ISSN (Online): 2583-7052

International Journal of Advanced Trends in Engineering and Management (IJATEM)
pp. 19-25

Article Title: **Effectiveness Of Machine Learning Algorithms In The Diagnosis Of Cardiovascular Disorders**

Effectiveness of Machine Learning Algorithms in the Diagnosis of Cardiovascular Disorders

T. Benila Christabel[1], K. K. Thanammal[2]

[1] Research Scholar, Department of Computer Science and Research Centre, S.T. Hindu College, Tamilnadu, India. benila.christabel88@gmail.com

[2] Assistant Professor, Department of Computer Science and Research Centre, S.T. Hindu College, Nagercoil, Tamilnadu, India. thanaravindran@gmail.com

ABSTRACT

In the entire world, cardiovascular diseases (CVDs) are the main cause of death. According to estimates, 17.9 million deaths worldwide in 2019 were attributable to CVDs, or 32% of all fatalities. Heart attack and stroke deaths accounted for 85% of these fatalities. The majority of CVD fatalities occur in low- and middle-income nations. In 2019, non communicable illnesses caused 17 million premature deaths (before the age of 70), and 38% of those fatalities were attributable to CVDs. By addressing behavioural risk factors like tobacco use, unhealthy eating and obesity, inactivity and problematic alcohol consumption, the majority of cardiovascular illnesses can be avoided. Early detection of cardiovascular disease is crucial in order to start treatment with counseling and medication. This study examines the effectiveness of machine learning algorithms in the diagnosis of cardiovascular diseases.

Keywords: Machine learning, Cardiovascular disease, Random forest, Support Vector Machines, K Nearest Neighbor.

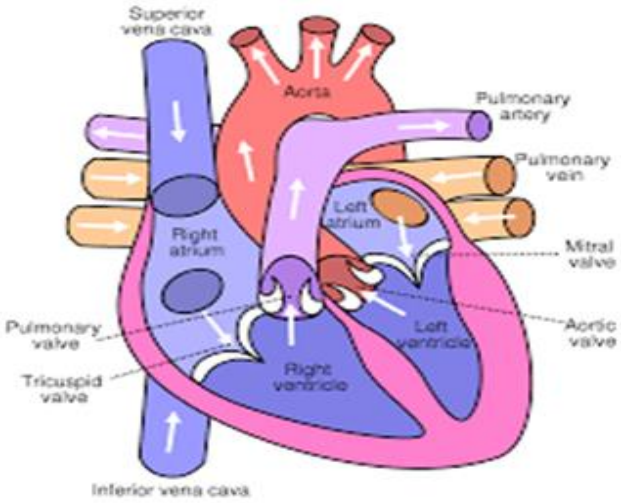
1 Introduction

Cardiovascular diseases (CVDs) refer to a variety of heart and blood vessel conditions. They consist of the following diseases. An illness of the blood arteries supplying the heart muscle is coronary heart disease, while a condition of the blood vessels supplying the brain is cerebrovascular disease. A condition of the blood arteries supplying the arms and legs known as peripheral arterial disease. Rheumatic heart disease is caused by streptococcal bacteria, which also cause rheumatic fever, which damages the heart muscle and heart valves. Deep vein thrombosis and pulmonary embolism are blood clots in the leg veins that can become dislodged and travel to the heart and lungs. Congenital heart disease is a birth abnormality that affects the normal development and functioning of the heart.

ISSN (Online): 2583-7052

International Journal of Advanced Trends in Engineering and Management (IJATEM)
pp. 19-25

Article Title: Effectiveness Of Machine Learning Algorithms In The Diagnosis Of Cardiovascular Disorders



2 Current System

To discover cardiac disease among different age groups, current systems entail doing numerous tests and applying algorithms, questionnaires, and surveys. The doctor's clinical judgement and the patient's medical history solely determine the patient's diagnosis. Cardio-vascular disease predictions and diagnoses are becoming more difficult for medical professionals and institutions to make systems that use algorithms for detection and classification do not offer adequate tests for diagnosing disease. The collection of parameters used to detect cardio vascular disease do not yield good findings and are insufficient to draw any conclusions because of information gaps that could lead to wrong conclusions.

