

# Critical Review: Prediction of Heart Disease using K-Means and Artificial Neural Network as Hybrid Approach to Improve Accuracy

Talha Hanif Butt, L18-1864, FAST-NUCES, Lahore

October 12, 2018

At the age above 30, the heart attack or CVD is a common problem can be seen in all human beings. Along with changing lifestyle there are many such factors such as smoking, alcohol, cholesterol level, obesity, high blood pressure, diabetes etc. which are responsible factors for the risk of having heart problems. Data mining plays a vital role in healthcare domain. Data Mining and Machine learning comes up as an emerging field of high importance for providing prognosis and a deeper understanding of medical data.

The proposed system is a hybrid approach to predict or diagnose heart disorders using UCI heart disease dataset by combining K-means and ANN algorithm. This technique uses medical terms such as age, weight, gender, blood pressure and cholesterol rate etc for prediction. To perform grouping of various attributes it uses k-means algorithm and for predicting it uses Back propagation technique in neural networks. The main goal is to obtain high accuracy rate of prediction.

The core modules of the proposed system consist of : a) Understanding the input data and selecting the attribute related to heart disease. b) Data Preparation: transformation and pre-processing of missing data is carried out. c) Processing Module: it specifies about the algorithmic approach applied over the system to obtain high accuracy result. Pre-processing modules are separately discussed in upcoming section. d) Evaluation and deployment: Final Analysing modules provide information related to generated output. It compares and conclude about measurable resultant artefacts like sensitivity, accuracy etc.

Instead of only using K-Means for finding attributes to be used, some other techniques could have been applied to work as validation for their importance. Nothing about the train-test split of data has been provided.

It is analysed that artificial neural network algorithm is best for classification of knowledge data from large amount of medical data. Death rate due to cardiovascular diseases is also increasing. The only solution to control this is to predict the heart disease and medicate it before it gone worse. The hybrid approach proposed gives higher accuracy rate of 97 percent of disease detection than earlier proposed methods.

In conclusion, it can be said that the proposed system addresses the limitations in the related work and also provides better performance on accuracy as well as time taken for predictions but it would be better to test some other methods for selecting attributes so that the confidence on the attributes can be increased. Nothing about the train-test split of data has been provided. Some details about the statistical analysis of the dataset could have proved useful in further understanding of the problem and consequently the solution.