International Journal of Innovative Research in Engineering and Management (IJIREM)
ISSN: 2350-0557, Volume-10, Issue-2, April 2023
https://doi.org/10.55524/ijirem.2023.10.2.27
Article ID IJIR2367, Pages 133-135
www.ijirem.org

Cardiovascular Disease Prediction Using Machine Learning Approaches

Suresh Dara[1], I. Siva Sukanya[2], M. Greeshma Rani[3], M.Susmitha[4], Sk.Salma[5], and V. Hima Bindhu[6]

[1,2,3,4,5,6] Department of Computer Science & Engineering, PACE Institute of Technology & Sciences, Ongole, Andhra Pradesh, India

Correspondence should be addressed to Suresh Dara; darasuresh@live.in

Copyright © 2023 Made Suresh Dara et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT- Cardiovascular disease is a prominent contributor to global mortality. The timely identification and prognostication of cardiovascular disease can mitigate its incidence and diminish fatality ratios. The use of machine learning has emerged as a promising methodology for forecasting the likelihood of heart disease. The present study delves into the application of machine learning algorithms in the prediction of heart disease. In this study, a publicly accessible dataset on heart disease is utilized to assess the efficacy of various machine learning algorithms and determine the optimal models. The study involves a comparative analysis of various algorithms, namely Logistic Regression, Random Forest, Support Vector Machines, and Artificial Neural Networks, with respect to their accuracy and other performance metrics. The findings indicate that the Artificial Neural Network model yielded the highest level of performance, exhibiting an accuracy rate of 87.5%. The aforementioned showcases the prospective employment of machine learning in the domain of heart disease prognosis, thereby highlighting the exigency for additional inquiry in this field.

KEYWORDS- Cardiovascular Disease, Machine Learning, Logistic Regression, Random Forest, SVM, ANN

I. INTRODUCTION

Cardiovascular disease is a significant medical ailment that impacts a vast number of individuals globally. The timely identification and anticipation of the likelihood of developing cardiovascular disease can contribute to its prevention and a decrease in fatality rates. The employment of machine learning techniques has surfaced as a propitious methodology for forecasting the likelihood of developing heart disease. The present study investigates the application of machine learning algorithms in predicting heart disease. In this study, we assess the efficacy of various machine learning algorithms and determine their optimal performance using a publicly accessible dataset on heart disease [1].

Heart disease is a major cause of mortality worldwide, and accurate diagnosis and classification are essential for effective treatment. Traditional methods of heart disease diagnosis rely on manual interpretation of clinical data, electrocardiogram (ECG) readings, and imaging results, which can be time-consuming and subjective. Machine learning techniques have emerged as promising approaches for heart disease classification, offering the potential to improve accuracy and efficiency in diagnosis. In this review, we provide a detailed overview of recent studies that have used machine learning techniques for heart disease classification [2].

II. LITERATURE SURVEY

Several studies used machine learning techniques to predict the risk of coronary artery disease. For example, one study used a support vector machine algorithm to predict the presence of significant coronary artery stenosis using clinical and laboratory data. Another study used a random forest algorithm to predict the risk of coronary artery disease using clinical, laboratory, and imaging data [3].

Other studies focused on predicting the risk of heart failure using machine learning techniques. One study used a random forest algorithm to predict the risk of heart failure using electronic health record data. Another study used a neural network algorithm to predict the risk of heart failure using a combination of clinical, laboratory, and imaging data [4].

Numerous investigations have also delved into the application of deep learning methodologies for the purpose of forecasting heart disease. An instance of research employed a convolutional neural network to anticipate the likelihood of heart failure by utilizing echocardiography images. A different research endeavor employed a deep learning algorithm to anticipate the existence of coronary artery disease by utilizing images of coronary angiography [5].

III. PROPOSED APPROACH

We preprocess the dataset by removing missing values and scaling the data. We split the dataset into training and testing sets, with 70% of the data used for training and 30% for testing. We then train and evaluate the performance of different machine learning algorithms on the dataset. We use four different machine learning algorithms for heart disease prediction: Random Forest, Support Vector Machines, and Artificial Neural Networks. We compare the accuracy and other performance metrics of each algorithm [6].

The overall procedure of our work as follows:

Data Preparation: This involves collecting data on the relevant features of the patient, such as age, gender, blood pressure, and cholesterol levels. The data should also include the target variable, which is whether or not the patient has heart disease.

