
A Comparative Analysis of Cardiovascular Disease using Different Machine Learning Techniques

Mukesh Pandey¹, Satish Kumar Mahariya², Dr. Shiv Dayal Pandey³, Dr. Kapil Joshi⁴, Dev Baloni⁵, Jefferson A. Costales⁶
Assistant Professor Uttaranchal University Dehradun^{1,2,3,4,5}, Eulogio "Amang" Rodriguez Institute of Science and Technology,
Manila, Philippines⁶

sitmukesh@gmail.com¹, satishmaharia191095@gmail.com², sdpmmnit@gmail.com³, kapilengg0509@gmail.com⁴,
devbalini1982@gmail.com⁵, jacostales@earist.ph.education

Abstract. Every year, around 18.6 million people die from cardiovascular diseases, the most common of which are myocardial infarctions and heart failure. MI and heart attacks are caused by persistent ischemia and hypoxia, which causes irreparable damage to myocardial tissue. An Effective, accurate, and feasible method for the timely diagnosis of these diseases is therefore needed for proper treatment. Heart disease early state can be identified with the help of multiple methods to use Several data mining, machine learning, deep learning algorithms Techniques. To simplify broad and complicated data analysis, the different algorithms models and methods of Artificial Intelligence can be implemented to different healthcare datasets. Indeed, this review's Details Analysis will be required to provide a methodological framework for the development of additional effective prediction models. The purpose of this paper is to review some of the recent research on cardiovascular disease prediction using various learning techniques. The goal of this study is to illustrate different algorithms of machine learning for cardiovascular disease prediction and find the best possible learning technique to identify the particular objective of cardiovascular disease.

Keywords: Machine Learning, heart disease, accuracy, prediction. Data Mining, Artificial Neural Network

1. INTRODUCTION

Roughly 18.6 million people die each year from cardiac disease. Heart performs the most important part of bloodstream. Under continuous, diligent treatment, diabetes may prompt a collection of blood sugars, that could raise the danger of hazardous inconveniences including cardiovascular sicknesses. About 47% of unexpected deaths occur outside of an emergency department. Which suggests that several cardiovascular disease sufferers do not react to early warning signs. Heart problems have some primary signs and symptoms: chest pain and discomfort. Awe-inspiration. Nausea, flush or cold sweat. Installation's store enormous quantities of information in their directories, which is very complicated and demanding, due to the rapidly growing digital technologies. The techniques and algorithms on a dataset can be used directly to create specific models and identify important conclusions. Machine learning methods are therefore useful in predicting performance from actual data. Therefore, the whole work relates one such machine learning method, called classification, to predict risk from heart disease risk factors. This also seeks to increase the efficiency of cardiac risk assessment using a technique called an ensemble.

2. LITERATURE REVIEW

Numerous studies on disease prediction systems have been performed using various algorithms methods of Artificial Intelligence for the different kind of data sets

K. Polaraju et al, [7] Recommended cardiovascular infection study utilizing Multiple Clustering calculation and it exhibits that different straight relapse is suitable for foreseeing coronary illness chance. The analyses are carried out using the training data set, which contains 3000 occurrences of each of the 13 distinct attributes stated above. The data collection is divided into two areas, with 70% of the information used for planning and 30% for research. In light of the test, unmistakably the grouping precision of the Regression calculation is higher than some other AI calculations.

Marjia et al, [8] Expectation of coronary illness created utilizing WEKA programming, utilizing impression of Bayes Net, Multilayer K Star, j48 and SMO. The exhibition from one of a kind factor Bayes Net or SMO accomplishes ideal productivity than multilayer vision, KStar, and J48 procedures use k-fold cross approval. The exactness precision got by these calculations is as yet unsuitable. Thus, the proficiency of the precision is upgraded more to give better prescient ailment end.

Purushottam et al, [15] Suggested an effective system for predicting heart disease. This system will help health professionals make meaningful outcome-making dependent on the factor in issue.

S. Seema et al,[9] Focused on methods capable of detecting severe diseases by mining historical health information. A comprehensive algorithm analysis is conducted to test the increased output at an appropriate rate. SVM gives this investigation the most noteworthy exactness rating while Naïve Bayes offers the least accuracy.

Megha Shahi et al, [11] Emerging cardiovascular illness prescient framework using information mining systems; WEKA is utilized for programmed sicknessdistinguishing proof and administration quality conveyance in medicinal services communities The examination utilized different calculations, for example, Naïve Bayes, KNN, Alliance law , ANN, SVM and Tree of Judgment. The examination recommended by SVM is effective and gives more noteworthy accuracy contrasted with anotherparameters.

