

Predicting Season Outcomes for the NBA

Teno González Dos Santos Chunyan Wang
Niklas Carlsson Patrick Lambrix

Linköping University, Sweden

Abstract. 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. In this paper, we work with NBA data from 10 seasons and propose an approach for predicting game outcomes that is then used for predicting which team will be champion and which stages a team will reach in the playoffs. We show that our approach has a similar performance as the odds from betting companies and does better than ELO.

1 Introduction

In many sports, work has started on predicting game or season outcomes. From an entertainment point of view, this is important considering the amount of money spent on betting. For clubs, understanding the critical factors of winning games and championships is important for creating a competitive team and implementing winning strategies. This paper focuses on such predictions for the National Basketball Association (NBA).

Most of the work on predicting game or season outcomes for the NBA uses box score information. The Four Factors (effective field goal percentage, turnovers per possession, offensive rebounding percentage, and free throw rate, e.g., [8, 3]) which have an offense variant and a defense variant, are used as a basis in [9, 1]. In [6], 18 box score features and information about wins and losses were used for 778 games. The Naive Bayes-based method reached 67% accuracy for game outcome. Several neural networks were trained on data from 620 NBA games using 11 box score statistics in [4]. The best networks had a prediction accuracy of 74%. A Maximum Entropy principle-based approach used on data from 7 seasons obtained an accuracy of 74% [2]. In [10], data was collected from the NBA finals 1980-2017 and 22 mainly box score features were used. The most significant feature influencing game outcome was deemed to be defensive rebounds. Other important factors were three-point percentage, free throws made, and total rebounds. A method taking into account team strength with attention to home court advantage and back-to-back games is proposed in [5]. Different approaches tested on 8 seasons have a prediction accuracy between 66% and 72% for regular seasons and between 64% and 79% for playoffs. The progression of a basketball game is modeled by a Markov model using play-by-play data in [12] and by a probabilistic graphical model based on play-by-play data and tracking

data in [7]. Play-by-play data is also used for learning stochastic models for substitutions. In all cases, the models are used for game outcome prediction. There is also work on predicting the outcome of basketball games in other leagues, but techniques may need adjustment to be transferable between leagues (e.g., [11]).

In this paper, we propose an approach for predicting which team will become NBA champion and to which stage of the NBA playoffs a team will proceed. The data that we use is from 10 seasons of NBA games and is presented in Sect. 2. We first introduce an approach for game outcome prediction (Sect. 3). This approach is then used to simulate NBA seasons and to derive frequencies over 10,000 simulations for teams reaching the different stages of the playoffs or become NBA champion (Sect. 4). We show that our approach has a similar performance as the odds from betting companies and significantly outperforms ELO. The paper concludes in Sect. 5.

2 Data collection and preparation

We gathered data from 10 complete NBA seasons from 2008-2009 to 2017-2018. All the extracted information comes from web-scraping <https://www.basketball-reference.com/>, a website specialized on NBA stats. The site includes box scores providing information relevant to a team’s performance in a single game, including well-known performance measures such as points, assists, and rebounds, as well as performance data on team level and information on the current regular season record prior to a game. Also information about salaries, draft picks and performance during previous seasons is available.

Table 1 summarizes the kind of data that we used.¹ Team victory is the objective variable. It takes a value of 1 in case the team has won the current match. This is the value to be predicted by the different classification models. For the collected team data we have standardized the team names. Thus, the teams which have changed their denominations in the previous 10 seasons have been converted to their current team names, e.g., the New Jersey Nets are denoted as the Brooklyn Nets. Our approach for season prediction involves simulating the seasons using a game outcome model for each game and then updating the information for the next game. Therefore, we use only stats in the box scores that can be derived from the game outcome. This means that stats such as assists, blocks, and points are not used.

From the **box scores** we retained information about the games regarding which team is the home team, at which stage of the season the game is played and how many earlier games were played in that stage, how many wins and losses the team had up to the current game in the regular season or in a playoff round, whether the team won the last game, the number of wins and losses in the last 3, 8, 15 games, home games and away games, and whether the game is a back-to-back game. The latter is important as the performance of players usually decreases when playing consecutive games in such a short time period

¹ Explanations of all features can be found in the appendix.

[4]. The previous 3, 8, 15 games take into account the recent performance of the team. We also look at sequences of home and away games as teams often have road trips and time periods with many consecutive home games in a row.

