Prediction of Heart Disease using Machine Learning Techniques

Abstract: Heart disease is one of the most significant causes of mortality in the world today. Prediction of heart disease is a critical challenge in the area of clinical data analysis. Machine learning (ML) has been shown to be effective in assisting in making decisions and predictions from the large quantity of data produced by the healthcare industry. We have also seen ML techniques being used in recent developments in different areas of the Internet of Things (IoT). Various studies give only a glimpse into predicting heart disease with ML techniques. I would like to propose a novel method that aims at finding significant features by applying machine learning techniques resulting in improving the accuracy in the prediction of heart disease.

Introduction:

It is difficult to identify heart disease because of several contributory risk factors such as diabetes, 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 Neighbor 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 recent developments in machine learning ML techniques used for Internet of Things (IoT) as well.

The machine learning predictive models need proper data for training and testing. The performance of machine learning model can be increased if balanced dataset is used for training and testing of the model. Furthermore, the model predictive capabilities can improve by using proper and related features from the data. Therefore, data balancing and feature selection is significantly important for model performance improvement. Feature selection has great influence in numerous applications such as building simpler, increasing learning performance, creating clean and understandable data. The feature selection from big data is challenging job and create big problems because big data has many dimensions. Further, challenges of feature selection for structured, heterogeneous and streaming data as well as its scalability and stability issues. For big data analytics challenges of feature selection is very important to resolve.

On another hand, suitable machine learning model is necessary for good results. Obviously, a good machine learning model is a model that not only performs well on data seen during training (else a machine learning model could simply learn the training data), but also on unseen data. To evaluate all classifiers on data and find that they get an average 50% of the cases

might be right. Further, more appropriate cross validation techniques and performance evaluation metrics are critical necessary for a model when model is train and test on dataset.
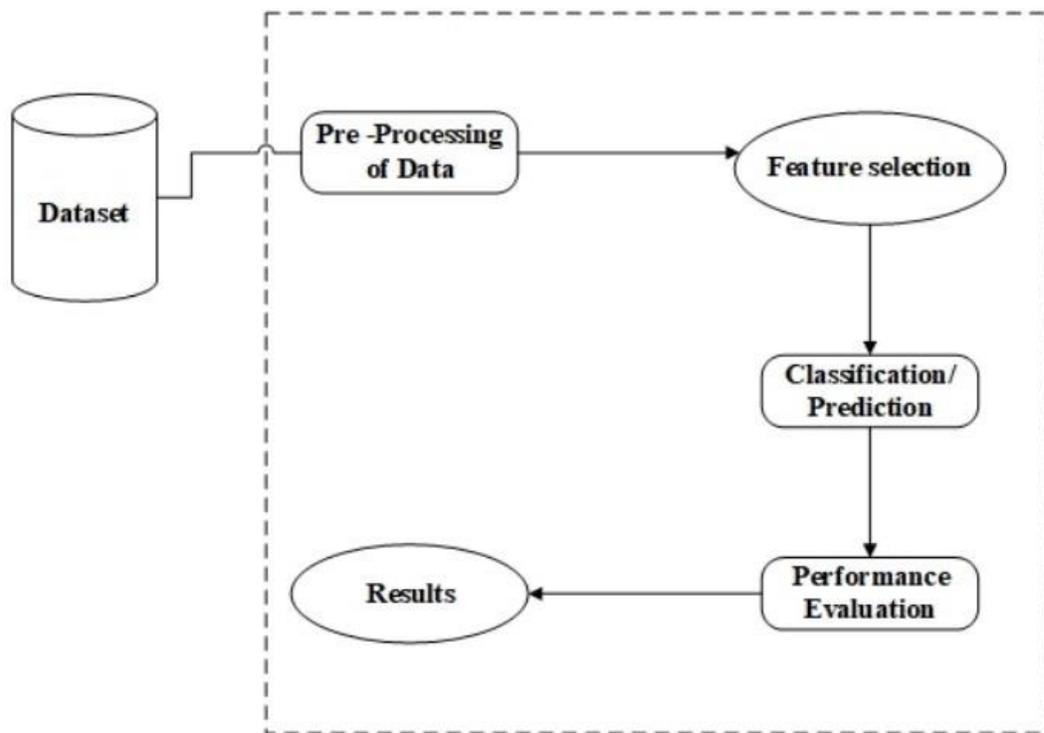
Work Flow for Prediction of Heart Disease using ML

Conclusion:

Machine learning (ML) techniques are crucial in different business fields. Healthcare field facing more problems and it is becoming more expensive. Several ML techniques are used to rectify them. Various ML techniques are being used for prediction of various diseases like heart disease, breast cancer, diabetic disease and thyroid disease. Identifying the processing of raw healthcare data of heart information will help in the long term saving of human lives and early detection of abnormalities in heart conditions. 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. So, I would like to do a detailed study on various existing algorithms being used in predicting heart diseases and analyze them. Then I would like to do my research with various machine learning algorithms or even try few hybrid techniques as well for improving the overall accuracy in the prediction of heart disease and reduce the overall deaths that occur due to heart disease.

by

G. Krishna Lava Kumar