Abstract

The diagnosis of disease is a difficult but critical task in medicine. Data mining is the process of extracting hidden interesting patterns from massive databases. In the healthcare industry, it plays a significant role in predicting disease. Heart disease is a single largest cause of death in developed countries and one of the main contributors to disease burden in developing countries. Data mining is a more convenient tool to assist physicians in detecting diseases by obtaining knowledge and information regarding the disease from patient’s data. By using data mining techniques, it takes less time for the prediction of the disease with more accuracy. This paper aims at analyzing the various data mining techniques namely Decision Trees, Naïve Bayes, Neural Networks, Random Forest Classification, and Support Vector Machine by using the Cleveland dataset for Heart disease prediction. Few of the supervised learning algorithms are used for the prediction of heart disease. It provides a quick and easy understanding of various prediction models in data mining and helps to find the best model for further work.


Index Terms

Computer Science  Artificial Intelligence

Keywords

Data mining, Supervised learning, Decision trees, Naïve Bayes, Neural Network, SVM.