

INTRODUCTION

Heart is an important organ of the human body. It pumps blood to every part of our anatomy. If it fails to function correctly, then the brain and various other organs will stop working, and within few minutes, the person will die. Change in lifestyle, work related stress and bad food habits contribute to the increase in rate of several heart related diseases. Heart diseases have emerged as one of the most prominent cause of death all around the world. According to World Health Organisation, heart related diseases are responsible for the taking 17.7 million lives every year, 31% of all global deaths. In India too, heart related diseases have become the leading cause of mortality. Predicting and diagnosing heart disease is the biggest challenge in the medical industry and it is based on factors like physical examination, symptoms and signs of the patient. Factors which influence heart diseases are cholesterol level of the body, smoking habit, and obesity, family history of diseases, blood pressure and working environment. Machine learning algorithms play a vital and accurate role in predicting heart disease. The advancement of technologies allows machine language to pair with big data tools to handle unstructured and exponentially growing data. Thus, feasible and accurate prediction of heart related diseases is very important. Medical organisations, all around the world, collect data on various health related issues. Heart disease is perceived as the deadliest disease in the human life across the world. In particular, in this type of disease the heart is not capable in pushing the required quantity of blood to the remaining organs of the human body in order to accomplish the regular functionalities. Some of the symptoms of heart disease include physical body weakness, improper breathing, swollen feet, etc. The techniques are essential to identify the complicated heart diseases which results in high risk in turn affect the human life. Presently, diagnosis and treatment process are highly challenging due to inadequacy of physicians and diagnostic apparatus that affect the treatment of heart patients. Early diagnosis of heart disease is significant to minimize the heart related issues and to protect it from serious risks. The invasive techniques are implemented to diagnose heart diseases based on medical history, symptom analysis report by experts, and physical laboratory report. Moreover, it causes delay and imprecise diagnosis due to human intervention. It is time consuming, computationally intensive and expensive at the time of assessment.

ABSTRACT

Today, heart diseases have become one of the leading causes of deaths in nationwide. The best prevention for this disease is to have an early system that can predict the early symptoms which can save more life. Recently research in data mining had gained a lot of attention and had been used in different kind of applications including in medical. The use of data mining techniques can help researchers in predicting the probability of getting heart diseases among susceptible patients. Among prior studies, several researchers articulated their efforts for finding a best possible technique for heart disease prediction model. This study aims to draw a comparison among different algorithms used to predict heart diseases. The results of this paper will helps towards developing an understanding of the recent methodologies used for heart disease prediction models. The use of analytics in healthcare improves care by facilitating preventive care and EDA is a vital step while analysing data. In this paper, the risk factors that causes heart disease is considered and predicted using K-means algorithm and the analysis is carried out using a publicly available data for heart disease. The dataset holds

209 records with 8 attributes such as age, chest pain type, blood pressure, blood glucose level, ECG in rest, heart rate and four types of chest pain.

CONCLUSION

Heart stroke and vascular disease are the major cause of disability and premature death. Chest pain is the key to recognize the heart disease. In this work, the heart diseases are predicted by considering major factors with four types of chest pain. K-means clustering is one of the simplest and popular unsupervised machine learning algorithms. Here the datasets are clustered and based upon the clusters the happening of chest pain is predicted. Systems based on machine learning algorithms and techniques have been very accurate in predicting the heart related diseases but still there is a lot scope of research to be done on how to handle high dimensional data and overfitting. A lot of research can also be done on the correct ensemble of algorithms to use for a particular type of data.

REFERENCES

Ramadoss and Shah B et al."A. Responding to the threat of chronic diseases in India". Lancet. 2005; 366:1744–1749. doi: 10.1016/S0140-6736(05)67343-6.

R.Kavitha and E.Kannan et al. "An Efficient Framework for Heart Disease Classification using Feature Extraction and Feature Selection Technique in Data Mining ", 2016

V. Manikantan & S.Latha,"Predicting the Analysis of Heart Disease Symptoms Using Medicinal Data Mining Methods", International Journal on Advanced Computer Theory and Engineering, Volume-2, Issue-2, pp.5-10, 2013.

Uma.K, M.Hanumathappa, "Heart Disease Prediction Using Classification Techniques with Feature Selection Method", Adarsh Journal of Information Technology, Volume-5, Issue-2, pp.22-29, 2016