GLIMPSE-Journal of Computer Science •Vol.3(1),JANUARY-JUNE2024,pp. 25-28

EXPLORING THE IMPACT OF MACHINE LEARNING ALGORITHMS ON CARDIO VASCULAR HEALTH ASSESSMENT

Nishant Kumar Pathak[1], Varun Kumar[2], Ashish Kumar[3]
Assistant Professor, Computer Science Engineering, Ajay Kumar Garg Engineering College, Ghaziabad

Abstract*—After the brain, the human heart stands as the second most vital organ in the body. Atherosclerosis, a leading cause of death globally and a significant health concern, emphasizes the need for precise medical procedures. Unfortunately, relying solely on experience and instinct may not always yield successful outcomes. To address medical errors and unfavorable results, computer-based diagnostic techniques become imperative. Implementing these techniques not only reduces the risk of fatal medical errors but also enhances patient safety, ultimately saving lives. This research aims to identify the most efficient and accurate machine learning method for predicting cardiac illnesses. The study categorizes machine learning techniques based on their effectiveness in identifying cardiac issues. Preferred classification methods in this study encompass logistic regression, decision trees, neural networks, support vector machines, and rudimentary Bayes. The overarching goal is to determine the effectiveness of these methods in improving cardiac health *diagnostics."

***Keywords*—Humanheart, Atherosclerosis, Logistic regression, Decision trees, Neural networks, Support vector machines, Rudimentary Bayes, Cardiac health diagnostics**

I. INTRODUCTION

The heart demands special attention given its size and significance, and it is crucial to take proper care of it. Considering the heart's involvement in numerous diseases, there is a pressing need to anticipate cardiac ailments, necessitating research and analysis in the field. It becomes imperative to comprehend more effective sickness prediction algorithms, as many deaths result from diseases that go undetected until they have advanced, often due to instrument inaccuracy [1].The objective of artificial intelligence (AI) technology is to enhance the intelligence of computers. Learning, being a fundamental aspect of intelligence, establishes machine learning as a subfield within artificial intelligence (AI). Notably, machine learning (ML) stands out as one of the rapidly expanding domains within artificial intelligence (AI). Its applications extend beyond the medical industry. Considering the substantial data generated by the healthcare sector, the implementation of machine learning emerges as an advanced technique for comprehensive data analysis, offering significant advantages to the industry. The digital revolution of recent years has resulted in the recording and storage of massive volumes of data [2]. In the human body, various organs perform distinct functions. When the heart, one of these crucial organs, fails to adequately pump blood throughout the body, it poses a risk to the person's health. Presently, coronary artery disease stands as the leading cause of death in the modern era [3]. Research has demonstrated that employing a prediction model to identify high-risk individuals and detect cardiac disease early can reduce death rates and improve decision-making concerning additional treatment and prevention measures. In assessing the risks of cardiovascular disease and prescribing medications to mitigate those risks[4,5], Clinical Decision Support Systems (CDSS) utilize a risk prediction model. The implementation of CDSS has shown to enhance clinical decision-making, preventative care, and overall decision quality, as indicated by data from numerous studies. As plaque accumulates on the walls of a patient's coronary arteries, the arteries gradually narrow. The reduced blood flow to the heart's muscles leads to a slower heartbeat. If an artery is partially or totally blocked, there is an increased likelihood of a myocardial infarction, commonly known as an ischemic stroke. Several factors elevate the risk of developing coronary heart disease, including smoking, aging, high blood pressure, high LDL cholesterol, uncontrolled diabetes, and hypertension [6].

