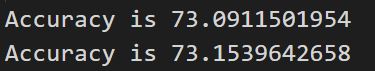
**Decision Tree for predicting horse races**

The model I worked on used the Decision Tree Classifier. The first step in using this model required converting ‘finish\_pos’ from the values of 1 to 8 to the values 0 or 1, positions greater than 3 received a 1 while scores lower or equal to 3 received a 1. Afterwards the data is entered in it’s respectable x and y category; the x data contains all the features besides finish\_pos,row\_id, entry\_id, and morning\_line, while y contains the finish\_pos or finishing positions for each horse/jockey. The usual train/test occurs after, then the training data is ran through the decision tree classifier with a criterion of gini. Gini is used to minimize misclassifications. Entropy is also used in the second model that my code produces; entropy is for exploratory analysis but is a tad slower to compute compared to gini. The overall results for both models were similar, both having scored in the 73~% accuracy. Using grid search to improve the stock parameters I found that the ideal max depth was 10,higher depths resulted in worse scores(usually a percent or two off), and changing the minimum sample had a miniature improvement of .2%.

****