insight into how to use the proper treatment and diagnosis for a specific patient. Researchers study enormous volumes of complex healthcare data using various ML approaches, which improves healthcare professionals in disease prediction.
CONCLUSION: The goal of this study is to summarize some of the current research on predicting heart diseases utilizing machine learning and data mining techniques, analyze the various mining algorithm combinations employed, and determine which techniques are useful and efficient. Future directions in prediction systems have also been considered.

Keywords: Heart Diseases, Machine Learning, Ensemble Models, Data Mining, Dataset, Classification

Received on 08 December 2023, accepted on 07 March 2024, published on 13 March 2024

doi: 10.4108/eetpht.10.5411

1. Introduction

The heart is one of the most vital organs in human body. It is a muscular organ that pumps blood into the body and serves as the heart of the cardiovascular system. The cardiovascular system contains all blood vessels, including veins, capillaries, and arteries, which together constitute difficult blood circulation throughout the body [1-3]. Any restriction or aberration in normal blood circulation flow from the heart can lead to some significant heart disease complications. These are referred to as CVDs, and they are among the deadliest illness in the world. CVDs include heart disorders, vascular diseases of the brain, and blood vessel diseases [4, 5]. Although CVDs can be managed through lifestyle changes and other related measures, all indicators indicate that they are on the rise daily, as mentioned in several WHO studies. However, many WHO publications have shown a global increase in CVDs, which is highly concerning. CVDs kill more people than any other cause worldwide, killing an estimated 17.5 million individuals in 2012 [6, 7]. According to numerous WHO estimates, mortality from heart disease is on the rise, which is mostly linked to insufficient preventative actions despite rising risk factors.
Clinical information has revealed that some risk factors create a person's chances of building CVD. Some of these risk factors include a family heritage of cardiovascular disease, a bad LDL cholesterol level, a low good cholesterol level, a high-fat diet, hypertension, obesity, and a lack of active

*Corresponding author. Email: srinivasarao.22phd7042@vitap.ac.in

9. Conclusion

The use of several data mining and ML methods for predicting the occurrence of heart disease has been summarized. Determine every algorithm's prediction accuracy and apply the suggested approach to the required area. Improve algorithm accuracy by using more relevant attribute selection methods. If a patient is diagnosed with a specific type of cardiac disease, there are numerous treatments available. Data mining and ML may provide an enormous amount of information from an appropriate dataset. In conclusion, the literature survey revealed that only marginal success is reached in the development of predictive models for heart disease patients, indicating the need for combinational and more complex systems to improve the accuracy of identifying the early onset of heart disease. The more data that is supplied into the database, the more intelligent the system becomes.

An automated model that can assist in the choice of suitable treatment approaches for a heart disease patient may be developed in the future. A lot of studies have already been put into establishing systems that can detect whether an individual is going to develop cardiovascular disease or not. When an individual is identified with a specific form of heart disease, he or she has several treatment options. Machine Learning can be quite useful in selecting the type of treatment to be taken by retrieving information from such relevant databases. In addition, we are interested in classifying it as a multi-class problem to determine the disease's level. In order to more accurately and effectively anticipate heart problems, an ideal method needs to be developed.

References

[1] U. Nagavelli, D. Samanta & P. Chakraborty, "Machine learning technology-based heart disease detection models," *Journal of Healthcare Engineering*, 2022.

[2] A. Rath, D. Mishra, G. Panda & S. C. Satapathy, "Heart disease detection using deep learning methods from imbalanced ECG samples," *Biomedical Signal Processing and Control*, vol. 68, 102820, 2021.

[3] S. Ketu & P. K. Mishra, "Empirical analysis of machine learning algorithms on imbalance electrocardiogram based arrhythmia dataset for heart disease detection," *Arabian Journal for Science and Engineering*, pp. 1-23, 2022.

[4] R. Arnaout, L. Curran, Y. Zhao, J. C. Levine, E. Chinn & A. J. Moon-Grady, "An ensemble of neural networks provides expert-level prenatal detection of complex congenital heart disease," *Nature Medicine*, vol. 27(5), pp. 882-891, 2021.

[5] F. HASSAN, A. U. Rahman, A. Javed, A. Alhazmi & M. Alhazmi, "CNN-CardioAssistant: Deep Convolutional Neural Network and Recursive Feature Elimination Method for Heart Disease Detection," 2022.

[6] F. F. Firdaus, H. A. Nugroho & I. Soesanti, "Deep neural network with hyperparameter tuning for detection of heart disease," In 2021 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob) IEEE, pp. 59-65, 2021.

[7] M. Alkhodari & L. Fraiwan, "Convolutional and recurrent neural networks for the detection of valvular heart diseases in phonocardiogram recordings," *Computer Methods and Programs in Biomedicine*, vol. 200, 105940, 2021.

[8] T. S. Roy, J. K. Roy & N. Mandal, "Classifier identification using deep learning and machine learning algorithms for the detection of valvular heart diseases," *Biomedical Engineering Advances*, vol. 3, 100035, 2022.

