# StarCraft Data Mining

Ben Weber

July 6, 2017

Strategy Prediction

The first chart shows the results for predicting Protoss build orders at different times during the game. It's an attempt at reproducing Figure 4 in my IEEEE CIG 2009 paper. I applied classifiers from the following packages: nearest neighbor (`class`), decision tree (`rpart`), neural networks (`nnet`), and `xgboost`.
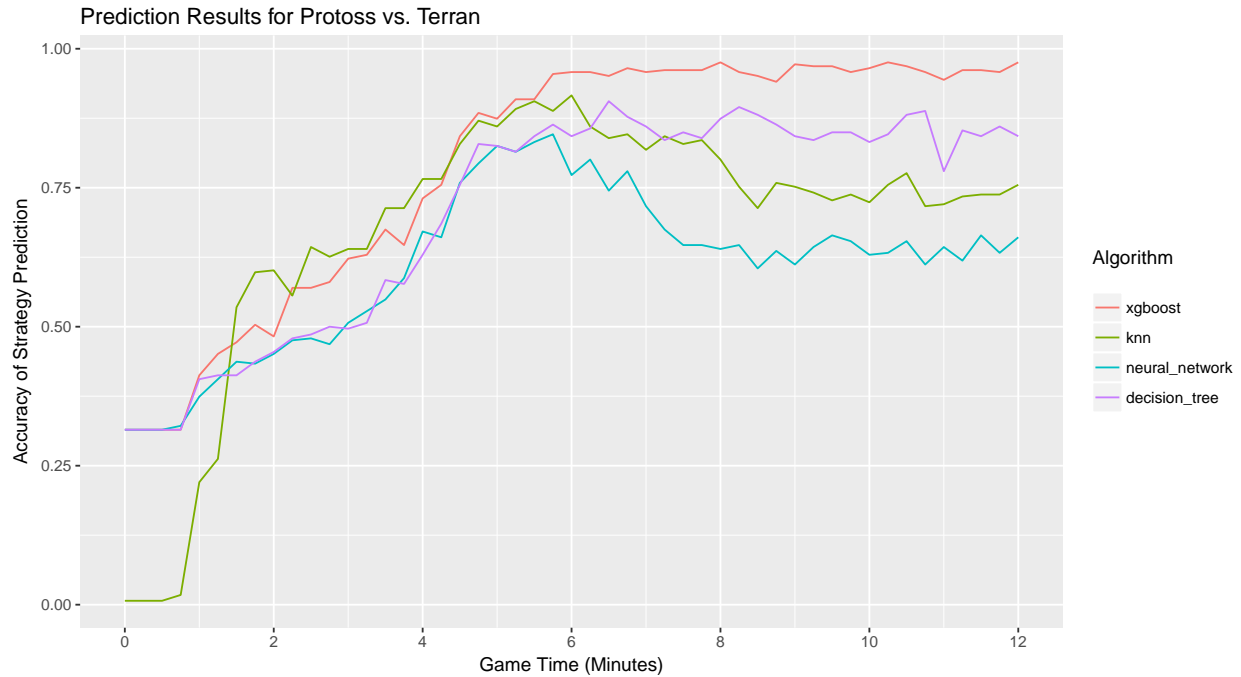
To compare the accuracy of `xgboost` with my prior results, I've also included the raw data for the chart:

```
## minutes    xgboost          knn neural_network decision_tree
##       0  0.3146853 0.006993007      0.3146853     0.3146853
##       1  0.4125874 0.220279720      0.3741259     0.4055944
##       2  0.4825175 0.601398601      0.4510490     0.4545455
##       3  0.6223776 0.639860140      0.5069930     0.4965035
##       4  0.7307692 0.765734266      0.6713287     0.6293706
##       5  0.8741259 0.860139860      0.8251748     0.8251748
##       6  0.9580420 0.916083916      0.7727273     0.8426573
##       7  0.9580420 0.818181818      0.7167832     0.8601399
##       8  0.9755245 0.800699301      0.6398601     0.8741259
##       9  0.9720280 0.751748252      0.6118881     0.8426573
##      10  0.9650350 0.723776224      0.6293706     0.8321678
##      11  0.9440559 0.720279720      0.6433566     0.7797203
##      12  0.9755245 0.755244755      0.6608392     0.8426573
```

Feature Importance

The second chart shows the importance of different features for predicting build orders at 8 minutes into the game, based on the xgboost model. This type of analysis was not covered in the original paper.