3 Proposed System

Data mining and machine learning techniques are crucial for disease prediction. Cardio-vascular disease is detected using machine learning algorithms like K Nearest Neighbour and Artificial Neural Networks (ANN). Researchers are working quickly to create software that will aid doctors in making decisions regarding the prediction and diagnosis of cardiac disease. This software will use a machine learning algorithm.

3.1 Algorithms and Techniques Used

3.1.1 Random Forest

It is a notion included in the broad category of random decision-making [3]. This technique works by combining decision trees that indicate the mode of classes or the mean forecasts of a single tree based on training duration and class. Trees cast votes during classifications, and a highly well-known class is returned. By averaging the deep decision trees trained with various subsets [7] of the training data, the variance can be decreased using the Random forests. To create random forests, tree predictors must be combined so that each tree depends equally and

ISSN (Online): 2583-7052

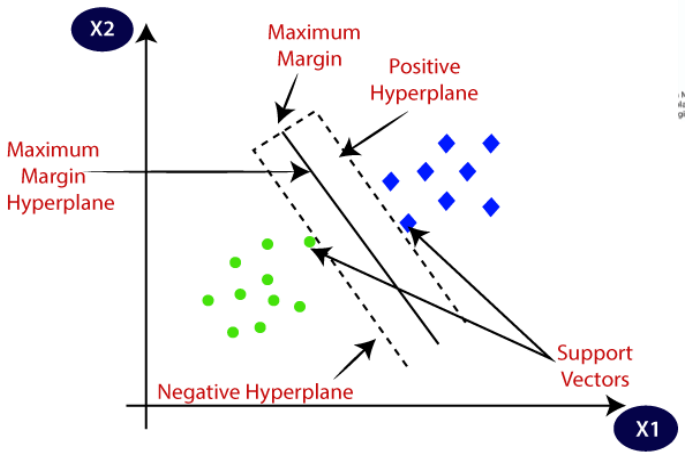
International Journal of Advanced Trends in Engineering and Management (IJATEM)
pp. 19-25

Article Title: Effectiveness Of Machine Learning Algorithms In The Diagnosis Of Cardiovascular Disorders

uniformly on the values of a random vector [5]. This method is how we search our database for delays. Random Forest is another well-liked supervised machine learning algorithm concept used for classification and regression tasks, however it often excels at classification tasks. [6] As the name implies, Random Forest approaches take into account numerous decision trees that provide results [4].

3.1.2 Support Vector Machines

It exists in both linear and non-linear forms. This classifier is supervised. SVM in this situation often uses two different datasets: a training set and a test set [2]. The classes become linearly separable at the ideal time. In these circumstances, a line can be established, precisely dividing classes. However, it doesn't divide lines in datasets in a systematic way; instead, many lines do [4]. The "separating line" is chosen as the best of these lines.



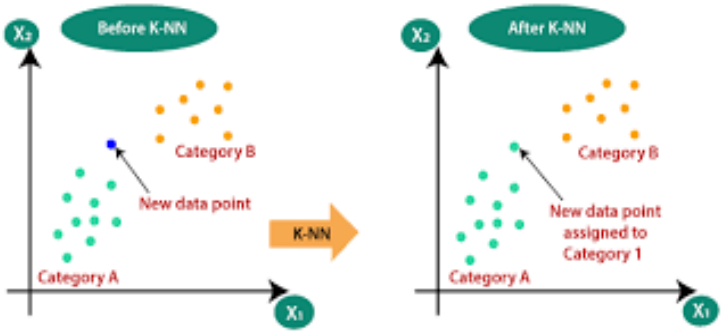
ISSN (Online): 2583-7052

International Journal of Advanced Trends in Engineering and Management (IJATEM)
pp. 19-25

Article Title: Effectiveness Of Machine Learning Algorithms In The Diagnosis Of Cardiovascular Disorders

3.1.3 K Nearest Neighbor

K nearest neighbours is a well-known classification technique due to its accuracy and simplicity. Because it stores every available case and categorises current cases based on similarity metrics, it is a straightforward algorithm. This method assumes that things that are similar to one another exist nearby. Which are similarly existing together as things. KNN has a K and is a collection of nearby neighbours' numbers. As a decision-factor for neighbours in the cores. K has typically remained an unusual statistic, with 2 classes. Or In simplest circumstances, these algorithms are referred to as nearest-neighbor algorithms when K=1.



