

A Data Mining Model for Coronary Artery Disease Detection using Noninvasive Clinical Parameters

Luxmi Verma¹ and Sangeet Srivastava^{2*}

¹The NorthCap University, Gurugram - 122017, Haryana, India; luxmi.verma@gmail.com

²Department of Applied Sciences, The NorthCap University, Gurugram - 122017, Haryana, India; sangeetsrivastava@ncuindia.edu

Abstract

Coronary Artery Disease (CAD) is one of the major cause of death as well as disability worldwide. Suspected cases of CAD go through invasive and non-invasive tests to get CAD detected. Angiography is a noninvasive method for detection. Not only it is costly, time consuming and risky, but it also needs technical expertise and not suitable for the screening of large population. Hence researchers are looking for better alternatives using non-invasive clinical tests. **Objectives:** To construct Artificial Neural Network based model for CAD identification and adjudging its accuracy with respect to other models. **Method:** Data mining techniques are being employed to identify CAD cases based on non-invasive clinical tests. Early detection of disease is necessary in order to avoid the risk being exaggerated further. Benchmark Cleveland heart disease data is collected from UCI machine repository and Probabilistic Neural Network is employed and trained and tested using non-invasive clinical parameters. **Finding:** Neural Network based model is presented that uses non-invasive clinical parameters of the subjects to model CAD cases and achieves the diagnosis accuracy of 96% and misclassification error rate of 4%. The models' performance is also compared with other classifiers such as RBF Network, AD Tree. **Improvement:** Neural network based model showed the superiority over other methods in terms of accuracy. Results are promising and reproducible and therefore the model can be valuable adjunct tool in clinical practices.

Keywords: Coronary Artery Disease, Data Mining, Decision Tree, Probabilistic Neural Network

1. Introduction

Data mining is defined as extraction of valuable and interesting patterns from enormous data¹⁻⁶. Data mining is used in various fields namely crime data analysis⁷, steganography⁸, education⁹, weather forecasting¹⁰, traffic management¹¹, product quality management¹², retail business¹³ and health sectors¹⁴⁻¹⁸. During the past few decades' researchers have applied data mining methods into health sectors for making clinical decision for prognosis, health care management, treatment planning, prediction of the effectiveness of surgical procedures and identification of various diseases such as cancer, diabetes, cardiovascular diseases etc. CAD is a category of heart disease. It is one of the major reason for death as well

as disability worldwide as per WHO¹⁹. CAD is chronic disease in which accumulation of plaque in coronary arteries gradually hardens and narrowing of coronary artery can lead to heart attack and death. The diagnose of CAD is a complex clinical procedure in which number of factors needs to be considered such as evaluation of risk factors, results of laboratory test and physical examination of the patients. Moreover, diagnosis consumes enormous amount of time, cost, equipment's and requires highly skilled physicians having experience in the field. The cost of care and follow-up of patients is very high making it imperative to identify CAD cases with high order of accuracy. The early diagnosis and prediction helps in reducing the mortality rate and morbidity rate of the disease. There are number of non-invasive methods

* Author for correspondence