

Heart Disease Diagnosis using Machine Learning

Jainav Amit Mutha
School of Electronics Engineering
Vellore Institute of Technology, Chennai, India

Jishnu Saurav Mittapalli
School of Computing Science and Engineering,
Vellore Institute of Technology Chennai, India

Abstract— Heart disease is one of the most important reasons for the death of people in the world. One human dies every 36 seconds in the United States of America alone due to a heart condition. About six hundred thousand people die per annum due to cardiovascular conditions. In this paper we propose a model to predict heart conditions beforehand so that precautionary measures can be taken.

Keywords— Heart attack, heart disease, machine learning, heart disease prediction.

I. INTRODUCTION

Heart is one of the most vital organs in the body. AMI is one of the most important causes of death in many countries in the world. Scientists today are doing a lot of research to help doctors predict such conditions beforehand. Analyzing data of different diseases has been a hot science topic today and it is really important and will help many people. An analysis of the data related to different health problems and its functions can help in predicting the diseases beforehand and understand the health of the organ with effect. wellness of this organ with a degree of certainty. Identification of heart disease is difficult because there are many other riskyIt is difficult to identify heart disease because of several contributory risk factors such as BP, diabetes and a few others. Today many techniques are being used for the analysis of these problems like DM and many other Neural network techniques to decide the severity and complexity of the condition. is complex and hence, the disease must be handled carefully. Analysing the problem ahead of time will avoid an effect on the heart much and prevent premature deaths. Classification algorithms along with data mining and medical science predictions can be made to predict heart diseases.

Especially in the last ten years cardiovascular diseases or heart problems have become the main reason of death worldwide. By the WHO it has been estimated, over 20 million deaths occur each year in the world due to heart problems. More than 80 percent are due to coronary diseases. In low and middle income countries it is even higher. Many pre-existing factors affect the condition of the heart like, eating conditions, Habits and other risk factors like, caffeine, alcohol, smoking etc. Many other factors like, blood pressure, hypertension, high cholesterol, obesity etc play a really important role in causing heart diseases and if we are able to predict this using these factors it will be really useful and can save many lives.

Till now a lot of research has been done to predict heart diseases mostly data mining. Data mining basically refers to

the collection of important information from large collections of data especially used in business, medical, educational fields etc. We can use health data mining here in this case. The next thing that can be used in machine learning. Machine learning is the most quickly burgeoning field of artificial intelligence.

These machine learning algorithms can be used to predict and analyze data in different fields in the world. Using these in the field of health especially to predict heart conditions a lot of benefit can be gotten from these technologies. Data mining helps in exploring huge datasets, and then ML techniques can then be used to analyse and predict new data. The field of medicine has huge amounts of data, and this needs good, efficient algorithms for the purpose of quick and accurate analysis and prediction. Professionals in healthcare can take the right decision but these predictions can be used to give warnings to take actual tests and talk to doctors and tak required measures as these are generally sudden and dangerous[1].

So in this paper, we have discussed the various machine learning algorithms that can be used for the purpose of prediction of heart conditions and their efficiencies. We feel this will help save a lot of lives, help doctors and medical professionals as well.

II. RELATED WORKS

Senthilkumar Mohan, Chandrasegar Thirumalai and Gautam Srivastava in their paper, “Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques” , have used different hybridised Machine learning techniques for predicting heart problems[2].

Sellappan Palaniappan and Rafiah Awang in their paper, “Intelligent heart disease prediction system using data mining techniques”, have used data mining techniques for analysing heart diseases in human beings[3].

Nidhi Bhatla and Kiran Jyoti in their paper, “An Analysis of Heart Disease Prediction using Different Data Mining Techniques”, have analysed different data mining techniques that can be used for cardiac problems[4].

Jyoti Soni, Ujma Ansari, Dipesh Sharma and Sunita Soni in their paper, “Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction” have given an overview on how data mining can be used for the purpose diagnosis of heart problems[5].

Purushottam, Kanak Saxena and Richa Sharma in their paper, "Efficient Heart Disease Prediction System" have given a unique method to predict and prevent heart conditions [6].

A. N. Repaka, S. D. Ravikanti and R. G. Franklin in their paper, "Design And Implementing Heart Disease Prediction Using Naives Bayesian," have shown how naive bayes can be used for the purpose of implementing a heart prediction system [7].