Chala Beyene et al, [12] Supported Cardiac Disease Detection and Evaluation can be done by Data Mining Techniques. The primary objective is to identify the incidence of cardiovascular disease within a short period of time. The method suggested is also important in many organizations of healthcare who have no more knowledge and skills It offers different clinical highlights, for example age, sex, pulse are a portion of the traits included for deciding if the individual has illness or not. Dataset analysis is computed using WEKAtools.

R. Sharmila et al, [13] Suggested usage of classification algorithm for the analysis of dynamical cardiac disease. It is advised to use bigdata resources including Distributed File System of Hadoop, Map reduction alongside expectation of coronary illness with optimized attribute selection.This research conducted an inquiry into the use of various data mining techniques to predict heart disease It recommends using HDFS to store big data in multiple regions and continuously execute.

Sharma Purushottam et al, [15] New rules c45 and partial tree prediction methods for cardiovascular risk. This research will uncover rules to estimate patients ' risk levels based on their health factor. Output can be measured in measurements of classification of performance, classification of failures, rules generated and the performance. And analysis was performed using partial tree and C4.5. The outcome reveals that prediction is potential and more efficient.

P.Sai Chandrasekhar Reddy et al, [17] Suggested cardiovascular risk prediction taking data mining method ANN. As the expense of diagnosing patients with cardiac failure grew, new technologies needs to be established which can diagnose heart attacks.After measurement, predictor method is being used to evaluate the wellbeing of the patient dependent on various factors, The system is known to be powerful in java.

S. Prabhavathi et al, [23] Suggested Neural Fuzzy System (DNFS) tree-based decision-making technique for study and prediction of various heart diseases. This report discusses work into the treatment of cardiovascular disease. DNFS stands for Decision Neural Fuzzy System based on a Decision Tree. This work is aimed at creating a smart and value-effective system, as well as improving the current system efficiency. In this report in particular, Data mining techniques are being utilised to enhance cardiovascular disease prediction. The conclusion of this study indicates that the neural networks and SVM have a positive way of predicting heart problems. Even data mining tools are ineffective in predicting heartillness.

Sairabi H.Mujawar et al, [24] K-means and naïve coves used for heart attack prediction. This work is conceived to develop the program utilizing a prior heart dataset that offers care. Thirteen qualities were considered for product design. Techniques for data processing, including clustering, Database details may be obtained from the classificationprocesses. Cleveland Heart Registry used 13 attributes for 300 records in total. This system is built to assess how often the individual has heart failure or depending on the 13 values of the attributes.

Sharan Monica.L et al, [25] The study of heart attack has been suggested. This research suggested methods for predicting the disease using data mining. It's also intended to choose the analysis of existing methods to retrieve data from the database,

and will be useful to health care professionals. The output can be achieved depending on the time it takes for the program to construct the decision tree.

B Gnaneswar et al,[27] explored numerous simple and hybrid data mining methods, which were analysed to determine which one is best for accurately predicting heart disease. They demonstrated the merits and limitations of existing systems such as the (SVM), decision trees Naive Bayes methods, , K-Means algorithms, and Neural Networks

G. Shanmugasundaram et al,[29] looked into the many aspects and their importance in detecting cardiac disorders. They discovered that several predicted features were not taken into consideration by the researchers, and that key aspects such as drinking and smoking could have been added. They came to the conclusion that lowering the number of variables did not accurately predict cardiovascular diseases in order to increase accuracy. This statement might not be true if the reduced qualities have no bearing on the forecast. What matters here is the significance of the traits picked, not their quantity

3. OBSERVATION OBTAINED ABOUT CVD

Table I shows the various kind of coronary illness and their description and Table II shows the observations were analyzed and summarized in terms of technology and Accuracy rate.

