NBA Game Prediction by Machine Learning

team members: Wang, Jiuqi. Lin, Tianze. Lyu, Yue.

October 2018

1 Introduction of NBA Game Prediction Program

1.1 Main Idea

Our program applies basic machine learning algorithms and statistical models to predict the most likely winner of a future NBA game based on the database of last year NBA regular season games for 30 NBA teams.

Dean Oliver is a statistician who is prominent contributor to the statistical evaluation of basketball. In 2004, Dean Oliver expanded upon his "Four Factors" philosophy from his 2002 book, Basketball on Paper. In his book, he states that the result of an NBA game highly depends on four factors: Effective Field Goal Percentage, Turnover Rate, Offensive Rebounding Percentage and Free Throw Rate. Based on his theory of "Four Factors Philosophy", we can identify how four important strategies relate to success in basketball.

We also collected the data of 2017-2018 season from stats.nba.com/teams/four-factors. We found that the result of an NBA game is highly dependent on the ranking of the opponent. This statement can be illustrated by the following calculations.

Correlation: Game Result, Opponent Ranking

Correlations

Pearson correlation	0.404
P-value	0.000

The p-value is close to 0, so we fail to reject the null hypothesis, and the correlation coefficient is statistically significant. The correlation coefficient is positive (0.404), so the chance of winning a game gets higher as your opponent ranks lower.

Furthermore, from the data of www.statista.com, we found that the ranking of a team is correlated to the salary of a team.

Correlation: Rank, Salary

Correlations

Pearson correlation	0.648
P-value	0.000

The rank gets higher if the team is paying more salaries to players. The rank also reflects the power of a team, which is related to the result of a game.

From our previous observations, we found out 6 possibly most important factors influencing the success in a NBA basketball game.

1. **Effective Field Goal Percentage (EFG%)**
2. **Turnover Rate (TOV%)**

3. **Offensive Rebounding Percentage (OREB%)**
4. **Free Throw Rate (FTR)**
5. **Rank**
6. **Team Total Salary**

Thus, in our database, we recorded the relative advantage of two participant teams (denoted as team A and team B) as a Boolean variable (1 for A is higher, 0 for B is higher) and the result of the game.

1.2 Detailed Math Theory and Statistics Mechanism

We try to apply Naïve Bayes theorem to train the computer about our data set and predict the most likely winner in a future NBA game. We also integrate Laplace Smoothing to avoid outliers.

Observe: discrete variable X=<X1,X2,...Xm> each with 2 outcomes.

Output: discrete output Y with 2 outcomes(team A wins, or team A loses).

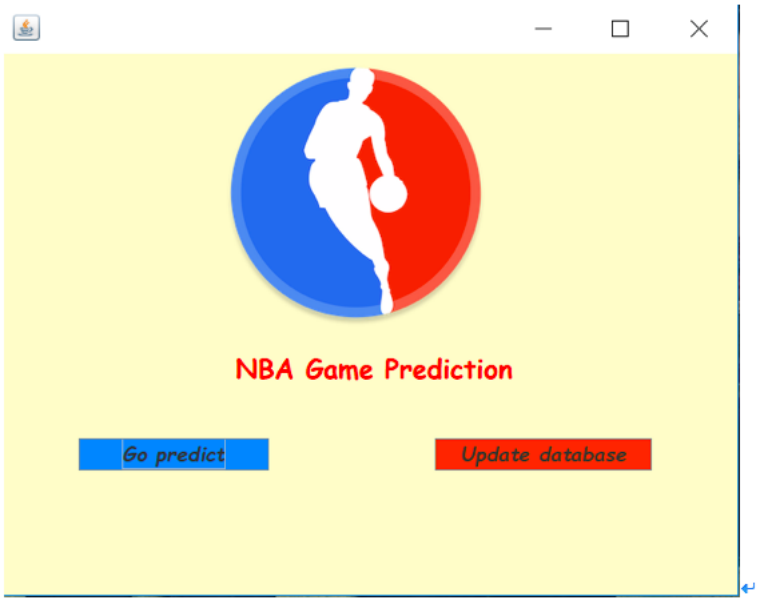
Classification: $\hat{Y} = argmax_y \hat{P}(Y|X) = argmax_y \hat{P}(X,Y) = argmax_y \hat{P}(X|Y)\hat{P}(Y)$.

