

A Survey Techniques Used for Prediction of Heart Attack with Machine Learning and Medical Text Mining

Divya Yadav^{*1}, Prof. Gayatri Pandi (Jain)²

^{*1}L.J Institute of Engineering and Technology, Gujarat Technological University, Ahmedabad, Gujarat, India

²H.O.D., L.J Institute of Engineering and Technology, Gujarat Technological University, Ahmedabad, Gujarat, India

ABSTRACT

Heart attack is one of the most critical heart disease in the world and affects human life very badly. In heart attack, the heart is unable to push the required amount of blood to other parts of the body. Accurate and on time diagnosis of heart attack is important for heart failure prevention and treatment. The diagnosis of such condition through traditional medical history has been considered as not reliable in many aspects. To classify the healthy people and people with heart attack causes and related problems, noninvasive-based methods such as machine learning are reliable and efficient. In the proposed study, we developed a machine-learning-based diagnosis system for heart attack prediction by using heart disease dataset. We used popular machine learning algorithms for performance evaluation metrics such as classification accuracy, sensitivity and correlation coefficient. The proposed system can easily predict and classify people with heart attack possibilities from healthy people.

Keywords : Heart Disease, Machine Learning, Deep Learning, MLP, Stacking Approach

I. INTRODUCTION

One of the most common disease in recent times is related to heart. Heart diseases are not just coronary diseases but they vary for a more inner parts which are connected to heart. congenital heart disease, Arrhythmia, coronary artery disease are some of the major diseases. Data gathering about such disease has been the part of the study since a long time. The medical data such as Blood pressure, hypertension, diabetes, cigarette smoked per day and so on is taken as input and then these features are modelled for prediction. This model can then be used to predict future medical data. The algorithms like K-nearest neighbour, Naïve Bayes, support vector machine and decision tree are used in such various researches.

The number of people affected by heart disease increases irrespective of age in both men and women. But other factors like gender, diabetes, BMI also contribute to this disease. In this paper, we have tried prediction and analysis of heart disease by considering the parameters like age, gender, blood pressure, heart rate, diabetes and so on. Since numerous factors are involved in heart disease, the prediction of this disease is challenging. Some of major symptoms of heart attack are: Chest tightness, Shortness of breath, Nausea, Indigestion, Heartburn, or stomach pain, Sweating and Fatigue, Pressure in the upper back Pain that spreads to the arm. [3] The following are the type of heart disease: Heart means “cardio”. Hence all heart diseases concern to category of cardiovascular diseases. The different kinds of heart disease are: Coronary heart diseases, Angina pectoris, Congestive

heart failure, Cardiomyopathy, Congenital heart diseases. [9] Coronary heart disease or coronary artery disease is the narrowing of the coronary arteries. The coronary arteries supply oxygen and blood to the heart. It causes a large number of people to become ill or to face death. It is one of the popular type of heart disease. High blood glucose from diabetes can damage blood vessels and nerves that control heart and blood vessels. If a person has diabetes for a longer time, there are high chances for that person to have heart disease in future. With diabetes, there are other reasons which contribute to heart disease. They are smoking which raises the risk of developing heart disease, high blood pressure makes the heart work harder to pump blood and it can strain heart and damage blood vessels, abnormal cholesterol levels also contribute to heart disease and obesity. Also, family history of heart disease can be a cause of having heart disease. But this history is not considered in this paper for prediction of heart disease

II. LITERATURE SURVEY

Effective heart attack prediction machine learning techniques [1], is a paper focusing on ML techniques being used in recent developments in different areas of the Internet of Things (IoT). Data pre processing is the first stage in the proposed algorithm. The multi-class variable is used to check the presence or absence of heart disease. In the instance of the patient having heart disease, the value is set to 1, else the value is set to 0 indicating the absence of heart disease in the patient. The pre-processing of data is carried out by converting medical records into diagnosis values. The results of data pre-processing for 297 patient records indicate that 137 records show the value of 1 establishing the presence of heart disease while the remaining 160 reflected the value of 0 indicating the absence of heart disease. In second stage feature selection and reduction, naive bayes, GLM, DL and SVM is used to satisfy the requirements of features. The third stage consisting of classification modelling which is

presenting the results after applying the best performing techniques of classification in machine learning[1].

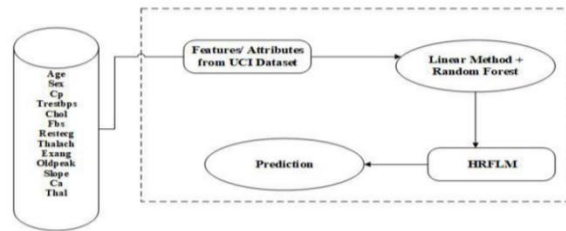


Figure 1. The HRFLM approach

The paper titled as “ensemble approach for developing smart heart disease prediction system” [2] using classification algorithms is focusing on ensemble approach in combining the prediction techniques to yield better results. Basically it uses neural network to apply better naive Bayesian or SVM or Random forest classifier to accurately predict the data. Their experimental analysis is showing 98% accuracy with SVM and Random Forest.[2].