The limited accuracy and extended processing time of current heart disease diagnosis methods contribute to their inadequacy in early detection. Researchers are actively working to devise a more effective method for the early detection of cardiac disorders[7]. Various studies have utilized data to address these concerns. Many investigators have employed a range of prediction algorithms documented in the literature; however, these techniques lack reliability in predicting cardiac issues. Standardizing the data used by machine learning classifiers is essential to enhance their efficacy. Various alternative standardization techniques, including min-max tensor and standard tensor (SS), are utilized to address instances in the dataset with missing feature values. Conducting tests for heart disease prediction is a complex process, and the identification of an appropriate timeframe poses a challenge. Predicting cardiovascular disease is inherently difficult, with the challenge being exacerbated in nations facing shortages of trained medical

personnel, diagnostic tools, and other resources necessary for the identification and treatment of heart disease [8]. Due to the unpredictable nature of the disease course, active heart failure (HF) therapy is frequently maintained until the later stages of illness, leading to a significant upsurge in resource demand in the last six months of life [9]. The identification of novel elements in diagnostic criteria and referral requirements for various illnesses has emerged from recognizing gender differences between men and women. Research suggests that in the realm of heart disease treatments, including percutaneous coronary intervention (PCI) and coronary artery bypass graft surgery (CABG), women exhibit better outcomes compared to men [10]. Artificial intelligence (AI), particularly in the form of deep learning, utilizes neural networks and mathematical calculations to perform specific tasks. Cardiovascular disease is highly prevalent within certain populations and can be influenced by factors such as cigarette smoking, diabetes, high blood pressure, obesity, hypertension, cholesterol, and other issues associated with an unrestricted diet high in saturated fat and a sedentary lifestyle. This syndrome is further compounded by elevated blood glucose levels during fasting, increased blood pressure, elevated cholesterol, and an abnormal resting electrocardiogram (ECG). Conditions like heart failure, angina pectoris, high blood pressure, cardiovascular disease, and stroke are among the cardiac issues that may contribute to the development of this illness [11].

II. DIFFERENT TECHNIQUES USED BY ML

Prior knowledge serves as the foundation for a thorough comprehension and analysis of any topic and is essential for success in any educational discipline[14]. Therefore, before delving into the actual content of this document, it is necessary to explore and understand the fundamental ideas associated with it. This approach will enable us to fully grasp and appreciate the article.

A. A Support Vector Machine (SVM) is a supervised machine learning algorithm used for classification and regression tasks. Its primary objective[15] is to find a hyperplane in an N-dimensional space (where N is the number of features) that distinctly classifies the data points into different classes [16]. SVM is effective in high-dimensional spaces, making it suitable for various applications, including image classification, text classification, and bioinformatics [17].

B. Naïve Bayes is a well-liked probabilistic machine learning algorithm for classification tasks[18]. It's based on the Bayes theorem and the feature independence assumption, which makes it easier to compute and hence deemed "naïve[19]." Despite its simplicity, Naïve Bayes is surprisingly effective in many real-world applications, particularly in text classification and spam filtering[20].

(i) Bayes' Theorem: The Bayes theorem, which determines a hypothesis's probability based on past knowledge, forms the foundation [21] of the algorithm. It is used to determine the likelihood of a specific class given a set of features in the context of Naïve Bayes.

$$P(X \mid Y)=P(Y \mid X)\cdot P(X)/p(X)$$

Where:
P(X|Y) is the posterior probability of class X given feature Y.
P(Y|X) is the likelihood of feature Y given class X.
P(X) is the prior probability of class X.
P(Y) is the probability of feature Y.

C. Machine learning makes use of a statistical method called logistic regression. This strategy may be used in binary classification, which uses only two categories to make distinctions. The aim of both logistic and linear regression is to determine the proper coefficient values for each input variable. Instead of using linear regression to create the prediction in this case, a logistic function—a non-linear function—is utilized [12]. Machine learning makes use of a statistical method called logistic regression. This strategy may be used in binary classification, which uses only two categories to make distinctions. The aim of both logistic and linear regression is to determine the proper coefficient values for each input variable. Instead of using linear regression to create the prediction in this case, a logistic function—a non-linear function—is utilized [12].

D. Random Forest This method combines multiple decision trees to provide a prediction that is more reliable and accurate. The random forest predicts the outcome based on which predictions received the most votes from the population, as opposed to depending just on one decision tree [13].

