

Report
Shengheng Xu (shenghex)
Hao Gao (haog)

Did you receive any help whatsoever from anyone in solving this assignment? No
Did you give any help whatsoever to anyone in solving this assignment? No

1 Tuning the ADTree classifier learner

1.1 What is ADTree

weka.classifiers.trees.ADTree is a class for generating an alternating decision tree. An alternating decision tree (ADTree) is a machine learning method for classification. It generalizes decision trees and has connections to boosting.

1.2 Parameters we varied

We change the two parameters.

-B(number of boosting iterations)	10 to 100
-E(Expand nodes)	-3(all), -2(weight), -1(z_pure), >=0 seed for random walk

Neither of these parameters has significant influence on the performance in terms of error rate. But the larger the number of boosting iteration, the longer it takes to train the model.

1.3 Final settings

We get the best robustness of the classifier with parameters "-B 61 -E 32". The average error/errorNB is 0.669, and the max is 1.032. It takes 0.112 seconds on average to train the models.

2 Tuning the DecisionTable classifier learner

2.1 What is DecisionTable

weka.classifiers.rules.DecisionTable is a class for building and using a simple decision table majority classifier.

2.2 Parameters we varied

We change the two parameters.

-X(number of folds)	1, 2, 5, 10
-S(search method specification)	"weka.attributeSelection.BestFirst -D 1 -N 5", "weka.attributeSelection.GeneticSearch -Z 20 -G 20 -C 0.6 -M 0.033 -R 20 -S 1", "weka.attributeSelection.GreedyStepwise -T -1.7976931348623157E308 -N -1", "weka.attributeSelection.LinearForwardSelection -D 0 -N 5 -I -K 50 -T 0", "weka.attributeSelection.RankSearch -S 1 -R 0 -A weka.attributeSelection.GainRatioAttribute Eval --", "weka.attributeSelection.ScatterSearchV1 -T 0.0 -Z -1 -R 0 -S 1 -D"

Neither of these parameters has significant influence on the performance of the classifier.

2.3 Final settings

We get the best robustness of the classifier with parameters "-X 2 -S weka.attributeSelection.BestFirst -D 1 -N 5". The average error/errorNB is 0.861, and the max is 1.258. It takes 0.090 seconds on average to train the models.

3 Tuning the Decorate classifier learner

3.1 What is Decorate

weka.classifiers.meta.Decorate is a meta-learner for building diverse ensembles of classifiers by using specially constructed artificial training examples. Comprehensive experiments have demonstrated that this technique is consistently more accurate than the base classifier, Bagging and Random Forests. Decorate also obtains higher accuracy than Boosting on small training sets, and achieves comparable performance on larger training sets.

3.2 Parameters we varied

We change the three parameters.

-E(Desired size of ensemble)	10 to 20
-I(Number of iterations)	25 to 75
-W(Full name of base classifier)	"weka.classifiers.trees.J48 -C 0.25 -M 2", "weka.classifiers.trees.RandomForest -I 10 -K 0 -S 1", "weka.classifiers.trees.REPTree -M 2 -V 0.0010 -N 3 -S 1 -L -1", "weka.classifiers.trees.RandomTree -K 0 -M 1.0 -S 1", "weka.classifiers.rules.JRip -F 3 -N 2.0 -O 2 -S 1"

Decorate seems to be sensitive to the base classifier.

3.3 Final settings

We get the best robustness of the classifier with parameters "-E 12 -I 30 -S weka.classifiers.trees.RandomForest". The average error/errorNB is 0.729, and the max is 1.278. It takes 1.349 seconds on average to train the models.

4 Summary

The best classifier we have found is ADTree. We set the parameters to ""-B 61 -E 32" and improve the average error/errorNB from 0.804 to 0.669, and the max from 1.184 to 1.032.