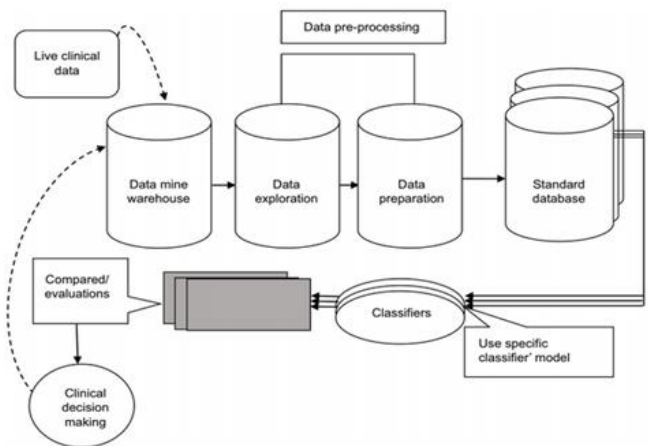


Figure 2. Methodology of Mining Heart Disease Data

Ensemble classification approach improves accuracy of prediction [3] is the key point after developing the novel approach for prediction of risk of heart disease. It further leads towards comparative analytical approach that determines the ensemble technique can be applied. Techniques like boosting, bagging, stacking, majority vote are applied for ensemble classifying technique implementation.

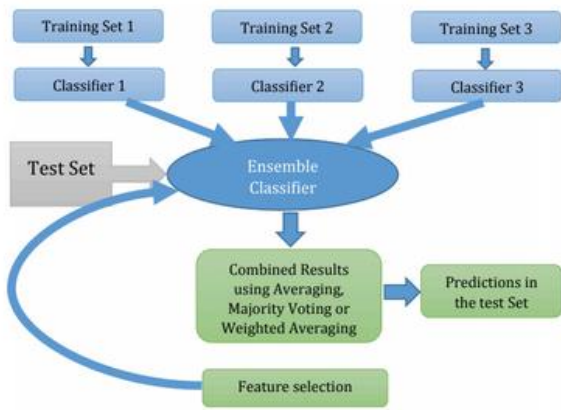


Figure 3. The Ensemble Approach

The results found encouraging for real time applications. At the same time it performs more than 20% better than the NB, SVM, C4.5 and other major used classifiers.

Classification via regression classifier and random forest classifier are mostly being ensemble in algorithms which have proven better accuracy.

Paper demonstrates five classification algorithms which were used for developing the DM model. For sampling the training and testing dataset, 10-fold cross-validation was applied. The performance and the accuracy of each experiment are evaluated through performance measures such as true positive rate, precision, F-measure, receiver operating characteristic (ROC) area, Kappa statistics and root mean square (RMS) error.[3]

On the concept that humans don't start their thinking from zero level every time, a recurrent model is applied for ensemble approach to accurately prediction for cardiovascular diseases. LSTM (Long term dependency based recurrent neural network) is used to achieve the remarkable results with RNNs. Written down as a set of equations, LSTMs look pretty intimidating. Hopefully, walking through them step by step in this field has made them a bit more approachable information.[8] As a result, in this paper, the researcher has proposed an algorithm to predict the risk factors of the CVD using the attention module based Long Short- Term Memory (LSTM),

which has almost 95% accuracy and 0.90 Matthews Correlation Coefficient (MCC) scores; better than any other previously proposed methods.[4]

The paper is proposing a model that can demonstrate it as a end-user tool

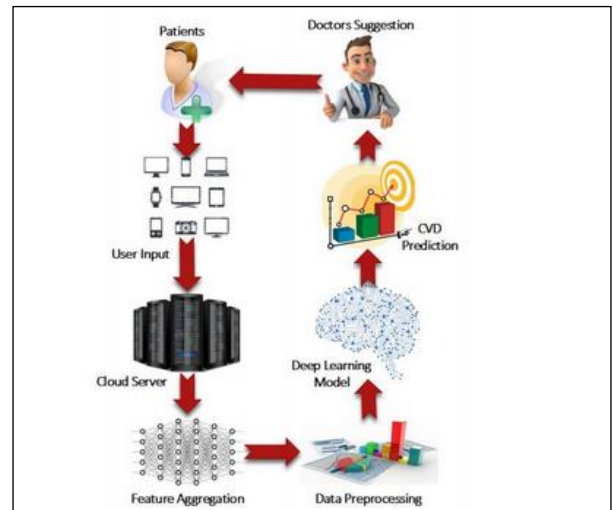


Figure 4. Intelligent Healthcare Platform

How helpful it would be, if we can notify about upcoming heart attack as alert or message to the person. This pre attack analysing and proposing system can really help in saving a lot of lives by timely medication and help. The paper presents a system for early detection and alerting of the onset of a heart attack. The system consists of a wireless and mobile ECG biosensor, a data centre, smart phone and web applications, and a remote 24 h health care.[5]