Nikhil Gawande and Alka Barhatte in their paper, "Heart diseases classification using convolutional neural network", have used neural networks for the purpose of heart attack prediction [8].

A. Lakshmanarao, A. Srisaila and T.Srinivasa Ravi Kiran in their paper, "Heart Disease Prediction using Feature Selection and Ensemble Learning Techniques" have used ensemble learning and feature selection for the purpose of disease prediction [9] .

R. Katarya and P. Srinivas in their paper, "Predicting Heart Disease at Early Stages using Machine Learning: A Survey," have basically given a survey of ML techniques used for heart attack prediction to their best level at their time[10].

A. Chauhan, A. Jain, P. Sharma and V. Deep, in their paper, "Heart Disease Prediction using Evolutionary Rule Learning," have used evolutionary rule learning for the purpose of cardiac conditions[11].

T. Kasbe and R. S. Pippal, in their paper, "Design of heart disease diagnosis system using fuzzy logic," have shown how to use fuzzy logic for the purpose of heart disease diagnosis[12].

M. I. H. H and M. R. H. Mondal, in their paper, "Effectiveness of Data Driven Diagnosis of Heart Disease," has proved how effective data can be to diagnose heart disease [13].

Purushottam, K. Saxena and R. Sharma, in their paper "Efficient heart disease prediction system using decision tree," has used decision trees for the purpose of heart disease diagnosis[14].

M. Quijano, J. P. Tello and A. Cadena, in their paper, "Heart disease predictor system based on Artificial Neural Network," have given a method using ANN how heart diseases are predicted using ANNs [15].

III. PROPOSED MODEL

In this paper we have discussed the different machine learning models and algorithms that can be used for the purpose of heart disease prediction. We can use either supervised or unsupervised algorithms for the purpose of cardiac condition diagnosis and prediction. Machine

learning algorithms and techniques are majorly used for binary classification of emotions that is either positive or negative. Machine learning algorithms are even subdivided into the following categories: Supervised algorithms. In these kinds of algorithms, the datasets used for training have pre-labeled classes given to the data. On the basis of

This given dataset, the ML model is trained and the inputs after being processed give out the desired outputs. A trained classifier helps these algorithms to classify the input data as positive or negative. Unsupervised Algorithm algorithms of this kind take in unlabeled data and then with different kinds of algorithms find a hidden pattern or structure and then classified into the different fields, this method does not use pre-labeled data for training. This method can further be classified into association and clustering. The most common algorithms in Unsupervised are Apriori and K-Means algorithms. In this paper we have used the algorithms, K nearest neighbors, Naive bayes, support vector machines, decision trees, random forest and a few more.

For the purpose of development and research in this paper after doing a lot of research we have decided to use the following features, age, sex, Chest Pain, BP, cholestoral, fasting, blood sugar, Electrocardiogram, Max Hear Rate, exercise induced angina, oldpeak, slope, Thalassemia, maximum heart rate, exang, Exercise induced angina.

The architecture of this has been shown in figure 1.

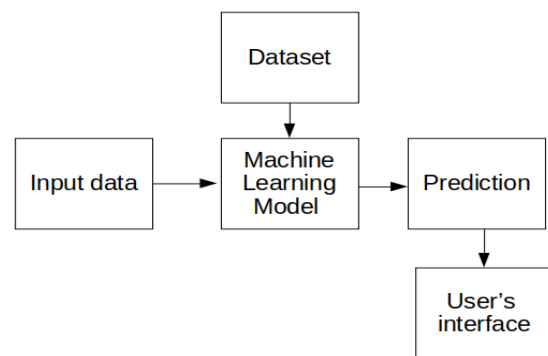


Fig. 1 - Architecture of the model proposed

IV. RESULTS AND DISCUSSION

As discussed above we have implemented these algorithms, 5 algorithms that have been used to predict heart conditions beforehand. Using these algorithms we can predict heart conditions with a good accuracy, at least a warning can be given to people at risk so that they can take precautionary and required measures. The accuracies of each of these models have been shown in table - 1. The models basically take in the pre existing conditions and returns if the person may have a heart condition.

Table - 1 - Algorithm wise accuracy

Algorithm	Accuracy (in percentage)
Naive Bayes	85.24
K Nearest Neighbours	60.81
Support Vector Machine	86.88
Decision tree	79.02
Random Forest	84.46

As we can see SVM gives the highest accuracy of 86.88 percent.

