NBA PLAYER CLASSIFICATION

Hogan Byun
hoganbyun@gmail.com

INTRODUCTION

GOAL:
Create a model that classifies NBA players beyond the traditional 5 positions (PG, SG, SF, PF, C)

Provide insight on team makeup of the top and bottom teams to see what good teams have that bad teams lack

PROBLEMS TO ADDRESS

Player A

0/0.2 3P per game
14.3 3P%
Position: Center

2015 Brook Lopez

Player B

1.8/5.2 3P per game
34.6 3P %
Position: Center

2016 Brook Lopez

PROBLEMS TO ADDRESS

Differences Within Positions

A player's position is not always a good indicator of that player's playstyle

Manually (or automatically) generate types of playstyles to classify players better

Player Development

Players are always evolving, which can result in different playstyles between seasons

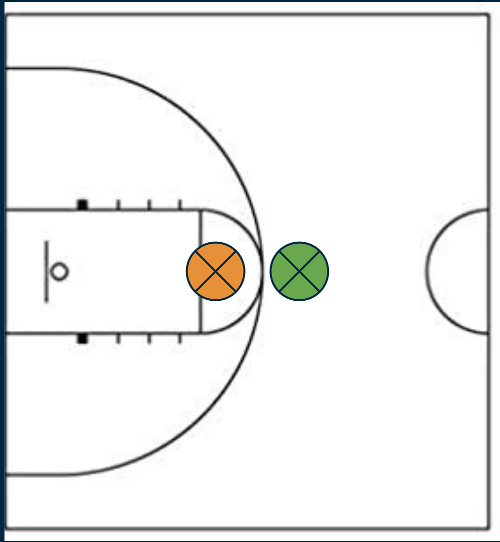
Count each season of a player as a different player
Eg. 2015 Brook Lopez and 2016 Brook Lopez would be different players

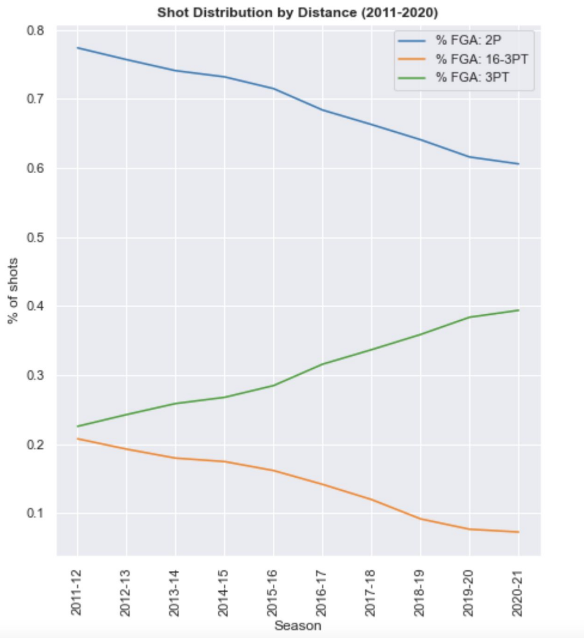
UNDERSTANDING THE PROBLEM

3 Point Era

Teams have started to trade the “long 2” (midrange) for the 3 pointer

2015
Stephen Curry’s 1st MVP
Warriors best record in NBA
Led NBA in 3P% (.398)





3PT per game had been increasing but seeing the Warriors’ success as a mainly 3PT shooting team in 2015 was what triggered a spike

This project only uses player data from 2015-2021

THE DATA

Gathered/Scraped from:

- NBA.com
- Basketball-reference

Contains:

- Traditional player stats (points, rebounds, FG %, etc.)
- Shot location/tendencies (midrange shooting, drives, postups, etc.)

