KOBE BRYANT SHOT SELECTION

STAT 348: Kaggle Competition Final Project

BACKGROUND

- Kobe Bryant is well regarded as one of the greatest basketball players of all time. His "Mamba Mentality" of chasing perfection governed his life on and off the court.
- During his 20-year NBA career, Kobe put up over 30,000 shots.
- The problem for this Kaggle competition: will Kobe make a shot?
- The data for this competition come from the 30,697 shots he took during his career, split into a train and test set.
- Model results are measured based on mean log loss.

FEATURE ENGINEERING

- Removed variables that may have confounding effects (such as lat and lon, when loc_x and loc_y already exist).
- Converted minutes- and seconds_remaining to a single time_remaining variable.
- Created a dummy variables for whether a game was home or away, the period of the game, the opponent, and whether a game was a playoff game.
- Added a variable for euclidian shot distance.
- Re-encoded the season variable to include numbered seasons instead of years.

MODEL COMPARISON

PENALIZED LOGISTIC REGRESSION

- Used a cross-validated mixture and penalty (tuning parameters).
- Best score: 0.61182

K NEAREST NEIGHBORS

- Used a cross-validated number of neighbors (tuning parameter).
- Best score: 0.62364

RANDOM FORESTS

- Built 1000 trees and used cross-validated values for min_n and mtry (tuning parameters).
- Best score: 0.60054

BEST MODEL – RANDOM FORESTS

- Tuning parameters:
  - Trees: Number of trees: set by me at 1000.
  - Mtry: Number of randomly sampled predictors: cross-validated to 8.
  - Min_n: Minimum number of data points required for a split: cross-validated to 50.
- Used a five-fold cross validation setup.
- Ran in approximately 1 hour and 18 minutes.


FINAL KAGGLE SCORE
0.60054