

NBA PREDICTIONS

MIGUEL PÉREZ CARO

Abstract

In this research I want to make a model which predict the champion of the NBA based on the winner of the MVP weeks. I used the dataset NBA player of the week from the website kaggle, and I used a DecisionTree model to solve my research question. Finally, I discovered that the accuracy of the model is high but during the year many players from different teams win the MVP of the week therefore the task to predict the champion based on the week MVP is really difficult.

Motivation

I'm a big fan of the NBA, and today, the NBA is considered as the best league in the world, and is dominated by many star players in different teams and a team that has several all-stars, so I asked myself if in spite of the great variety of amazing players that are capable of winning the MVP of the week, at the end of the season based on these data one could predict which would be the NBA champion. The final result is improvable, but it is a model with an acceptable accuracy and that can be used to predict the final result of the playoffs after the regular season.

Dataset(s)

I chose from the website kaggle the dataset `NBA_player_of_the_week.csv` where we can find from 1985 to 2017, the MVP winner of every week, with the exact date, the season, the position, the weight, the height and the team. I decided to use only the team and the season, and I added a column which is called `NBAchampion` storing the winner of the NBA per year.

Data Preparation and Cleaning

First of all, I checked that all the columns were complete, and I deleted those that were not going to be necessary. I also had to change the format to the Season column to be able to operate with it. I did all this, with functions studied during the course such as `dropna ()`, `isnull ()` or `any ()`. The only problem I had was that I needed the champion per year to be able to create the model, for which I created a table with the year and the champion and I filtered the `NBA_player_of_the_week` dataset for every year adding in all rows the corresponding champion.

Research Question(s)

Can you predict the NBA champion based on the MVP champions of the week?

Methods

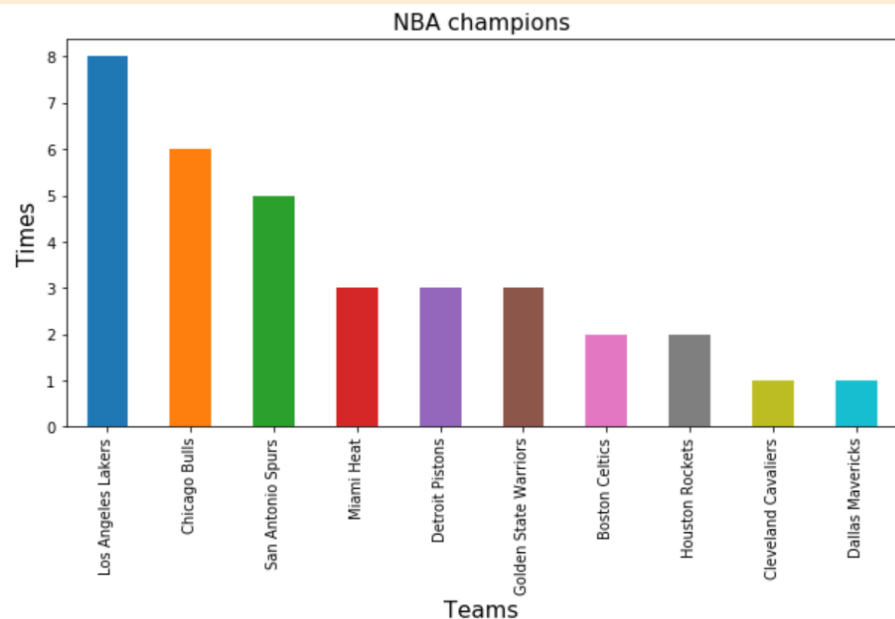
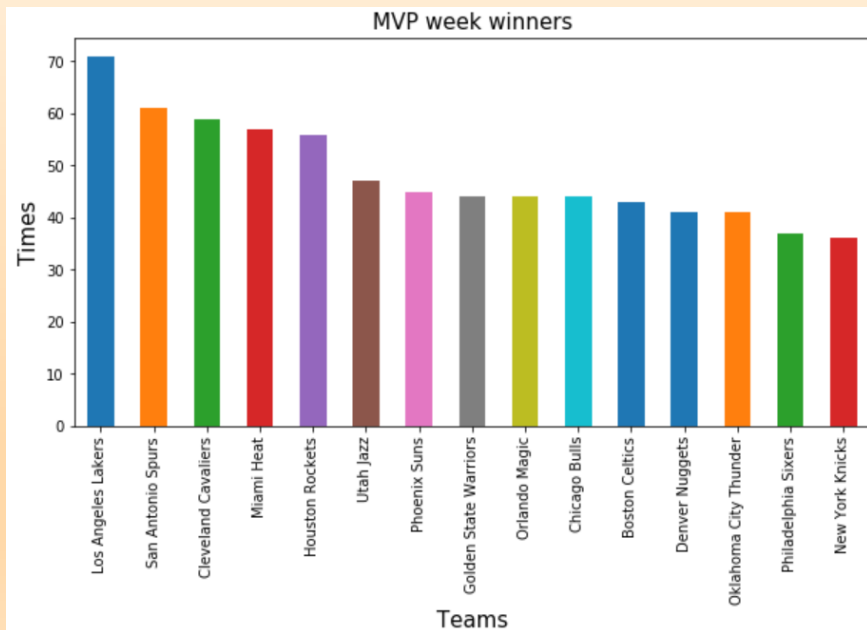
First of all, I checked if the datasets are cleaned with the function `isnull().any().`, and I printed the first 5 rows to check which column is interesting for me, deleting which are not interesting.

Once I had my dataset cleaned I made the graphs in order to see which player have win the MVP week more times, which team, and the position of the player. With this data I had an overview of the dataset.

Later, I added the NBAchampion dataset and I made the graph which I compared with the others, which is very useful to analyze research question because I saw how many times a team have won the NBA and the MVP week, so I think is the most appropriate way to analyze my question.

Findings I

Once my dataset was ready and I added the NBA champion column, I made the graphs in order to compare them.



Findings II

After that, I mapped the data, and created the model, which gave me the following result:

```
In [67]: accuracy_score(y_true = Y_test, y_pred = predictions)
```

```
Out[67]: 0.7724867724867724
```

The accuracy is not very high but as we can see on the previous slide, for example, Miami won only one title but Miami is the fourth team that has won the most MVP of the week.

Limitations

The great limitation of my model, is that as happened in some year, during the regular season a player stands out above the rest by winning MVP of the week many times, but finally the team of that player does not win the NBA.

Conclusions

My final conclusion is that the NBA champion can be predicted with high accuracy based on the winners of the week's MVP, except when a player stands out during the regular season and during the playoffs he lowers the level.

Acknowledgements

I got my data from kaggle website and the nba champion from the NBA official website.

I did not get feedback.

References

I used the documentary from the Python for Data Science course and the information from the NBA official website.