NBA Game Prediction using Machine Learning techniques

September 27, 2022

Project Synopsis

Machine Intelligence

BACHELOR OF TECHNOLOGY-V Sem CSE
Department of Computer Science & Engineering

SUBMITTED BY

Name: Neeraj Gopalakrishnan	SRN: PES1UG21CS824
Name: Pracheth Thamankar	SRN: PES1UG20CS469
Name: Tushar N Kumar	SRN: PES1UG20CS473

1 Abstract and Scope

The sports market has exploded in the 21st Century with the introduction of new media formats such as streaming broadcasting and social media. With the growth of sports and the monetary value associated with it sports analytics and sports team management has become a popular topic.

Game outcome prediction plays a major role in sports performance analysis. Game prediction significantly influences many parts of the sports market such as viwerbase,team management and strategy and also sports betting.

One of the most popular sports that lures betting and attracts millions of fans worldwide is basketball, particularly the National Basketball Association (NBA) of the United States.Total revenue across the organization is estimated to have reached $8.76 billion in the 2018-2019 season with each team being worth atleast $1 billion [1].

The primary goal of this project is to build a machine learning model to predict outcome of NBA games. Efficient predictor model can be built and improved on from existing models such as clustering, classification and regression. The Model's goal is to accurately predict the outcome of the game given the characteristics of the game. These characteristics can include a wide variety of factors ranging from individual player statistics such as PER (Player Efficiency Rating) [3], Player season performance, Recent Player performance,Offensice productive efficiency to complicated engineered features such Elo rating [2], [4]. The Machine learning model can be trained on existing datasets consisting of these variables to predict game outcome for unseen values of home and away team statistics.

2 Feasiblity Study

Predicting game or season outcomes is important for clubs as well as for the betting industry. Understanding the critical factors of winning games and championships gives clubs a competitive advantage when selecting players for the team and implementing winning strategies.

3 Design Approach

3.1 Gathering the Data

There is an abundance of data related to NBA statistics. Some of the best sources for Basketball data are the official NBA site `https://www.nba.com`, sports data analytics sites such as `https://www.basketball-reference.com/`(basketball-reference.com) and `www.kaggle.com`

3.2 Data pre-processing

The gathered dataset may have inconsistencies, missing values, noisy data. Further the data can be in formats not suitable for analytics such as categorical/nominal variables. This has to be first converted to dummy variables representing numerical data. Python and popular python libraries such as pandas, numpy, seaborn can be used for cleanind and pre-processing

3.3 Researching the model that will best fit the data

NBA game prediction can be done my modelling the data against conventional supervised learning models such as Linear and Logistic regression model, Neural Networks. Random Forest regression, Naive Bayes and Maximum Entropy principle models have also been used specifically for sports analytics and game-outcome prediction.

3.4 Training, Testing and Evalution

The data is split into three datasets 'Training', 'Validation', and 'Testing'. Accurate Models can be built through K-Fold Cross validation techniques The Best models can be found by comparing evalution metrics such as accuracy, precision, recall etc.

References

[1] https://www.investopedia.com/articles/personal-finance/071415/how-nba-makes-money.asp

[2] https://towardsdatascience.com/predicting-the-outcome-of-nba-games-with-machine-learning

[3] https://en.wikipedia.org/wiki/Player_efficiency_rating

[4] https://projects.fivethirtyeight.com/complete-history-of-the-nba/#warriors