MODELING

Semi-supervised learning

Determine playstyles for a small group of players to use to classify the rest

2 Stage Modeling

Unsupervised Learning

Model separates players into groups, after which, give each group a playstyle

1 Stage Modeling

MODEL RESULTS: SEMI-SUPERVISED

Accuracy: % of correct classifications

Precision: % of actual positives from classified positives

Recall: % of true positives classified

F1: Balance between precision and recall

Micro: Aggregates contributions of all classes to get an average metric

Macro: Independently calculates metrics for each class, then averages them

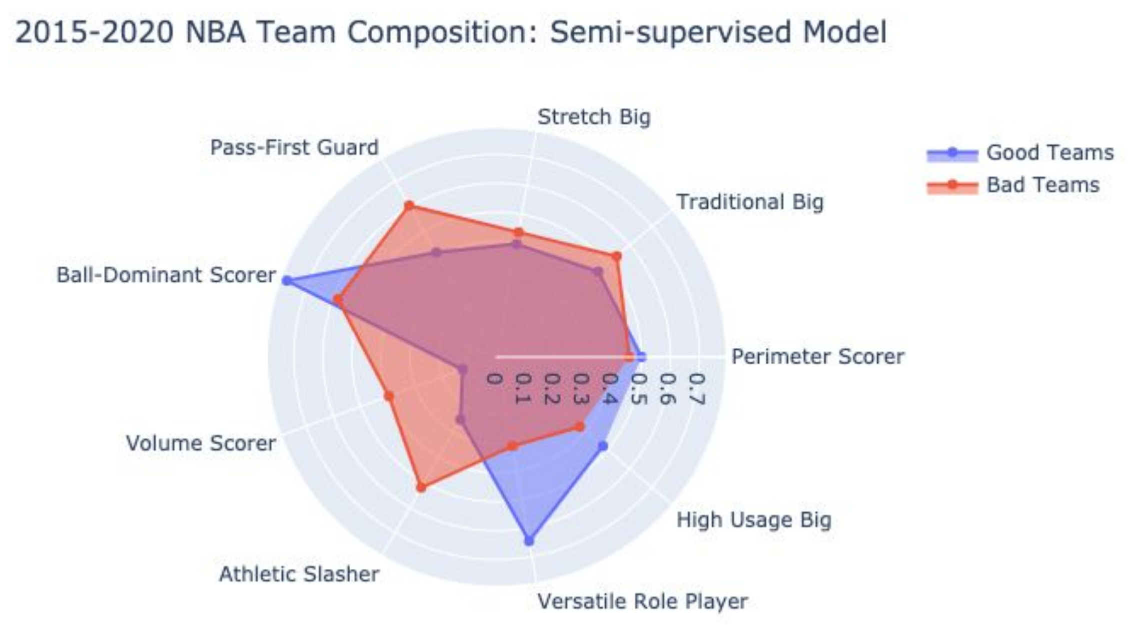
Weighted: Independently calculates metrics for each class, then averages them after giving weights corresponding to each class' proportion in the dataset

```
Accuracy:  0.8208

Micro Precision:  0.8208
Micro Recall:  0.8208
Micro F1:  0.8208

Macro Precision:  0.8165
Macro Recall:  0.8123
Macro F1-score:  0.8060

Weighted Precision:  0.8244
Weighted Recall:  0.8208
Weighted F1-score:  0.8170
```

MODEL RESULTS: SEMI-SUPERVISED

Comments:

- Good teams have more
  - Ball-dominant scorers
  - Versatility in role players
- Bad teams have more
  - Slashers
  - Volume Scorers
  - Pass-First Guards



MODEL RESULTS: UNSUPERVISED

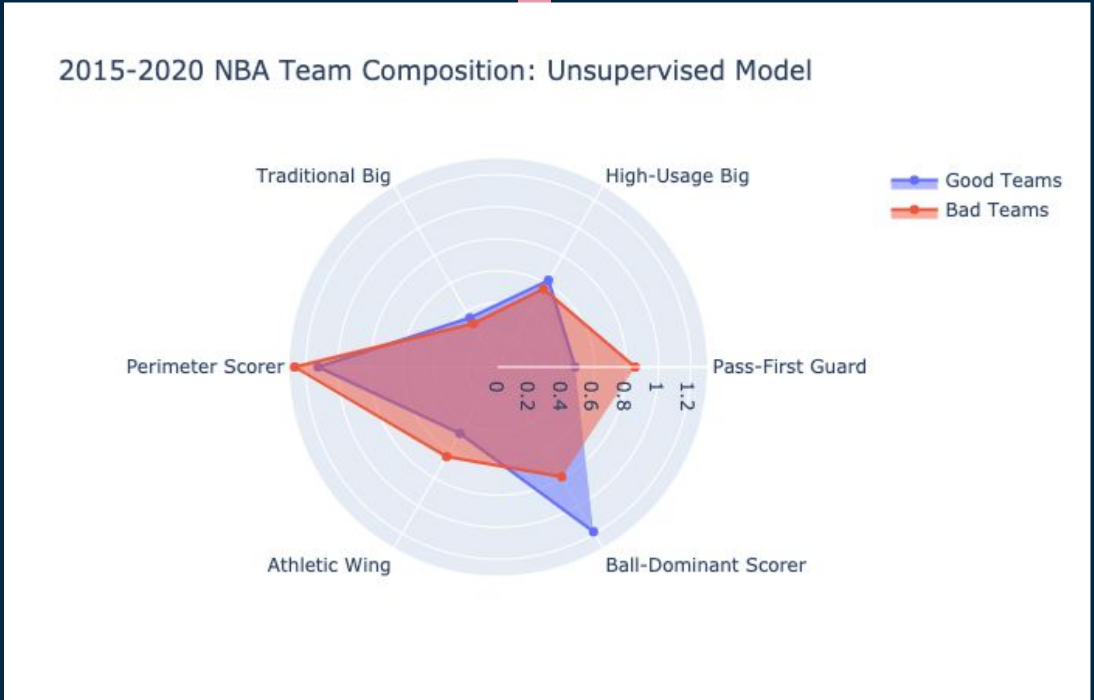
Wanted more than 5 classifications, ended up with 7 total

- Role Player
- Traditional Big
- High-Usage Big
- Pass-First Guard
- Ball-Dominant Scorer
- Athletic Wing
- Perimeter Scorer

MODEL RESULTS: UNSUPERVISED

Comments:

- Good teams have more
  - Ball-dominant scorers
- Bad teams have more
  - Athletic Wings
  - Pass-First Guards



FUTURE WORK

DASHBOARD: CLASSIFIER

Used to generate predicted playstyle based on user-inputted stats

DASHBOARD: PLOTTING

Create interactive tool that allows users access radar plots by toggling years and team-type (good, bad, average)

DEFENSIVE CLASSIFICATION

Most of these classifications mainly concern offense. More defensive stats could show which positions need the best defense.

Do you have any questions?

hoganbyun@gmail.com
410 236 4663

THANK YOU!

CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, and infographics & images by Freepik