

Abstract

Accuracy from model of data mining algorithms is to determine good or not the model, a model that has a low accuracy could be fail to classify a data and this can be fatal if it used in decision support system(DSS), bagging is ensemble method that manipulates the data training into several bagthen build models from the bags, when the time to classify they do voting based on the Output from the models. Bagging can improve accuracy because it can reduce variance and overfitting of the model. Bagging suitable for unstable learning algorithms which models could be changed if the data training was changed, the examples are CART and C4.5.

In this final task,will be build a system that can analyze bagging influence for CART and C4.5 algorithms using 12 dataset taken from UCI repository. Based on the result,CART's and C4.5's Accuracy improved by bagging method, but accuracy not always improve if the number of bag increase, the tendency of many bags are suitable for CART is 25 and C4.5 is 50. Bagging method is not suitable for small dataset and little attribute.

Keywords : Accuracy, bagging, Classification and Regression Tree(CART), C4.5.