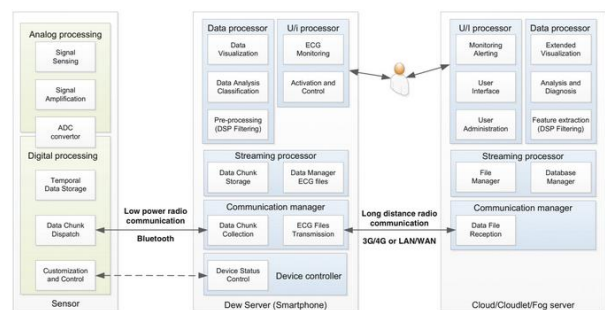


Figure 5. The system architecture for early detection of heart attack

The system proposes the detection by past ECG data which is further classified with available trained ECG datasets. IOT device developed for the proposed

system sends ECG signals via sensors instead of continuous recording. [5] Complexity is managed by the intervals of sending ECG signals to the data center where wave detection process is taken place and diagnosis is sent to the doctor.

Artificial neural network can play vital role in analysing heart attack from past records. This result can then be used to predict future of heart attack for one patient. With machine learning there are several techniques and classifiers are introduced. Data analytics is useful for prediction from more information and it helps medical centre to predict of various disease. [6]

Data mining techniques for both classification and clustering is used in this area. Fuzzy logic, Random Forest, SVM, K-NN and naive bayes are some of the highly used algorithms. A comparative study is performed to measure accuracy of the models. The accuracy is then dependent over multiple criterias like number of attributes, ratio of noisy data, type of data and pre processing applied over data.

The proposed an analysis of cardiovascular disease. This paper proposed data mining techniques to predict the disease. It is intend to provide the survey of current techniques to extract information from dataset and it will useful for healthcare practitioners. The performance can be obtained based on the time taken to build the decision tree for the system. The primary objective is to predict the disease with less number of attributes. [6]

III. CONCLUSION

Survey of multiple techniques proposed and applied by different researchers has given me a lot of insight about heart disease prediction scenario. Our referred papers have met higher level results in case of cardiovascular disease and heart attack prediction. In the same plethora we have tried to compare the various algorithms like decision tree classifier, SVM, K-NN, NB and k-means. Machine learning and deep

learning techniques like RNN, ANN and CNN are also used in many cases when there are good level of results are required.

IV. REFERENCES

- [1] Sethilkumar Mohan, Chandrasegar Thirumalai, Gautam Shrivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques," IEEE- journal IEEE-Access - 2019
- [2] Mustafa Jan, Akber A Awan, Muhammad S Khalid, Salman Nisar, "Ensemble approach for developing a smart heart disease prediction system using classification algorithms" – Dove Press – 2019.
- [3] C. Beluah Christian Ledha, S. Caroleen Jeeva, "Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques" – Elsevier-2019
- [4] MD Samiul Islam, Haider Muhammd Imran, Samir M Umran, Mohammad Karim, "Intelligent Healthcare Platform: Cardiovascular Disease Risk Factors Prediction Using Attention Module Based LSTM" – IEEE-2019
- [5] Marjan Gusav, Aleksander Stojmiski, Ana Guseva, "ECGAlert: A Heart Attack Alerting System" – Springer – 2017
- [6] M. Marimuthu, M. Abinaya, K.S.Harish, K. Madhankumar, V. Pavithra "A Review on Heart Disease Prediction using Machine Learning and Data Analytics Approach" – IJCA-2018
- [7] 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.
- [8] 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.

- [9] A. S. Abdullah and R. R. Rajalaxmi, "A data mining model for predicting the coronary heart disease using random forest classifier," in Proc. Int. Conf. Recent Trends Comput. Methods, Commun. Controls, Apr. 2012, pp. 22–25.
- [10] C. A. Devi, S. P. Rajamhoana, K. Umamaheswari, R. Kiruba, K. Karunya, and R. Deepika, "Analysis of neural networks based heart disease prediction system," in Proc. 11th Int. Conf. Hum. Syst. Interact. (HSI), Gdansk, Poland, Jul. 2018, pp. 233–239.

Cite this article as :

Divya Yadav, Prof. Gayatri Jain, "A Survey Techniques Used for Prediction of Heart Attack with Machine Learning and Medical Text Mining ", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 6 Issue 6, pp. 108-112, November-December 2019. Available at doi : <https://doi.org/10.32628/IJSRSET196630>
Journal URL : <http://ijsrset.com/IJSRSET196630>