4 Methodology

The core components of the proposed systems are classification components based on K Nearest Neighbors, Random Forest, and Artificial Neural Network, [4]feature selections and collections are based on Fasts Correlation-Based Feature Selection (FCBF). A pre-clipped training dataset is used to render a binary class classification issue.Features are chosen from the training dataset, then the best feature subsets are enhanced using our combined PSO/ACO algorithm, and finally [2] by choosing the best feature, new features are classified using WEKA[17]. Software for data mining is used. The modules of our suggested system are defining the specifics of the following subunits.

5 Data Collection

A database has a total of 76 properties, however all of them are being used in the experiment by employing subsets of 14.The only database utilising ML academic to the dates is a Clevelands database. The "goals" entry mentions a patient's presence of cardio-vascular disorders, with integer values ranging from 0 No Present and 1 present.

5.1 Testing and Results

Test and effects analyzing the characteristics that are shown below, a prediction of cardio-vascular disorders is made after assigning inputs and receiving correct results predict an output based on desired output contribution norms.

ISSN (Online): 2583-7052

International Journal of Advanced Trends in Engineering and Management (IJATEM)
pp. 19-25

Article Title: Effectiveness Of Machine Learning Algorithms In The Diagnosis Of Cardiovascular Disorders

Feature No	Feature Name
1	age
2	sex
3	cp
4	trest
5	chol
6	fbs

Input Values

1. age: age (from 29 to 77 years old)
2. SEX: (Male = 1; Female = 0)
3. CP: kind of chest pain (0 to 3)
4. TREST: (94 to 22) mm Hg of resting blood pressure at hospital admission
5. CHOL:Cholesterols' (126 to 564)mg serum
6. Fasting Blood Sugar > 120 (mg).1=True,0=False

Output

Either 1 or 0.

1 = Indicates Cardiovascular Disease is detected.

0 = No disease has been identified.

All machine learning algorithms use input attributes from the classification process to predict this attribute. The values of this property are split into two groups: values 0 indicate that the patient has no cardiovascular disease, and values 1 through 4 are regarded as values 1, which indicate the presence of heart disease.

6 Conclusion

This study offers a thorough understanding of machine learning techniques for categorising cardiovascular disorders [3]. The use of classifiers is essential in the health care sector so that the outcomes may be utilized to foretell the treatments that patients will get. To identify the most accurate and efficient systems, the current approaches are examined and contrasted. Machine learning approaches increase the precision of heart disease risk detection, allowing patients to benefit from treatment in the early stages of the condition [2]. Machine learning algorithms have enormous potential for predicting heart-related or cardiovascular disorders [5]. Each of the aforementioned algorithms has done well in testing.

ISSN (Online): 2583-7052

International Journal of Advanced Trends in Engineering and Management (IJATEM)
pp. 19-25

Article Title: Effectiveness Of Machine Learning Algorithms In The Diagnosis Of Cardiovascular Disorders

References

1. B. L. Deekshatulu; Priti Chandra, Year: 2017, “Classification of Heart Disease Using K- Nearest Neighbor and Genetic Algorithm”, M.Akhil jabbar* International Conference on Computational Intelligence: Modeling Techniques and Appli- cations 2017.

2. Nura Esfandiari; Mohammad Reza Babavalian; Amir-Masoud Eftekhari Moghadam; Vahid Kashani Tabar, Year: 2017, “Knowledge discovery in medicine: Current issue and future trend”, Expert Syst. 2017.

3. Michael W. Berry; Murray Browne, Year: 2017, “Lecture notes in data mining”, World Scientific.

4. Shilaskar Swati; Ashok Ghatol, Year: 2018, “Feature selection for medical diagnosis Evaluation for cardio vascular diseases”, Expert Syst. Appl., Vol: 40.

5. T. H. M. Prerana; N. C. Shivaprakash; N. Swetha, Year: 2018, “Prediction of Heart Disease Using Machine Learning Algorithms- Na¨ıve Bayes, Introduction to PAC Algorithm, Comparison of Algorithms and HDPS”, International Jour- nal of Science and Engineering, Vol: 23.

