

Heart Disease Prediction using Machine Learning Techniques and Data Mining

¹Akshaykumar R. Dhotre, ²Sandip N.Ahire, ³Vivek G. Valvi, ⁴Aniket D. Landge, ⁵Prof .S.V.Athawale

Abstract: The machine learning can be referred to as the discovery of relationships in large databases automatically and in some cases, it is used for predicting relationships based on the results discovered. Machine learning plays an important role in various applications such as business organizations, e-commerce, health care industry, scientific and engineering. In the health care industry, data mining is mainly used for Disease Prediction. The objective of our works to predict the diagnosis of heart disease with a reduced number of attributes. Here fourteen attributes involved in predicting heart disease. But fourteen attributes are reduced to six attributes by using a Genetic algorithm. Subsequently, three classifiers like Naive Bayes, Classification by Clustering and Decision Tree are used to predict the diagnosis of heart disease after the reduction of the number of attributes.

Keywords: Heart disease prediction; Prediction model; Classification algorithms

Introduction:

It is difficult to identify heart disease because of several contributory risk factors such as high blood pressure, high cholesterol, abnormal pulse rate and many other factors. Various techniques in data mining and neural networks have been employed to find out the severity of heart disease among humans. The severity of the disease is classified based on various methods like K-Nearest Neighbour Algorithm (KNN), Decision Trees (DT), Genetic algorithm (GA), and Naive Bayes (NB). The nature of heart disease is complex and hence, the disease must be handled carefully. Not doing so may affect the heart or cause premature death. The perspective of medical science and data mining are used for discovering various sorts of metabolic syndromes.

Data mining with classification plays a significant role in the prediction of heart disease and data investigation. We have also seen decision trees be used in predicting the accuracy of events related to heart disease. Various methods have been used for knowledge abstraction by using known methods of data mining for prediction of heart disease.

Related Work:

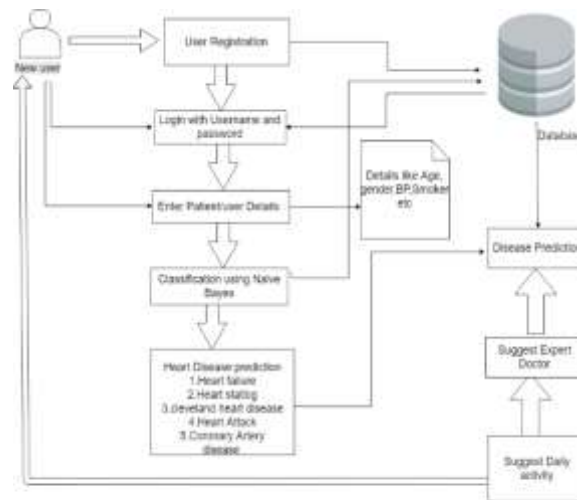
Numerous works [3, 2] related to heart disease diagnosis using Machine learning techniques have motivated this study. The dataset, algorithms, methods used by the authors and the observed results along with the future work are studied for each paper. Large number of work is carried out in finding out efficient methods of medical diagnosis for various diseases. Our work is an attempt to predict efficiently diagnosis with reduced number of factors (i.e. attributes) that contribute more towards the cardiac disease using classification. Sellapan et al (2008), Asha et al(2010) developed an Intelligent Heart Disease Prediction System to predict the heart disease using three classifiers Support vector machine, Naive Bayes. Naive Bayes performed with good prediction probability of 96.6%. Also, 14 attributes were used for prediction.

1. Methodology

To diagnose the disease at early stage at affordable cost is the important aim of this paper. Using machine learning Technique we can detect disease at first stage. We can completely cure the disease by proper diagnosis. Health care industry collect huge amount of information. Which are not mined to locate secret information. Solution of this problem is machine learning technique.

Motivation:

The main motivation of this project is to provide an insight about detecting and curing heart disease using machine learning technique. We turn data into useful information that can enable healthcare practitioners to make intelligent clinical decisions.

System Architecture:**Fig: system overview**

The new user needs to register and login to fill the patient details. Then classify the patient's symptoms and predict the disease. The accuracy of classifiers can be identified by comparing the algorithms and according to that we will classify the disease and recommend the remedies, daily activity for the user.

Algorithm Used:-**1) Naive Bayes:-**

In the Naïve Bayes classifier, we predicate the result, depending upon the training dataset. Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. For example, a fruit may be considered to be an apple if it is red, round, and about 3 inches in diameter.

Even if these features depend on each other or upon the existence of the other features, all of these properties independently contribute to the probability that this fruit is an apple and that is why it is known as Naïve Bayes. Naive Bayes model is easy to build and particularly useful for very large data sets. Along with simplicity, Naive Bayes is known to outperform even highly sophisticated classification methods.

Conclusion:-

Machine learning techniques were used in this work to process raw data and provide a new and novel discernment towards heart disease. Heart disease prediction is challenging and very important in the medical field. However, the mortality rate can be drastically controlled if the disease is detected at the early stages and preventative measures are adopted as soon as possible.

References

- 1) A. Devi, S. Rajamhoana, C. K. Umamaheswari, R. Kiruba, K. Karunya and R. Deepika, Analysis of Neural Networks Based Heart Disease Prediction System, 2018 11th International Conference on Human System Interaction (HSI), Gdansk, 2018, pp. 233-239.
- 2) Cheng, C. Chiu, H., 2017. An Artificial Neural Network Model for the Evaluation of Carotid Artery Stenting Prognosis Using a National- Wide Database. 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp.2566-2569.
- 3) H. A. Esfahani and M. Ghazanfari, "Cardiovascular disease detection using a new ensemble classifier," 2017 IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI), Tehran, 2017, pp. 1011-1014.
- 4) Gavhane, A., 2018. Prediction of Heart Disease Using Machine Learning. 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), (Iceca), pp.1275-1278.
- 5) Jpdlo, V. et al., 2018. Heart diseases prediction with Data Mining and Neural Network Techniques. , 6(7 2), pp.1-6.

- 6) Shafenor Amin, M., Kia Chiam, Y. Dewi Varathan, K., 2018. Identification of significant features and data mining techniques in predicting heart disease. Telematics and Informatics. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S0736585318308876>.
- 7) Shah, S.M.S. et al., 2017. Feature extraction through parallel Probabilistic Principal Component Analysis for heart disease diagnosis. Physica A: Statistical Mechanics and its Applications, 482, pp.796-807. Available at: <http://dx.doi.org/10.1016/j.physa.2017.04.113>.
- 8) Tarle, B., 2017. An Artificial Neural Network Based Pattern Classification Algorithm for Diagnosis of Heart Disease. 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA), pp.1-4.
- 9) Tran, V.P. Al-jumaily, A.A., 2017. Non-Contact Doppler Radar Based Prediction of Nocturnal Body Orientations Using Deep Neural Network for Chronic Heart Failure Patients. , pp.3-7.
- 10) Uyar, K. Ilhan, A., 2017. Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks. Procedia Computer Science, 120, pp.588-593.