
Predict NBA Playoffs Results With Machine Learning

Group 1

Mike Ryu

Nestor Aviles

Tyler Ingebrigtsen

Will Hua

Overview

Part 1: Intro

Part 2: Data Collection and Preparation

Part 3: Model Training and Testing

Part 4: GUI

Part 1: Intro

Goal:

Using Machine Learning to predict the 2020-2021 NBA Playoffs Game Results

Data we used:

- 2008-2021 NBA Teams Regular Season Stats
- 2008-Current NBA Playoffs Schedule and Results
- 2008-Current NBA Playoffs Betting Money Line

Part 1: Intro

Some libraries we used:

Basketball_reference_scraper, Itertools, Tkinter, Sklearn

Algorithms we tested:

- Random Forest Classifier
- Decision Tree Classifier
- GaussianNB
- Logistic Regression

Part 2: Data Collection and Preparation

Data Collection:

- 2008-2021 NBA Regular Season Stats
 - nbastats.net
- 2008-Current NBA Playoffs Schedule and Results
 - Basketball_reference_scraper
- 2008-Current NBA Playoffs Betting Money Line
 - sportsbookreviewsonline.com

* We started with nba_api library to get the data we wanted but the library was lack of documentation and updates, nbastats.net contains all the stats we wanted and can be easily downloaded as a CSV file.

Part 2: Data Collection and Preparation

Data Preparation:

- Data Cleaning:
Since we got our data from different source and a span of 12 years, we spent a lot of time concatenating, merging, filtering and sorting our data with Pandas and Vlookup function in Excel.
- Data Splitting:
 - Training: 2008-09 --- 2019-20
 - Testing: 2020-21

Part 3: Model Selection, Training and Performance

Model Selection:

Supervised learning → **Classification or Regression**

Output → **Class** → **Classification or Logistic Regression**

Classification : Random Forest Classifier, GaussianNB, Decision Tree Classifier

Logistic Regression : Logistic Regression

Part 3: Model Selection, Training and Testing

Model Training and Testing

x_train: 2008-2009 --- 2019-2020 NBA Playoffs face-off teams' regular season stats and Home Team's game betting money line.

y_train: 2008-2009 --- 2019-2020 NBA Playoffs game results

x_test: 2020-2021 NBA Playoffs face-off teams' regular season stats and Home Team's game betting money line.

y_test: 2021 NBA Playoffs game results

Part 3: Model Selection, Training and Testing

Accuracy Scores:

- Random Forest Classifier :0.671
- Decision Tree Classifier : 0.557
- GaussianNB :0.456
- Logistic Regression: 0.595

Other data for reference:


From 1983-84 to 2018-19, home teams won 65% of their games in the playoffs.

Best published model had an accuracy of 74.4%, it's a Maximum Entropy Model.

Part 4: GUI

NBA Predictor

2020-21 NBA Playoffs Game Result Predictor



Away

Home

Money Line


0


Predict

Part 4: GUI

NBA Predictor

2020-21 NBA Playoffs Game Result Predictor



Away	Home	Money Line
MIL	PHO	-230
		Win!!!
<div>Predict</div>		

Future Considerations

1. Add Series Score

- We wanted to add series scores for each series, because in fact, performance under pressure is mediated by several factors that may have an effect on players' mindset, such as leading or losing a series in a potential decisive game or playing in front of supportive audiences (Tauer, Guenther, & Rozek, 2009). This inherent pressure of decisive moments tends to increase self-awareness and conscious attention, which may lead to performance decrements (Beilock & Gray, 2007; Wallace, Baumeister, & Vohs, 2005). Under this scope, game-related statistics have been used to evaluate technical and tactical behaviours within different sporting contexts in basketball (Sampaio, et al., 2015), and are considered relevant when predicting the game outcome (Percy, 2015).

Future Considerations

2. Instead of predicting win or loss, we can try to predict the W/L spread. Probably we will use a linear regression model.
3. Add the stats/data of each team's key players.
4. Consider adding NLP into our machine learning model.
 - NBA, as a for-profit organization, if ever controlled the game results, popularity/profitability/ratings of the franchises and players would be in their considerations.

Works Cited

- Nuno Mateus, Bruno Gonçalves, Eduardo Abade, Nuno Leite, Miguel Angel Gomez, and Jaime Sampaio, EXPLORING GAME PERFORMANCE IN NBA PLAYOFFS,
https://www.researchgate.net/publication/322530345_Exploring_game_performance_in_NBA_playoffs
- Tristan Smith, And the 2020 NBA Champion is...,
<https://tjlsmith0831.medium.com/and-the-2020-nba-champion-is-7fa1a60c7fae>
- NBA Teams Regular Season Stats, <http://nbastats.net>
- NBA Playoffs Schedule and Results, <https://www.basketball-reference.com/>
- NBA Playoffs Betting Money Line, <https://sportsbookreviewsonline.com>

Thank you!