E. The k-nearest neighbor algorithm is one straightforward but efficient categorization method. As it does not rely on any simplifying assumptions, it is useful for addressing classification problems when the time series is unknown. Next, an average value derived from the recovered data points is assigned to the k training set data points that most closely resemble the information item with the missing target value. Supervised machine learning has been using the k-nearest neighbor algorithm successfully for many years. The KNN method assigns a new instance or input to the group that is most similar to the groups that came before it based on the assumption that it is related to earlier instances [26].

F. Decision tree: Based on the observed data points, the decision tree classifier constructs a tree and uses it to make decisions. In this diagram, target values are indicated by leaves on branches. Using a decision tree to categorize cardiac diagnoses The stems and branches depict the class, while the leaves are labeled with the appropriate class. Defining traits for every class. Data subsets are distinguished by their visual characteristics[16]. Users can evaluate the decision tree model's dependability with the aid of the XTrain parameters

and the YTrain fit form. By passing the XTest and YTest parameters to the system's score() function, one can determine the score of a decision tree model

III. RELATED WORK

One of the main reasons scientists are focusing on this subject is that the heart is one of the body's most vital organs. It significantly accelerates the pumping of blood, which is as crucial to the body as oxygen. This is why a considerable number of researchers are currently dedicated to it. Analyzing heart-related issues is always necessary, whether the goal is to detect, predict, or prevent heart disease. Data mining, machine learning, and artificial intelligence are just a few of the disciplines influencing this research. Some examples of modern scholarly works are as follows:

Anurag Jain and his associates developed a hybrid decision support system in 2021. Based on clinical evidence, this approach may aid in the early diagnosis of heart disease. Its accuracy of 86.6% is higher than that of any other heart disease prediction technology currently in use. By employing a random forest classifier, the algorithm has demonstrated its ability to yield the majority of accurate results. A Python-based simulation environment was used to analyze the suggested hybrid system. The analyses were conducted using the Cleveland heart disease dataset that was made available in the UCI machine learning repository [24].

Maad M. Mijwil and colleagues (2023) assert that machine learning (ML) techniques are predictive and play a crucial role in disease investigation and identification. This article compares and contrasts the five most popular methods for diagnosing and determining the causes of cardiac problems. The support vector machine method completed the task in 317 seconds, while the artificial neural network method took the longest to implement, requiring 501 seconds. The article identifies best and worst practices based on performance. Machine learning techniques are deemed essential for the healthcare industry as they contribute to informed decision-making [25]. 2023 will see researchers Binit Patel et al. By developing a range of machine learning algorithms based on personal medical information from the UCI data set, this study aims to enhance the early identification of cardiovascular disease. The techniques demonstrated have the highest accuracy to date, at 95.08%. Both the xgboost boost and the classifier using random forests were successful in achieving this. The proposed technology can only identify the existence of a cardiac condition; it cannot determine the severity of heart disease [26].

IV. CONCLUSION AND FUTURE WORK

This section describes how to use the given situation to predict cardiac disease using Python. Python is a high-level, object-oriented programming language with a focus on reusability, lively, active construction possibilities, and short development cycles. The cardiovascular care sector makes efficient use of the data that is available by gathering information froma range of hospitals and patients. Professionals are also just exhibiting this enhanced treatment plan, which will improve the system's overall quality of healthcare delivery. The experiment's algorithms performed fairly well using the given attributes. Lastly, it can be said that machine learning may mitigate the damage to an individual's physical and mental well-being by forecasting cardiac illness. We conclude that different machine learning and data mining techniques have been used in combination with different instruments to anticipate the onset of heart disease based on an analysis of several recent studies that used diverse data mining and learning methods and codes to predict the onset of heart attacks. Different patient datasets with heart disease are used in multiple trials. The majority of experiments use a small, homogeneous dataset to train prediction models.

Therefore, we need to gather real data from a sizable number of heart disease patients from prestigious medical institutions across the nation in order to train and validate our prediction models.Next, we need to evaluate the accuracy of our prediction models using large datasets.The suggested system also has a number of problems. The severity of cardiac disease cannot be determined using this method. Larger datasets with more features and a wider range of data mining techniques will be used in future research.