[9] A. Thakur, H. Kaur, N. Goel, P. Paul, P. Asopa, S. Goswami & M. K. Das, "A Hybrid Approach for Heart Disease Detection using K-Means and K-NN Algorithm, "*American Journal of Electronics & Communication*, vol. 4(1), pp. 14-21, 2023.

[10] I. D. Mienye & Y. Sun, "Improved heart disease prediction using particle swarm optimization based stacked sparse autoencoder," *Electronics*, vol. 10(19), 2347, 2021.

[11] P. Srinivas & R. Katarya, "hyOPTXg: OPTUNA hyper-parameter optimization framework for predicting cardiovascular disease using XGBoost," *Biomedical Signal Processing and Control*, vol. 73, 103456, 2022.

[12] P. Rani, R. Kumar, N. M. S. Ahmed & A. Jain, "A decision support system for heart disease prediction based upon machine learning," *Journal of Reliable Intelligent Environments*, vol. 7(3), pp. 263-275, 2021.

[13] M. W. Nadeem, H. G. Goh, M. A. Khan, M. Hussain & M. F. Mushtaq, "Fusion-Based Machine Learning Architecture for Heart Disease Prediction. Computers," *Materials & Continua*, vol. 67(2), 2021.

[14] K. Budholiya, S. K. Shrivastava & V. Sharma, "An optimized XGBoost based diagnostic system for effective prediction of heart disease," *Journal of King Saud University-Computer and Information Sciences*, vol. 34(7), pp. 4514-4523, 2022.

[15] X. Y. Gao, A.Amin Ali, H. Shaban Hassan & E. M. Anwar, "Improving the accuracy for analyzing heart diseases prediction based on the ensemble method," *Complexity*, pp. 1-10, 2021.

[16] M. N. Uddin & R. K. Halder, "An ensemble method based multilayer dynamic system to predict cardiovascular disease using machine learning approach," *Informatics in Medicine Unlocked*, vol. 24, 100584, 2021.

[17] T. R. Mahesh, V. Dhilip Kumar, V. Vinoth Kumar, J. Asghar, O. Geman, G. Arulkumaran & N. Arun, "AdaBoost ensemble methods using K-fold cross validation for survivability with the early detection of heart disease," *Computational Intelligence and Neuroscience*, 2022.

[18] M. Kavitha, G. Gnaneswar, R. Dinesh, Y. R. Sai, & R. S. Suraj, "Heart disease prediction using hybrid machine learning model," In 2021 6th international conference on inventive computation technologies (ICICT) IEEE, pp. 1329-1333, 2021.

[19] A. Yazdani, K. D. Varathan, Y. K. Chiam, A. W. Malik & W. A. Wan Ahmad, "A novel approach for heart disease prediction using strength scores with significant predictors," *BMC medical informatics and decision making*, vol. 21(1), 194, 2021.

[20] J. Premsmith & H. Ketmaneechairat, "A predictive model for heart disease detection using data mining techniques," *Journal of advances in Information Technology*, vol. 12(1), 2021.

[21] R. Valarmathi & T. Sheela, "Heart disease prediction using hyper parameter optimization (HPO) tuning," *Biomedical Signal Processing and Control*, vol. 70, 103033, 2021.

[22] D. Deepika & N. Balaji, "Effective heart disease prediction using novel MLP-EBMDA approach," *Biomedical Signal Processing and Control*, vol. 72, 103318, 2022.

[23] A. A. Nancy, D. Ravindran, P. D. Raj Vincent, K. Srinivasan & D. Gutierrez Reina, "Iot-cloud-based smart healthcare monitoring system for heart disease prediction via deep learning, "*Electronics*, vol. 11(15), 2292, 2022.
[24] A. A. Malibari, "An efficient IoT-Artificial intelligence-based disease prediction using lightweight CNN in healthcare system," *Measurement: Sensors*, 26, 100695, 2023.
[25] M. M. Yaqoob, M. Nazir, A. Yousafzai, M. A. Khan, A. A. Shaikh, A. D. Algarni & H. Elmannai, "Modified Artificial Bee Colony Based Feature Optimized Federated Learning for Heart Disease Diagnosis in Healthcare," *Applied Sciences*, vol. 12(23), 12080, 2022.
[26] A. Mehmood, M. Iqbal, Z. Mehmood, A. Irtaza, M. Nawaz, T. Nazir & M. Masood, 'Prediction of heart disease using deep convolutional neural networks," *Arabian Journal for Science and Engineering*, vol. 46(4), pp. 3409-3422, 2021.
[27] P. Dileep, K. N. Rao, P. Bodapati, S. Gokuruboyina, R. Peddi, A. Grover & A. Sheetal, "An automatic heart disease prediction using cluster-based bi-directional LSTM (C-BiLSTM) algorithm*," Neural Computing and Applications*, vol. 35(10), pp. 7253-7266, 2023.