Table 1 Various kinds of Coronary Illness

Illness	Description
Stroke	A stroke is a neurological disorder where insufficient blood circulation through the brain leading to cell death
Cardiac arrest	Cardiac arrest is a whole cessation of blood deliver springing up from the heart's incapability to successfully pump. Signs consist of loss of manage and anomalous or incomplete respiration
Arrhythmia	A Problem with the charge or temper of your pulse. It means that your coronary heart pulsates too rapidly, too gradually, or with a sporadic example. At the factor whilst the coronary heart pulsates faster than ordinary, it's far referred to as tachycardia.
High Blood Pressure	This will happen when your strain augmentations to disastrous levels. In this circulatory process strain estimation thinks about how much blood is experiencing your veins and the proportion of check the blood meets while the heart is siphoning. Confined veins increase deterrent.
Coronary artery disease	Coronary artery disorder is coronary artery narrowing or blockage, typically triggered by atherosclerosis. The accumulation of cholesterol and fatty compounds (called plaques) on lies inner walls of the arteries is atherosclerosis (sometimes named "hardening" or "clogging" of arteries).
Peripheral artery disease	It develops in the arteries that bring blood to the brain, lungs, and limbs. Plaque consists of food, glucose, magnesium, fibrous tissue, and other blood-borne compounds.
Congestive heart failure	Congestive cardiovascular breakdown (CHF) is a relentless powerful condition that impacts the siphoning force of your heart muscles. While routinely insinuated similarly as "cardiovascular breakdown," CHF expressly suggests the stage where fluid creates around the heart and makes it siphon inefficiently. We have four heart chambers.
Congenital heart disease	Congenital cardiac disorder (congenital heart defect) is one or more defects within the anatomy of the heart from which you are born.

TABLE II A STUDY OF DIFFERENT TECHNIQUES USED FOR CVD

S.NO	AUTHOR	OBJECTIVE	TECHNIQUES USED	ACCURACY/CONCLUSION
1	Noura Ajam et al	Artificial Neural Network is Used for Predicting Heart Attack.	Artificial Neural Network	By using ANN we get accuracy of 88%.
2	Sharma Purushottam et al,	Evaluation of the heart attack prediction method using C4.5 principles and a partial list.	Naive Bayes algorithm and C4.5 rules	Naive Bayes gives less accuracy as compare to C4.5 in predicting heart attack.

3	S. Seema et al,	Data mining is being used to detect chronic diseases.	Naïve Bayes SVM	Naïve Bayes ' average precision in the case of heart attack is 91.56% Best precision reached by SVM was 95.56% for heart disease
4	Ashok al,Kumar Dwivedi et al,	Efficiency evaluation of various machine learning methods for detecting heart attack	Logistic Regression Naïve Bayes KNN Classification Tree	85% 83% 80% 77%
5	K. Gomathi et al,	Data processing techniques for multi-disorder analysis.	Naïve Bayes	Detect Breast Cancer: 82.5% Diabetes: 77.6% Detect Heart Disease: 79%
			J48	Diabetes: 100% Cardiac disease: 77% Detect Breast cancer: 75.5%
6	Jayamin Patel et al,	Forecasting heart attack utilizing the computer learning and data mining methodologies.	Random Forest Algorithm	LMT gives us 55.75% of accuracy while J48 give us 56.76% accuracy which is better than LMT.
7	Ashwini Shetty A et al,	Similar approaches to data analysis for forecasting heart attack.	Neural Network	By using WEKA tool and Matlab accuracy got 88%
8	Prajakta Ghadge et al,	Intelligent program for forecasting cardiac failure using big data	Naïve Bayes.	The computerization of this system makes very favorable
9	S. Prabhavathi et al,	Comparative coronary illness conclusion and expectation utilizing DNFS methods .	Naïve Bayes, Decision Tree	Accurate similar to the forms of heart disorder. Treatment of CVD= around 85% to 99%. CHD Diagnosis= from 82 to 92 %
10	Sharan Monica. L et al,	Mining Techniques prediction analysis for Cardio-Vascular	Simple CART	92.2%
			J48	91.1%
			Naïve Bayes	88.5%
11	Jayami Patel et al,	Mining and ML techniques for predicting heart attacks	LMT, UCI	UCI, compared with LMT, gives better precision.
12	P. Sai Chandrasekhar Reddy et al,	The use of ANN and data mining algorithms to forecast cardiac cases	ANN	The accuracy has been proved in java
13	Chala Bayen et al,	A Review to predict the incidence of heart disease	SVM, Naïve Bayes, J48,	It offers short-term result that helps to deliver quality programs and lower individual costs.