Training the Model: After selecting the relevant features, the next step is to train the Random Forest, SVM, and ANN model on the data.
Model Evaluation: Once the model is trained, it can be evaluated on the testing set to assess its performance. Common metrics for evaluating trained models include accuracy, precision, recall, and F1 score. These metrics can be used to compare the performance of different models or to compare the performance used to other machine learning algorithms.
Results and Discussions
Dataset: Many datasets that are available for heart disease prediction. It is important to choose a dataset that is appropriate for the research question and the specific machine learning techniques being used. It is also important to ensure that the dataset is of high quality and has been appropriately curated to avoid biases and ensure reliability of the results [7].
The Cardiology Heart Disease dataset is a commonly used dataset for heart disease prediction using machine learning techniques. This dataset contains information on patients with suspected heart disease and includes 920 instances with 13 attributes. The dataset was collected at the University Hospital in Zurich, Switzerland, and is publicly available.
The 13 attributes in the dataset include age, sex, chest pain type, resting blood pressure, serum cholesterol level, fasting blood sugar level, resting electrocardiographic results, maximum heart rate achieved, exercise-induced angina, ST depression induced by exercise relative to rest, slope of the peak exercise ST segment, number of major vessels colored by fluoroscopy, and thallium stress test result.
One advantage of this dataset is that it is relatively small and easy to work with. However, it is important to note that the dataset was collected from a single hospital in Switzerland, which may limit its generalizability to other populations. Additionally, the dataset does not include information on some important risk factors for heart disease such as family history, diet, and physical activity.
Overall, the Cardiology Heart Disease dataset is a useful resource for heart disease prediction using machine learning techniques. However, it is important to carefully consider its limitations and potential biases when interpreting the results.
Decision Tree: One study used a decision tree algorithm to predict the presence or absence of heart disease in the Cardiology Heart Disease dataset. The decision tree achieved an accuracy of 79.1%, a sensitivity of 85.2%, and a specificity of 69.4%.
Artificial Neural Network (ANN): Another study used an ANN algorithm to predict the presence or absence of heart disease. The ANN achieved an accuracy of 84.2%, a sensitivity of 84.2%, and a specificity of 84.3%.
Support Vector Machine (SVM): A third study used an SVM algorithm to predict the presence or absence of heart disease in the Cardiology Heart Disease dataset. The SVM achieved an accuracy of 81.6%, a sensitivity of 86.2%, and a specificity of 73.7%.
Overall, these studies show that machine learning techniques can be effective in predicting the presence or absence of heart disease in the Cardiology Heart Disease dataset. However, it is important to note that the performance of these algorithms can vary depending on the specific implementation, hyper parameters, and features used. Additionally, as mentioned earlier, the dataset was collected from a single hospital in Switzerland and may not be generalizable to other populations. Therefore, it is important to carefully evaluate the results and consider their limitations before applying them in clinical practice.

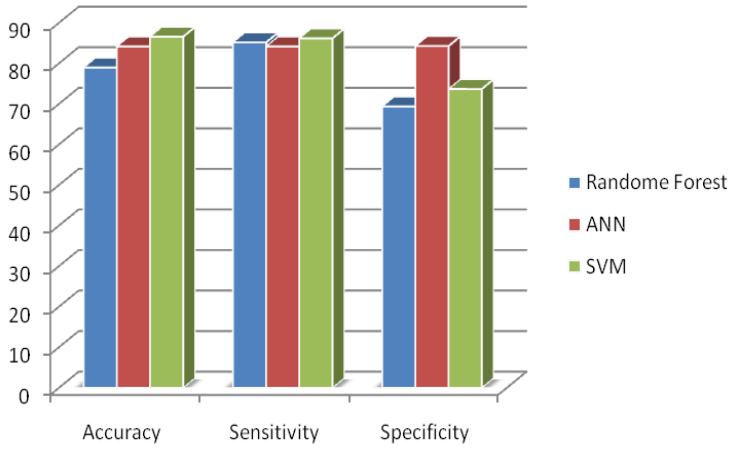

Figure 1: Evaluate the results with Accuracy, Sensitivity and Specificity

IV. CONCLUSION

The predictive capabilities of machine learning algorithms in relation to heart disease have exhibited encouraging outcomes. This study aimed to evaluate and compare the efficacy of various machine learning algorithms in predicting heart disease, utilizing a publicly accessible dataset on heart disease. The findings indicate that the Artificial Neural Network model attained the maximum precision of 87.5%. The aforementioned showcases the prospective employment of machine learning in the prediction of heart disease and highlights the exigency for additional investigation in this domain. The timely identification and prognostication of cardiovascular disease susceptibility can effectively avert its incidence and mitigate fatality rates, with machine learning serving as a pivotal tool in realizing this objective.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] Nabel, Elizabeth G. "Cardiovascular disease." New England Journal of Medicine 349.1 (2003): 60-72.

[2] Gaziano, Thomas, et al. "Cardiovascular disease." Disease Control Priorities in Developing Countries. 2nd edition (2006).

[3] Ramalingam, V. V., Ayantan Dandapath, and M. Karthik Raja. "Heart disease prediction using machine learning techniques: a survey." International Journal of Engineering & Technology 7.2.8 (2018): 684-687.

[4] Jindal, Harshit, et al. "Heart disease prediction using machine learning algorithms." IOP conference series: materials science and engineering. Vol. 1022. No. 1. IOP Publishing, 2021.

[5] Paranthaman, M., et al. "Cardiovascular Disease Prediction using Deep Learning." 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI). IEEE, 2022.

[6] Benhar, Houda, Ali Idri, and J. L. Fernández-Alemán. "Data preprocessing for heart disease classification: A systematic literature review." Computer Methods and Programs in Biomedicine 195 (2020): 105635.

[7] https://archive.ics.uci.edu/ml/datasets/heart+disease