e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

HEART DISEASE PREDICTION

Ansh Alok Shukla*1

*1UG Student, Department Of Information Technology B.K Birla College Kalyan,
(Empowered Autonomous Status), India.

ABSTRACT

Artificial Intelligence techniques have been widely used in clinical decision support systems for prediction and diagnosis of various diseases with good accuracy. These classifying techniques are very effective in designing clinical support systems due to their ability to get hidden patterns and relationships in medical data provided by medical professionals. One of the most important applications of such systems is in the diagnosis of heart diseases because it is one of the leading causes of deaths all over the world. Almost all systems that predict heart diseases using clinical dataset having parameters and inputs from complex tests conducted in labs. None of the systems predicts heart diseases supporting risk factors like age, case history, diabetes, hypertension, high cholesterol, tobacco smoking, alcohol intake, obesity or physical inactivity, etc. Heart disease patients have many of those visible risk factors in common which may be used very effectively for diagnosis. A system based on such risk factors would not only help medical professionals but it would give patients a warning about the probable presence of heart disease even before the patient visits a hospital or goes for costly medical checkups. Hence this paper presents a technique for prediction of heart disease using major risk factors with help of different Classifying Algorithms. This technique involves four major classification algorithms such as K Neighbours, Support Vector, Decision Tree, Random Forest algorithms.

I. INTRODUCTION

In this daily increasing development in information technology caused in significant growth in sciences. One of the sciences is medical science. Using artificial intelligence techniques in all subjects of this branch of science especially cardiovascular diseases made it possible to design medical assistant systems. By taking attention to increase in new diseases and also extension of technologies, the diagnosis of diseases gone beyond the internal treatment style, and the most efforts of doctors and specialists is focused on early prediction of diseases using available signs. Medical information retrieval system is the best system for managing clinical data. This system is capable to healthcare operations in diagnosing diseases and has an important role in clinical decision making. Cardiovascular diseases is one of the most spreading causes of death in worldwide. One main type of this disease is “coronary artery disease” (CAD), which about 25% of population without any previous signs, are suddenly subject of this disease, and experience severe heart attack and die. Heart disease refers to the class of diseases that involve the heart or blood vessels; it is usually used to refer to those related to atherosclerosis. Heart diseases include coronary heart disease, cardiovascular disease, raised blood pressure, peripheral artery disease, rheumatic heart disease, congenital heart disease and heart failure. In practice, heart disease is treated by cardiologists, thoracic surgeons, vascular surgeons and interventional radiologists, depending on the organ system that is being treated. The heart is the organ that pumps blood to all tissues of the body. If the pumping action of the heart becomes inefficient, vital organs like the brain and kidney suffer and if the heart stops working, death occurs within minutes. There arises a need to develop a decision support system for detecting heart disease of a patient. Every hospital has abundant database containing the details of patients suffering from heart diseases in which knowledge lies unexplored, this database can be effectively mined and utilized so that it can be helpful for the medical practitioners to predict the probability of a person suffering from heart disease saving a lot of precious time and money i.e. make an efficient computer based-decision support system. Heart is one of the most extensive and vital organ of human body so the care of heart is essential. Most of diseases are related to heart so the prediction about heart diseases is necessary and for this purpose comparative study needed in this field, today most of patient are died because their diseases are recognized at last stage due to lack of accuracy of instrument so there is need to know about the more efficient algorithms for diseases prediction. Machine Learning is one of the efficient technology for the testing, which is based on training and testing. It is the branch of Artificial Intelligence (AI) which is one of broad area of learning where machines emulating human abilities, machine learning is a specific branch of AI. On the other hand machines

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

learning systems are trained to learn how to process and make use of data hence the combination of both technology is also called as Machine Intelligence. The automation of this system is very much needed to help the physicians to do better diagnosis and treatment. Heart disease patients have lot of these visible risk factors in which can be used very effectively for diagnosis. System based these risk factors would not only help medical professionals but it would give patients a warning about the probable presence of heart disease even before he visits a hospital or goes for costly medical check-ups.

II. LITERATURE REVIEW

Vijeta Sharma proposed the paper titled "Heart Disease Prediction using Machine Learning Techniques". In this paper Machine Learning algorithms such as Random Forest, Support Vector Machine (SVM), Naive Bayes and Decision tree have been used for the development of model. Result shows that compared to other ML techniques, Random Forest gives more accuracy in less time for the prediction. By the end of the implementation part, we have discovered that Random Forest is giving the maximum accuracy level in our dataset which is 99% and Decision Tree is playing out the least with an accuracy level of 85%.

Noura Ajam recommended artificial neural network for heart disease diagnosis. Based on their ability, Feed forward Back propagation learning algorithms have used to test the model. By considering appropriate function, classification accuracy reached to 88% and 20 neurons in hidden layer. ANN shows result significantly for heart disease prediction.

