

A hybrid model of hierarchical clustering and decision tree for rule-based classification of diabetic patients.

ABSTRACT

Hybrid models in data mining have recently gained attention including in the study of medical research. Various studies in this domain using hybrid models have shown different results. This paper presents the new hybrid model by exploring Agglomerative Hierarchical Clustering and Decision Tree Classifier on Pima Indians Diabetes dataset. The experiments compared performance accuracy of the Decision Tree Classifier against the same classifier augmented with Hierarchical Clustering. Results showed that the hybrid model achieved higher accuracy with 80.8% as compared to 76.9% of the standard model. This is a promising result for adoption of hierarchical clustering in a rule-based classifier.

Keyword: Medical data mining; Hierarchical clustering; Classification; Decision tree; Hybrid model.