V. CONCLUSION AND FUTURE SCOPE

As we know many people lose their lives to heart disease each year and it is really sad, as people who die because of heart disease die without any warning in most cases. So, if we are able to provide a warning, it would be of great help as people would then be able to take precautionary measures and this may lead to saving many lives every year. People with symptoms like cholesterol, maximum heart rate, Chest Pain, blood sugar etc can check and based on these symptoms predictions are made. We have used five major algorithms for developing the models in order to make predictions. We feel that this model if used will be of great help to both doctors as well as people to check themselves at a basic level.

In the future, we can try to collect more data and improve the accuracy of the models and help predict heart conditions even better. Apart from these we can try out more models including hybrid models of machine learning, deep learning and also neural networks and other predictive algorithms can be used.

REFERENCES

[1] F. Bulut, "Heart attack risk detection using Bagging classifier," 2016 24th Signal Processing and Communication Application Conference (SIU), 2016, pp. 2013-2016, doi: 10.1109/SIU.2016.7496164.

[2] S. Mohan, C. Thirumalai and G. Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques," in IEEE Access, vol. 7, pp. 81542-81554, 2019, doi: 10.1109/ACCESS.2019.2923707.

[3] S. Palaniappan and R. Awang, "Intelligent heart disease prediction system using data mining techniques," 2008 IEEE/ACS International

Conference on Computer Systems and Applications, 2008, pp. 108-115, doi: 10.1109/AICCSA.2008.4493524.

[4] Nidhi Bhatla and Kiran Jyoti in their paper, "An Analysis of Heart Disease Prediction using Different Data Mining Techniques", International Journal of Engineering Research & Technology (IJERT), Vol. 1 Issue 8, October - 2012

[5] Jyoti Soni, Ujma Ansari, Dipesh Sharma and Sunita Soni, "Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction", International Journal of Computer Applications (0975 - 8887), Volume 17- No.8, March 2011.

[6] Purushottam, Kanak Saxena, Richa Sharma, Efficient Heart Disease Prediction System, Procedia Computer Science, Volume 85, 2016, Pages 962-969, ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2016.05.28>

[7] A. N. Repaka, S. D. Ravikanti and R. G. Franklin, "Design And Implementing Heart Disease Prediction Using Naives Bayesian," 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 2019, pp. 292-297, doi: 10.1109/ICOEI.2019.8862604.

[8] N. Gawande and A. Barhatte, "Heart diseases classification using convolutional neural network," 2017 2nd International Conference on Communication and Electronics Systems (ICCES), 2017, pp. 17-20, doi: 10.1109/CESYS.2017.8321264.

[9] A. Lakshmanarao, A. Srisaila and T. S. R. Kiran, "Heart Disease Prediction using Feature Selection and Ensemble Learning Techniques," 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), 2021, pp. 994-998, doi: 10.1109/ICICV50876.2021.9388482.

[10] 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), 2020, pp. 302-305, doi: 10.1109/ICESC48915.2020.9155586.

[11] A. Chauhan, A. Jain, P. Sharma and V. Deep, "Heart Disease Prediction using Evolutionary Rule Learning," 2018 4th International Conference on Computational Intelligence & Communication Technology (CICT), 2018, pp. 1-4, doi: 10.1109/CICT.2018.8480271.

[12] T. Kasbe and R. S. Pippal, "Design of heart disease diagnosis system using fuzzy logic," 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), 2017, pp. 3183-3187, doi: 10.1109/ICECDS.2017.8390044.

[13] M. I. H. Khan and M. R. H. Mondal, "Effectiveness of Data Driven Diagnosis of Heart Disease," 2020 11th International Conference on Electrical and Computer Engineering (ICECE), 2020, pp. 419-422, doi: 10.1109/ICECE51571.2020.9393055.

[14] Purushottam, K. Saxena and R. Sharma, "Efficient heart disease prediction system using decision tree," International Conference on Computing, Communication & Automation, 2015, pp. 72-77, doi: 10.1109/CCAA.2015.7148346.

[15] M. Quijano, J. P. Tello and A. Cadena, "Heart disease predictor system based on Artificial Neural Network," 2013 Pan American Health Care Exchanges (PAHCE), 2013, pp. 1-4, doi: 10.1109/PAHCE.2013.6568350.