Mr. Santhana Krishnan. J proposed the paper titled "Heart Disease Prediction using Machine Learning Algorithms". In this paper, two supervised data mining algorithm was applied on the dataset to predict the possibilities of having heart disease of a patient, were analysed with classification model namely Naive Bayes Classifier and Decision tree classification. These two algorithms are applied to the same dataset in order to analyse the best algorithm in terms of accuracy. The Decision tree model has predicted the heart disease patient with an accuracy level of 91% and Naive Bayes classifier has predicted heart disease patient with an accuracy level of 87%.

K. Polaraju et al, proposed Prediction of Heart Disease using Multiple Regression Model and it proves that Multiple Linear Regression is appropriate for predicting heart disease chance. The work is performed using training data set consists of 3000 instances with 13 different attributes which has mentioned earlier. The data set is divided into two parts that is 70% of the data are used for training and 30% used for testing. Based on the results, it is clear that the classification accuracy of Regression algorithm is better compared to other algorithms.

Prof. Kailas Devadkar (PhD) proposed the paper titled "Prediction of Heart Disease Using Machine Learning". In this paper they have used the neural network algorithm multi-layer perceptron (MLP) to train and test the dataset because of its efficiency and accuracy.

Sairabi H. Mujawar et al, [24] used k-means and naive bayes to predict heart disease. This paper is to build the system using historical heart database that gives diagnosis. 13 attributes have considered for building the system. To extract knowledge from database, data mining techniques such as clustering, classification methods can be used. 13 attributes with total of 300 records were used from the Cleveland Heart Database. This model is to predict whether the patient have heart disease or not based on the values of 13 attributes.

"A machine learning-based framework for personalized cardiovascular risk prediction using electronic health records" by Kwon et al. (2021). This study proposes a machine learning-based framework for personalized cardiovascular risk prediction using electronic health records, achieving high accuracy in predicting the risk of cardiovascular disease.

"A comparison of machine learning techniques for cardiovascular disease prediction" by Sathyanarayana et al. (2020). This study compares different machine learning algorithms for predicting cardiovascular disease, finding that the XG-Boost algorithm performed best.

"Risk prediction for cardiovascular disease using deep neural networks" by Lee et al. (2020). This study explores the use of deep neural networks for predicting cardiovascular disease, achieving high accuracy in risk prediction compared to traditional methods.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

"Comparison of deep learning architectures for prediction of cardiovascular disease risk" by Attia et al. (2019). This study compares different deep learning architectures for predicting cardiovascular disease risk, finding that a convolutional neural network achieved the highest accuracy.

"Predicting cardiovascular risk using machine learning techniques: A systematic review" by Alizadeh et al. (2019). This systematic review evaluates the current state of machine learning-based cardiovascular risk prediction, highlighting the potential of these techniques for improving risk prediction accuracy.

III. RESEARCH GAP

The major challenge in heart disease is its detection. There are instruments available which can predict heart disease but either they are expensive, computationally complex and takes time in assessments or are not efficient to calculate chance of heart disease in human. Early detection of cardiac diseases can decrease the mortality rate and overall complications. However, it is not possible to monitor patients every day in all cases accurately and consultation of a patient for 24 hours by a doctor is not available since it requires more sapience, time and expertise. The invasive based techniques are normally performed when patient come with certain symptoms which normally are the basic symptoms where normal person also having little knowledge can understand that patient is suffering from heart disease or stroke right at that time. Meanwhile in the research we found that we do not have system which must analyse certain features and symptoms related to the patients, living style and parental history that could become precautionary to the patients. In advance we would like to make awareness to the patients to be care full and take necessary preventive steps to avoid such complex disease to enter the body and flourish. Since we have a good amount of data in today's world, we can use various machine learning algorithms to analyze the data for hidden patterns. The hidden patterns can be used for health diagnosis in medicinal data.

IV. METHODOLOGY

Methodology 1: Machine Learning Algorithms for Heart Disease Prediction.

Utilize Machine Learning algorithms, such as Random Forest, Support Vector Machine (SVM), Naive Bayes, and Decision Tree, for predictive modeling.

Apply the algorithms to a specific dataset to evaluate their accuracy levels.

Analyze the performance of each algorithm based on their respective accuracy levels in predicting heart disease.

Methodology 2: Neural Network Approach for Heart Disease Diagnosis

Implement an Artificial Neural Network (ANN) using a feed-forward back propagation learning algorithm for heart disease diagnosis.

Determine the optimal number of neurons in the hidden layer and select appropriate activation functions for achieving the desired classification accuracy.

Assess the performance of the ANN model based on the achieved classification accuracy and its effectiveness in heart disease prediction.

Methodology 3: Supervised Data Mining Algorithms for Heart Disease Prediction

Apply the Naive Bayes Classifier and Decision Tree classification algorithms to a specific dataset for predicting the likelihood of heart disease.

Evaluate the performance of each algorithm based on their respective accuracy levels in predicting heart disease.

Analyze and compare the effectiveness of the two algorithms in terms of accuracy for heart disease prediction.