REFERENCES

[1] Soni, S. K., Madan Mohan Malaviya University of Technology, North Dakota State University, Institute of Electrical and Electronics Engineers. Uttar Pradesh Section, IEEE Industry Applications Society, & Institute of Electrical and Electronics Engineers. (n.d.). ICE3-2020 : International Conference on Electrical and Electronics Engineering : February 14-15, 2020.

[2] Soni, S. K., Madan Mohan Malaviya University of Technology, North Dakota State University, Institute of Electrical and Electronics Engineers. Uttar Pradesh Section, IEEE Industry Applications Society, & Institute of Electrical and Electron-ics Engineers. (n.d.). ICE3-2020 : International Conference onElectrical and Electronics Engineering : February 14-15, 2020.

[3] Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques. IEEE Access, 7, 81542–81554. https://doi.org/10.1109/ACCESS.2019.2923707

[4] Zhenya, Q., & Zhang, Z. (2021). A hybrid cost-sensitive ensemble for heart disease prediction. BMC Medical Informatics and Decision Making, 21(1). https://doi.org/10.1186/s12911-021-01436-7

[5] Alarsan, F. I., & Younes, M. (2019). Analysis and classification of heart diseases using heartbeat features and machine learning algorithms. Journal of Big Data, 6(1). https://doi.org/10.1186/s40537-019-0244-x

[6] Nabel, E. G., & Braunwald, E. (2012). A Tale of Coronary Artery Disease and Myocardial Infarction. New England Journal of Medicine, 366(1), 54–63. https://doi.org/10.1056/nejmra1112570

[7] Allen, L. A., Grady, K. L., Goldstein, N. E., Matlock, D. D., Arnold, R. M., Cook, N. R., Felker, ; G Michael, Francis, G. S., Hauptman, P. J., Havranek, E. P., Krumholz, H. M., Mancini, D., Riegel, B., & Spertus,

[8] J. A. (2012). AHA Scientific Statement Decision Making in Advanced Heart Failure A Scientific Statement From the American Heart Association Endorsed by Heart Failure Society of America, American Association of Heart Failure Nurses, and Society for Medical Decision Making Association Council on Quality of Care and Outcomes Research, Council on Cardiovascular Nursing, Council on Clinical Cardiology, Council on Cardiovascular Radiology and Intervention, and Council on Cardiovascular Surgery and Anesthesia. https://doi.org/10.1161/CIR.0b013e31824f2173/-/DC1

[9] Ghwanmeh, S., Mohammad, A., & Al-Ibrahim, A. (2013). Innovative Artificial Neural Networks-Based Decision Support System for Heart Diseases Diagnosis. Journal of Intelligent Learning Systems and Applications, 05(03), 176–183. https://doi.org/10.4236/jilsa.2013.53019

[10] Lesyuk, W., Kriza, C., & Kolominsky-Rabas, P. (2018). Cost-of-illness studies in heart failure: A systematic review 2004-2016. BMC Cardiovascular Disorders, 18(1). https://doi.org/10.1186/s12872-018-0815-3

[11] Kim, C., Redberg, R. F., Pavlic, T., & Eagle, K. A. (2007). A systematic review of gender differences in mortality after coronary artery bypass graft surgery and percutaneous coronary interventions. In Clinical Cardiology (Vol. 30, Issue 10, pp. 491–495). https://doi.org/10.1002/clc.20000

[12] di Giosia, P., Passacquale, G., Petrarca, M., Giorgini, P., Marra, A. M., & Ferro, A. (2017). Gender differences in cardiovascular prophylaxis: Focus on antiplatelet treatment. In Pharmacological Research (Vol. 119, pp. 36–47). Academic Press. https://doi.org/10.1016/j.phrs.2017.01.025

[13] Nashif, S., Raihan, Md. R., Islam, Md. R., & Imam, M. H. (2018). Heart Disease Detection by Using Machine Learning Algorithms and a Real-Time Cardiovascular Health Monitoring System. World Journal of Engineering and Technology, 06(04), 854–873. https://doi.org/10.4236/wjet.2018.64057

[14] Institute of Electrical and Electronics Engineers. (n.d.). 2019 2nd International Conference on New Trends in Computing Sciences (ICTCS) : proceedings : Amman, Jordan, 9-11 October 2019.

