Predictive Modeling

Arav Goyal and Kirtan Patel

Who are we?

Momentum

- Platform that allows users to predict the result of any given NBA/NFL game
- Takes in two NBA/NFL teams and outputs the winning team and their win probability using data from their last 5 games
- Stores over 13,000 games in a database for efficient querying

Workflow

Step 1 - Find your data

- Find a website/API with the data that you need for your model
  - Getting data from an API is much easier than using code to get it yourself, but there is a lot of documentation online if you want to learn how to webscrape!
- Research what data correlates with the result you want from your model (ie. points and efficiency correlate with wins, etc.)

Step 2 - Clean and engineer your data

- Remove any unnecessary or out-of-content columns/rows (ie. playoff games for a regular season model, etc.)
- Use your data to make new columns (ie. rolling averages, true shooting percentage, RAPTOR, Elo rating, win column, etc.)

Step 3 - Train the model

- Split your data into two pools: train and test (usually 80/20 split)
- Train your model using your training data
- Test your model using your test data
- Find the accuracy/precision of your model

Step 4 - Tweak your model

- You most likely will not make a model you're happy with the first time around
- Add/omit variables and see if your accuracy changes
- Gather more data or find new data
- Use a different model

Your Turn