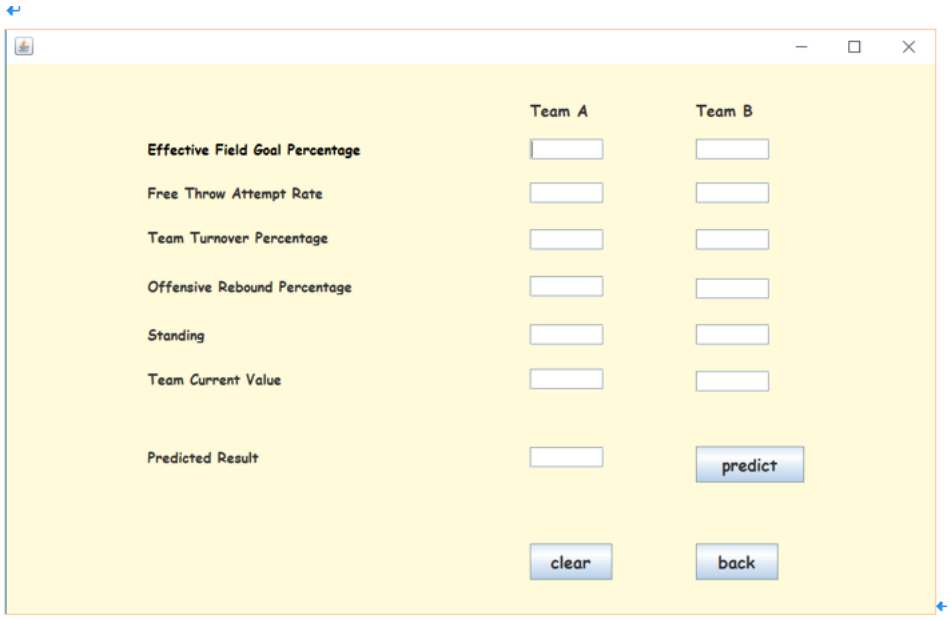
We have Y0="A wins the game" and Y1="A loses the game". We apply the above formula on our database to find a Y that can maximize $argmax_y \hat{P}(X|Y)\hat{P}(Y)$, then $argmax_y \hat{P}(Y|X)$.

The user can also update the data base based on the new known NBA game results and team information. The algorithm will build stronger by itself and have more precise predictions later.

2 NBA Game Prediction User Guide

1. Run the program by a Java IDE.
2. You will see the main menu. If you want to predict a future NBA game, click "Go predict".

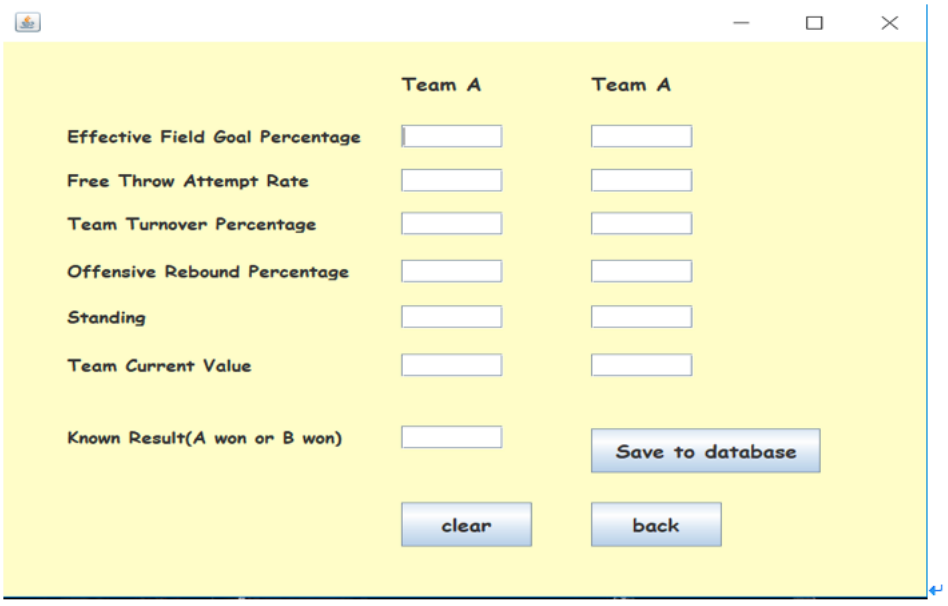




3.Input numbers into the boxes. You can find Effective Field Goal Percentage, Turnover Rate, Offensive Rebounding Percentage and Free Throw Rate for each team from *https://stats.nba.com/teams/four-factors*. You can find the standing from *https://www.basketball-reference.com/leagues/NBA-2019-standings.html*. Please input the rank in the Expanded Standing section of this site. The current value of a team is available at *https://www.basketball-reference.com/contracts/*.

4.Click "predict", and you will see the predict result. If you want to make another prediction, click "clear" and repeat step 3. If you want to go to the main menu, click "back".

5.If you want this software to give better predictions, you can update the database after a game. In the main menu, click "Update database".



6.Input the numbers into the boxes and click "Save to database". The prediction will become more accurate if there are more data in the database. If you want to save another game result into the database, click "clear" and input the values again. Click "back" if you want to go back to the main menu.