

# An Empirical Comparison of Ensemble Methods Based on Classification Trees

Mounir Hamza<sup>1</sup> and Denis Larocque<sup>1</sup>

<sup>1</sup> Department of Quantitative Methods, HEC Montréal, Montréal, Québec, Canada

**Keywords:** Bagging, Boosting, Arcing, Random forest, Classification tree, CART, Noise, Linear combination of variables, Splitting rule, Gini, Entropy, Twoing.

## Abstract

In this paper, we perform an empirical comparison of the classification error of several ensemble methods based on classification trees. This comparison is performed by using fourteen data sets that are publicly available and that were used in Lim, Loh and Shih (Machine Learning 40, 203-228, 2000). The methods considered are a single tree, Bagging, Boosting (Arcing) and random forests. They are compared from different perspectives. More precisely, we look at the effects of noise and of allowing linear combinations in the construction of the trees, the differences between some splitting criteria and, specifically for random forests, the effect of the number of variables from which to choose the best split at each given node. Moreover, we compare our results with those obtained in Lim et al. (2000). In this study, the best overall results are obtained with random forests. In particular, random forests are the most robust against noise. The effect of allowing linear combinations and the differences between splitting criteria are small on average, but can be substantial for some data sets.