

Abstract

Classification is one of predictive Data Mining Task, which used to predict the value form discrete variable target. Classification has been widely implemented in various fields, such as Direct Marketing, Spam Detection, Functional Classification of Proteins, Galaxy Classification etc. Methods in Data Mining Classification are Decision Tree, Neural Network, k-Nearest Neighbor, Case Based Reasoning, Genetic Algorithm etc. Decision tree itself is a functional that uses a tree representation to determine the rules of classification. One of Decision tree method is ID3 (Iterative Dichotomiser three). The algorithm is a function that is by performing weighting on each attribute. Then the tree is built based on the greatest weight to the weight of the smallest and the tree will be constructed by calling the function recursively. Although the tree is built pretty good, but there are better ways to improve the efficiency of time and size of trees, by not ignoring accuracy. Results of analysis found that the use of threshold values can provide a tree construction with efficiency in time, smaller tree size, and good enough accuracy depend on each characteristics that each dataset has.

Kata Kunci : data mining, classification, ID3