[15] Garate Escamilla, A. K., Hajjam El Hassani, A., & Andres, E. (2019). A Comparison of Machine Learning Techniques to Predict the Risk of Heart Failure (pp. 9–26). https://doi.org/10.1007/978-3-030-15628-2_2

[16] SKR Engineering College, Institute of Electrical and Electronics Engineers. Madras Section, & Institute of Electrical and Electronics Engineers. (n.d.). International Conference on Energy, Communication, Data Analytics & Soft Computing (ICECDS) - 2017 : 1st & 2nd August 2017.

[17] Gupta, A., Kumar, R., Singh Arora, H., & Raman, B. (2020). MIFH: A Machine Intelligence Framework for Heart Disease Diagnosis. IEEE Access, 8, 14659–14674. https://doi.org/10.1109/ACCESS.2019.2962755

[18] Gowsalya, M. (2020). Heart Disease Detection Using Machine Learning Chithambaram T Vellore institute of technology and Logesh Kannan N Vellore institute of technology. https://doi.org/10.21203/rs.3.rs-97004/v1

[19] Dayananda Sagar College of Engineering, Institute of Electrical and Electronics Engineers. Bangalore Section, & Institute of Electrical and Electronics Engineers. (n.d.). 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA 2020) : conference proceedings : 5-7 March, 2020.

[20] Bharti, R., Khamparia, A., Shabaz, M., Dhiman, G., Pande, S., & Singh, P. (2021). Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning. Computational Intelligence and Neuroscience, 2021. https://doi.org/10.1155/2021/8387680

[21] Guidi, G., Pettenati, M. C., Melillo, P., & Iadanza, E. (2014). A machine learning system to improve heart failure patient assistance. IEEE Journal of Biomedical and Health Informatics, 18(6), 1750–1756. https://doi.org/10.1109/JBHI.2014.2337752

[22] Diwakar, M., Tripathi, A., Joshi, K., Memoria, M., Singh, P., & Kumar, N. (2020). Latest trends on heart disease prediction using machine learning and image fusion. Materials Today: Proceedings, 37(Part 2), 3213– 3218. https://doi.org/10.1016/j.matpr.2020.09.078

[23] Alalawi, H. H., & Alsuwat, M. S. (2021). Detection of Cardiovascular Disease using Machine Learning Classification Models. www.ijert.org

[24] Alqahtani, A., Alsubai, S., Sha, M., Vilcekova, L., & Javed, T. (2022). Cardiovascular Disease Detection using Ensemble Learning. Computational IntelligenceandNeuroscience, 2022.

[25] Rani, P., Kumar, R., Ahmed, N. M. O. S., & Jain, A. (2021). A decision support system for heart disease prediction based upon machine learning. Journal of Reliable Intelligent Environments, 7(3), 263–275. https://doi.org/10.1007/s40860-021-00133-6

[26] Shukur, B. S., & Mijwil, M. M. (2023). Involving machine learning techniques in heart disease diagnosis: a performance analysis. International Journal of Electrical and Computer Engineering, 13(2), 2177–2185. https://doi.org/10.11591/ijece.v13i2.pp2177-2185.

ABOUT THE AUTHORS

Dr. Nishant Kumar Pathak boasts 13 years of teaching experience and has published 28 research papers in national and international journals. His specialization areas include software engineering, wireless sensor networks, artificial intelligence, and data science.

Mr Varun Kumar heaving 20 years of teaching Experience .Published 5 Research Papers National and International Journal. His Specialization area is wireless sensor networks/ Artificial intelligence/Data science.

Mr Ashish Kumar heaving 10 years of teaching Experience. He has Published 2 Research Papers. His Subject Specialization area is Networking,Artificial intelligence and Cloud Computing.