14	R. Sharmila et al,	A statistical method to boost diagnosis computer techniques for cardiac attack.	Parallel Fashion Support Vector Machine	Support Vector Machine offers 82.35% better accuracy. SVM provides greater accuracy in parallel mode than sequential SVM .
15	S. Carolin Jeeva et al,	Ensemble Classification techniques to predict risk-based heart disease and also improving the accuracy	ensemble techniques	By using ensemble techniques, using ensemble classification, a cumulative improvement of 7% accuracy was obtained for poor classifiers.
16	Aakash Chauhan et al,	Determine the criteria for predicting coronary artery disease.	Rule mining algorithm	Rule Mining improves results based on prediction techniques
17	Chaitanya Suvarna et at,	Data mining and optimization approaches were used to create a prediction algorithm for cardiac disease.	particle swarm optimization technique (PSO)	Data mining methods based on particle swarms successfully forecast cardiac disease.
18	Balasaheb Tarle et al,	Propose a medical data classification model for cardiac disease	Artificial Neural Network	Classification accuracy is 83 percent when cross-validation and ANNs are used together.
19	Sonam Nikhar et al,	classifier of decision tree and Naive Bayes approach are discussed to predict cardiac disease	Decision Tree, Naive Bayesian Classifier	The DTC is more precise than the Naive Bayes classifier.

4. CONCLUSION AND FUTURE SCOPE

A study is guided in this article on the many ways proposed to predict cardiovascular illnesses. A total of 19 attributes were used. The goals, data source, objectives, and outcomes of each strategy were all presented. The results were discussed, and it was highlighted to different techniques of machine Learning and Data Mining were used for these approaches as well as attributes. In the future, we will conduct a more in-depth investigation of these attributes that are thought to be possible threat factors in order to compile a comprehensive list of applicable attributes that can be used to forecast cardiovascular illnesses and then explore for probable links between them. As a result of this research, we will be able to present a framework for developing a predictive model..

Reference

1. PreetiGulia, SumitaSihag.**Enhance Security in MANET using Bacterial Foraging Optimization Algorithm. International Journal of Computer Applications.**2013 December, 84(1), pp. 32-35 (International Journal of Computer Applications (0975 – 8887)) .
2. Animesh Hazra, Arkomita Mukherjee, Amit Gupta, Asmita Mukherjee, “Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review”, Research Gate Publications, July 2017, pp.2137-2159.
3. V. Krishnaiah, G. Narsimha, N. Subhash Chandra, “Heart Disease Prediction System using Data Mining Techniques and Intelligent Fuzzy Approach: A Review”,International Journal of Computer Applications, February 2016
4. Guizhou Hu, Martin M. Root, “Building Prediction Models for Coronary Heart Disease by Synthesizing Multiple Longitudinal Research Findings”, European Science of Cardiology, 10 May 2005.
5. T.Mythili, Dev Mukherji, Nikita Padaila and Abhiram Naidu, “A Heart Disease Prediction Model using SVM- Decision Trees- Logistic Regression (SDL)”, International Journal of Computer Applications, vol. 68, 16 April 2013.
6. Devangi Kotak “Data Mining Classification Techniques for Cardiovascular Disease Prediction”I nternational Journal of Engineering Research&Technology (IJERT) ISSN: 2278-0181 IJERTV8IS110412

