

Decision Tree Report

Rui Wang

According to instructions, I split my 10000 instances into two set, 8000 in train set and 2000 in test set.

After building decision tree on the train set and used test set to do testing, I got the following data:

Correlation coefficient	0.4135
Mean absolute error	0.2704
Root mean squared error	0.3677

It looks like my decision tree is overfit to my train set. I guess the reason is the number of feature I choose to use is not enough.