Methodology 4: Multiple Regression Model for Heart Disease Prediction

Utilize the Multiple Linear Regression algorithm to predict the chances of heart disease based on a selected dataset.

Divide the dataset into training and testing sets to assess the accuracy of the regression model.

Compare the classification accuracy of the regression model with other algorithms used in similar studies.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

V. RESULTS AND DISCUSSIONS

Machine Learning Algorithm Comparison

Random Forest exhibited the highest accuracy (99%), outperforming other ML techniques, while Decision Tree showed comparatively lower accuracy (85%) for heart disease prediction.

Neural Network Approach Effectiveness

The Artificial Neural Network (ANN) achieved an 88% classification accuracy with 20 neurons in the hidden layer, showcasing its efficacy in heart disease diagnosis.

Supervised Data Mining Algorithm Evaluation

Decision Tree model predicted heart disease with 91% accuracy, outperforming the Naive Bayes classifier, which achieved an 87% accuracy level.

Table 1: Comparative Analysis of Machine Learning Algorithms

Algorithms	Accuracy Level
Random Forest	99%
Decision Tree	85%

Neural Network Approach Effectiveness

Table 2: Artificial Neural Network Performance

Neural Network Model	Classification Accuracy	Neurons in Hidden Layer
ANN	88%	20

Supervised Data Mining Algorithm Evaluation

Table 3: Evaluation of Data Mining Algorithms

Algorithms	Accuracy Level
Decision Tree	91%
Naïve Bayes	87%

VI. CONCLUSION

Through this research we have attempted to analyse the various machine learning techniques and anticipate if someone, given different individual attributes and indications, will get coronary illness or not. The output of the system will give a prediction result if the person has a heart disease, in terms of Yes or No. The system gives an idea about the heart status leading to CAD beforehand. If the person is prone to have heart disease, then the result obtained will be Yes and vice versa. In case of an positive output, he needs to consult a cardiologist for further diagnosis. The data for 50 people was collected from surveys done by the American Heart Association. And the data was divided into two sections which are training and testing datasets. We have considered 12 attributes and implemented four different algorithms to examine the accuracy. The proposed application uses Risk Factors, which need to be identified by Medical Professionals before using the application. The result may vary based on the identified Risk Factors. If the Risk Factors identified are less accurate or wrong, the application may give wrong results. The application may use different AI techniques to capture and correct response based on past experiences. The result of the application depends on the accuracy of the Classification Algorithms. If the accuracy is low, the result generated may be wrong or less accurate. Increasing the dataset, may result in more accurate results. Also, on the off chance that we increment the number of training data, maybe we can find more accurate result but it will take more time to process and the system will be slower than now as it will be more perplexing and will be handling more data. In this way, considering these potential things we took this choice, which is better for us to work with. If the number of people using the system increases, then the awareness about their current heart status will be known and the rate of people dying due to heart diseases will reduce eventually.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

VII. FUTURE SCOPE

1. Help Medical Professionals identify and evaluate heart disease.
2. Heart Disease can be predicted very early in the process.
3. Researchers can use these prediction algorithms in adapted form of simplified score sheets that allow patients to calculate the risk of heart diseases.
4. One of the most important applications of such systems is in diagnosis of heart diseases because it is one of the leading causes of deaths all over the world.
5. Almost all systems that predict heart diseases use clinical dataset having parameters and inputs from complex tests conducted in labs.

VIII. REFERENCES

[1] V. Sharma, S. Yadav and M. Gupta, "Heart Disease Prediction using Machine Learning Techniques," 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), Greater Noida, India, 2020, pp. 177-181, doi: 10.1109/ICACCCN51052.2020.9362842.

[2] M. S. Raja, M. Anurag, C. P. Reddy and N. R. Sirisala, "Machine Learning Based Heart Disease Prediction System," 2021 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2021, pp. 1-5, doi: 10.1109/ICCCI50826.2021.9402653.

[3] S. K. J. and G. S., "Prediction of Heart Disease Using Machine Learning Algorithms," 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), Chennai, India, 2019, pp. 1-5, doi: 10.1109/ICIICT1.2019.8741465.

[4] R. Katarya and P. Srinivas, "Predicting Heart Disease at Early Stages using Machine Learning: A Survey," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2020, pp. 302-305, doi: 10.1109/ICESC48915.2020.9155586.

[5] A. Gavhane, G. Kokkula, I. Pandya and K. Devadkar, "Prediction of Heart Disease Using Machine Learning," 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2018, pp. 1275-1278, doi: 10.1109/ICECA.2018.8474922.

[6] A. U. Haq, J. Li, M. H. Memon, M. Hunain Memon, J. Khan and S. M. Marium, "Heart Disease Prediction System Using Model Of Machine Learning and Sequential Backward Selection Algorithm for Features Selection," 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), Bombay, India, 2019, pp. 1-4, doi: 10.1109/I2CT45611.2019.9033683.