7. Nimai Chand Das Adhikari, Arpana Alka, and rajat Garg, "HPPS: Heart Problem Prediction System using Machine Learning".
8. K. Polaraju, D. Durga Prasad, "Prediction of Heart Disease using Multiple Linear Regression Model", International Journal of Engineering Development and Research Development, ISSN:2321-9939, 2017.
9. Marjia Sultana, Afrin Haider, "Heart Disease Prediction using WEKA tool and 10-Fold cross-validation", The Institute of Electrical and Electronics Engineers, March 2017.
10. Dr.S.Seema Shedole, Kumari Deepika, "Predictive analytics to prevent and control chronic disease", January 2016.
11. Ashok kumar Dwivedi, "Evaluate the performance of different machine learning techniques for prediction of heart disease using ten-fold cross-validation", Springer, 17 September 2016.
12. Megha Shahi, R. Kaur Gurm, "Heart Disease Prediction System using Data Mining Techniques", Orient J. Computer Science Technology, vol.6 2017, pp.457- 466.
13. Mr. Chala Beyene, Prof. Pooja Kamat, "Survey on Prediction and Analysis the Occurrence of Heart Disease Using Data Mining Techniques", International Journal of Pure and Applied Mathematics, 2018.
14. R. Sharmila, S. Chellammal, "A conceptual method to enhance the prediction of heart diseases using the data techniques", International Journal of Computer Science and Engineering, May 2018.
15. Jayami Patel, Prof. Tejal Upadhay, Dr. Samir Patel, "Heart disease Prediction using Machine Learning and Data mining Technique", March 2017.
16. Purushottam, Prof. (Dr.) Kanak Saxena, Richa Sharma, "Efficient Heart Disease Prediction System", 2016, pp.962-969.
17. K.Gomathi, Dr.D.Shanmuga Priyaa, "Multi Disease Prediction using Data Mining Techniques", International Journal of System and Software Engineering, December 2016, pp.12-14.
18. Mr.P.Sai Chandrasekhar Reddy, Mr.Puneet Palagi, S.Jaya, "Heart Disease Prediction using ANN Algorithm in Data Mining", International Journal of Computer Science and Mobile Computing, April 2017, pp.168- 172.
19. Ashwini Shetty A, Chandra Naik, "Different Data Mining Approaches for Predicting Heart Disease", International Journal of Innovative in Science Engineering and Technology, Vol.5, May 2016, pp.277- 281.
20. Jaymin Patel, Prof. Tejal Upadhyay, Dr.Samir Patel, "Heart Disease Prediction using Machine Learning and Data Mining Technique", International Journal of Computer Science and Communication, vol.7, September 2015- March 2016, pp.129-137.
21. Boshra Brahmi, Mirsaeid Hosseini Shirvani, "Prediction and Diagnosis of Heart Disease by Data Mining Techniques", Journals of Multidisciplinary Engineering Science and Technology, vol.2, 2 February 2015, pp.164- 168.
22. Noura Ajam, "Heart Disease Diagnoses using Artificial Neural Network", The International Insitute of Science, Technology and Education, vol.5, No.4, 2015, pp.7-11.
23. Prajakta Ghadge, Vrushali Girme, Kajal Kokane, Prajakta Deshmukh, "Intelligent Heart Disease Prediction System using Big Data", International Journal of Recent Research in Mathematics Computer Science and Information Technology, vol.2, October 2015 - March 2016, pp.73- 77.
24. S.Prabhavathi, D.M.Chitra, "Analysis and Prediction of Various Heart Diseases using DNFS Techniques", International Journal of Innovations in Scientific and Engineering Research, vol.2, 1, January 2016, pp.1-7.
25. Sairabi H.Mujawar, P.R.Devale, "Prediction of Heart Disease using Modified K-means and by using Naïve Bayes", International Journal of Innovative research in Computer and Communication Engineering, vol.3, October 2015, pp.10265-10273

26. Sharan Monica L, and Satheesh Kumar B (2016), "Analysis of Cardiovascular Heart Disease Prediction Using Data Mining Techniques", International Journal of Modern Computer Science (IJMCS), ISSN: 2320-7868(Online), Vol. 4, Issue 1, February 2016, pp. 55-58.
27. C. Beulah Christalin Latha, S. Carolin Jeeva "Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques" Informatics in Medicine Unlocked 16 (2019) 100203
28. G. B et E. J. M.R, "A Review On Prediction And Diagnosis Of Heart Failure," International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2017.
29. J. Maiga, G. G. Hungilo and Pranowo, "Comparison of Machine Learning Models in Prediction of Cardiovascular Disease Using Health Record Data," 2019 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS), 2019, pp. 45-48, doi: 10.1109/ICIMCIS48181.2019.8985205.
30. S. G, M. S. V, R. Saravanan et S. Balaji, "An Investigation of Heart Disease Prediction Techniques,"IEEE, vol. 10, 2018
31. M. A. P et S. P. Shaji, "Prediction and Diagnosis of Heart Disease Patients using Data Mining Technique," International Conference on Communication and Signal Processing, 2019.
32. S. N. Induja et C. G. Raji, "Computational Methods for Predicting Chronic Disease in Healthcare Communities,"chez International Conference on Data Science and Communication (IconDSC), 2019
33. A. Gavhane, I. Pandya, G. Kokkula et P. K. Devadkar, "Prediction of Heart Disease Using Machine Learning," chez ProInternational conference on Electronics, Communication and Aerospace Technology, 2018.
34. A. Rairikar, V. Kulkarni, V. Sabale, H. Kale et A. Lamgunde, "HEART DISEASE PREDICTION USING DATA MINING TECHNIQUES," chez International Conference on Intelligent Computing and Control, 2017.
- 35.



Mukesh Pandey is an Assistant Professor in CSE Department at Uttarakhand University, Dehradun. He is SimpliLearn Certified Trainer in Machine Learning. He holds B.Tech in Information Technology and an M.Tech in Computer Science. He is also doing his Ph.D in Computer Science from UTU, Dehradun. He has more than 9.8-year Experience in Government as well as Private Engineering Institutes in Academics'. He has a good number of Research Papers in Reputed Journals and Conferences.