6. Chun-Lang Chang; Chih-Hao Chen, Year: 2016, “Applying decision tree and neural network to increase quality of dermatologic diagnosis”, Expert Syst. Appl.

7. Ahmad Taher Azar; Shereen M. El-Metwally, Year: 2013, “Decision tree classifiers for automated medical diagnosis”, Neural Compute Appl., Vol: 23, no: 7 - 8, pp. 2387 – 2403.

8. Bo Jin; Y. C. Tang; Yan-Qing Zhang, Year: 2016, “Support vector machines with genetic fuzzy feature transformation for biomedical data classification,”.

9. Moloud Abdar; Mariam Zomorodi-Moghadam; Resul Das; I-Hsien Ting, Year: 2017, “Performance analysis of classification algorithms on early detection of liver disease”, Expert Syst Appl., Vol: 67, pp. 239 – 51.

10. Moloud Abdar; Neil Yuwen Yen; Jason Chi-Shun Hung, Year: 2017, “Improving the diagnosis of liver disease using multilayer perceptron neural network and boosted decision trees”, J Med Biol Eng., Vol: 10, pp. 1 – 13.

11. Paweł Pławiak, Year: 2018, “Novel genetic ensembles of classifiers applied to myocardium dysfunction recognition based on ECG signals, Swarm”, Evol Comput., Vol: 39, pp. 192 – 208.

12. Paweł Pławiak, Year: 2018, “Novel methodology of cardiac health recognition based on ECG signals and evolutionary-neural system”, Expert Syst Appl., Vol: 92, pp. 334 – 49.

13. Fahime Khozeimeh; Roohallah Alizadehsani; Mohamad Roshanzamir; Abbas Khosravi; Pouran Layegh; Saeid Nahavandi, Year: 2017, “An expert system for selecting wart treatment method”, Comput Biol Med., Vol: 81, no: 6, pp. 167 – 75.

14. Fahime Khozeimeh; Farahzad Jabbari Azad; Yaghoub Mahboubi Oskouei; Majid Safari; Shahrzad Tehranian; Roohallah Alizadehsani; Pouran Layegh, Year: 2017, "Intralesional immunotherapy compared to cryotherapy in the treatment of warts", Int J Dermatology., Vol: 56, pp. 474 – 8.

15. Roohallah Alizadehsani; Moloud Abdar; Seyed Mohammad Jafar Jalali; Mohamad Roshanzamir, Abbas Khosravi; Saeid Nahavandi, Year: 2018, "Comparing the performance of feature selection algorithms for wart treatment selection", Proc. Int. Workshop Future Technol, pp. 6 – 18.

16. Chieh-Chen Wu; Wen-Chun Yeh; Wen-Ding Hsu; Md Mohaimenul Islam; Phung Anh Alex Nguyen; Tahmina Nasrin Poly; Yao-Chin Wang; Hsuan-Chia Yang; Yu-Chuan Jack Li, Year: 2019, "Prediction of fatty liver disease using machine learning algorithms", Computer Methods Prog Biomed., Vol: 170, pp. 23 – 9.

17. Pavleen Kaur; Ravinder Kumar; Munish Kumar, Year: 2019, "A healthcare monitoring system using random forest and internet of things (IoT)", Multimed Tools Appl., Vol: 78, pp. 19905 – 16.

18. Jesmin Nahar; Tasadduq Imam; Kevin S. Tickle; Yi-Ping Phoebe Chen, Year: 2013, "Computational intelligence for heart disease diagnosis: a medical knowledge driven approach", Expert Syst Appl., Vol: 40, no: 1, pp. 96 – 104.

19. Luxmi Verma; Sangeet Srivastava; P. C. Negi, Year: 2016, "A hybrid data mining model to predict coronary artery disease cases using non-invasive clinical data", J Med Syst., Vol: 40, no: 7, pp. 1 – 7.

20. Randa El-Bialy; Mostafa A. Salamay; Omar H. Karam; M. Essam Khalifa, Year: 2015, "Feature analysis of coronary artery heart disease data sets", Proc Comput Sci., Vol: 65, pp. 459 – 68.