For **team performance in previous seasons** we gathered information on the stage that the team reached, the regular season record, the offensive rating in terms of points scored per 100 possessions, and the defensive rating in terms of points allowed per 100 possessions. We also collected the Four Factors metrics.

The **performance of individual players** has an impact on the team performance. This is particularly true in sports such as basketball, where there are only five players per team on the court at each moment in time and the top players often play the majority of the game. Due to the top players' significantly impacting the outcome of games, many NBA teams prioritize trying to recruit two or three top players to their roster. These players are often referred to as the "Big Two" or the "Big Three", and are generally considered the most important players for team success. An example of a high impact player is LeBron James. Before arriving to the Cleveland Cavaliers in 2006 (after being drafted), Cleveland had never won the NBA championship and performed poorly on a regular basis. After his arrival, they reached the playoffs for 5 consecutive years until his move to Miami in 2010 with an NBA final in 2007. The team did not qualify for the playoffs again until his return to the team in 2014, when they played four consecutive finals and won the title in 2016. During his four years in Miami, he also made it to the finals each year (and won two championships), while forming a feared "Big Three" together with Dwyane Wade and Chris Bosh. We collected data about the performance of players using a variant of eWS48 which is an estimate of the number of wins contributed by a player per 48 minutes (total time played in a game without overtime). The average value in the league is around 0.100. We normalized this by multiplying by the minutes played during the season and divided by the total number of games in the season (82) and the number of minutes in a standard game (48). We then aggregated player performances to a team level. We used information on the mean eWS48 for returning players (staying with the team) and players leaving and joining the team.

The features related to **player salaries** represent how much a team pays their players, how this quantity relates to the salary cap imposed by the league, and the importance of key players based on how much they are paid. The total salary - salary cap ratio can be a critical factor, since spending more money usually leads to better players on the roster. However, if a team pays their players over the salary limit, they need to pay also a luxury tax, which could influence the team's future development. The importance of the salary of the top players can be exemplified by the fact that, according to <https://hoopshype.com/salaries/>, in the 2008-2009 season the Boston Celtics paid 61 MUSD, i.e., 77% of the salary, only to 3 players. In general, at least half of the teams during each of the seasons considered in this paper spent over 50% of the salary to 3 players.

The features for the **NBA draft picks** represent the draft picks made by the teams in the previous 5 years. The draft is organized in two rounds of (usually) 30 players. Usually, the earlier the player gets picked, the better his expected

performance is. However, this has not always been the case, as several 1st draft picks left the league after a few years, due to injuries or poor performance.

3 Game Outcome Prediction

3.1 Methods

A first step in our approach is to compute a model for game outcome prediction. We used four different techniques: Logistic Regression (LR), Linear Support Vector Machines (LSVM), Random Forest (RF) and Multilayer Perceptron (MLP). For each of these techniques we did hyperparameter tuning to find the best fit to the data. Furthermore, when appropriate, we selected the features for the different algorithms that resulted in the best accuracy which is the ratio of correct predictions to all predictions or $(TP + TN) / (TP + TN + FP + FN)$, where TP is the number of true positives, TN the number of true negatives, FP the number of false positives and FN the number of false negatives. For every combination of hyperparameters and features, we fit the model and predict a season based only on the data from previous seasons, and report the averages using the 10 different resulting accuracies.

3.2 Results

For LR, we used a grid of values to tune the hyperparameter C, which stands for the inverse of regularization strength (see Fig. 1). The best accuracy was obtained by the model with $C = 0.1$, with a mean test accuracy score of 68.58%.

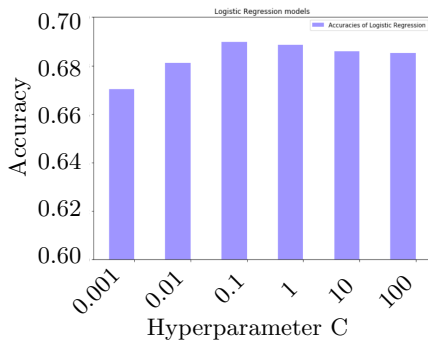


Fig. 1. Accuracies for the different models of LR, with different C.

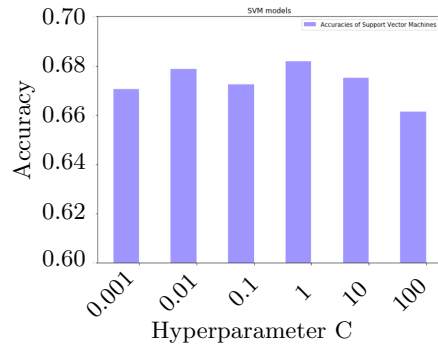


Fig. 2. Accuracies for the different models of LSVM, with different C.

For LSVM, we tried to optimize the C parameter, which adds a penalty for each misclassified data point (see Fig. 2). The best accuracy was obtained by the model with $C = 1$, with a mean test accuracy score of 68.18%.

Table 1: Features.

Box score data per game	<p>Home team Season stage Games played in Regular Season Wins in League Record Losses in League Record Games played in current play-offs round Wins in current play-offs round Losses in current play-offs round Won Last game Won Last Home game Won Last Away game Wins in previous 3, 8 and 15 games Wins in previous 3, 8 and 15 home games Wins in previous 3, 8 and 15 away games Back-to-back game</p>
Team performance in previous season	<p>Previous season furthest stage Previous season regular season record Previous season offensive rating Previous season defensive rating Offense Four Factors: eFG%, TOV%, ORB%, FT/FGA Defense Four Factors: eFG%, TOV%, DRB%, FT/FGA</p>
Player performance in previous season	<p>Staying players weighted mean eWS48 Signed players weighted mean eWS48 Leaving players weighted mean eWS48</p>
Player Salaries	<p>Total Salary Total Salary / Salary Cap Ratio Top-1 player salary ratio Top-2 players salary ratio Top-3 players salary ratio Top-5 players salary ratio</p>
NBA draft picks	<p>Previous season draft picks in positions 1 to 3 Previous season draft picks in positions 4 to 10 Previous season draft picks in positions 11 to 20 Previous season draft picks in positions 21 to end of 1st round Previous season draft picks in 2nd round Previous 3 seasons draft picks in positions 1 to 3 Previous 3 seasons draft picks in positions 4 to 10 Previous 3 seasons draft picks in positions 11 to 20 Previous 3 seasons draft picks in positions 21 to end of 1st round Previous 3 seasons draft picks in 2nd round Previous 5 seasons draft picks in positions 1 to 3 Previous 5 seasons draft picks in positions 4 to 10 Previous 5 seasons draft picks in positions 11 to 20 Previous 5 seasons draft picks in positions 21 to end of 1st round Previous 5 seasons draft picks in 2nd round</p>
Objective variable	<p>Team Win</p>

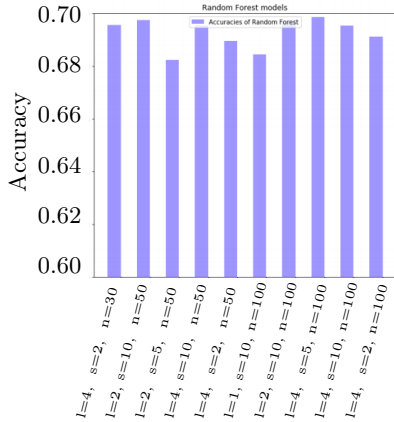


Fig. 3. Top 10 accuracies for the different models of Random Forest with different combinations of hyperparameters (min_samples_leaf, min_samples_split, and n_estimators).

For MLP, we used sets of different values for the different hyperparameters. We used single hidden layer networks with 50, 100, or 180 neurons in each layer. For the regularization term alpha (L2) we used 0.0001, 0.01, and 0.05. As activation functions we used hyperbolic tangent function, logistic sigmoid function and rectified linear unit function. The learning rate for the schedule for weight updates was kept constant at 0.001 or adaptive which kept the learning rate constant at 0.001 as long as training loss kept decreasing. Further, we used SGD and Adam for weight optimization. The best accuracy was achieved by the model with hyperbolic tangent function as the activation function, alpha = 0.05, a single hidden layer with 100 neurons, Adam solver and an adaptive learning rate (Fig. 4). This combination had a mean test accuracy score of 68.85%.

For RF, we used sets of different values for the different hyperparameters. For the number of estimators representing the number of trees in the forest we used the values 10, 15, 20, 30, 50 and 100. The minimal number of samples required to split a node was set to 2, 5 and 10, while the minimum number of samples in a leaf node was set to 1, 2 and 4. The top 10 accuracies are shown in Fig. 3. The best accuracy was achieved by the model with number of estimators = 100, minimum of samples in a leaf = 4 and minimum of samples in a split = 5. This combination got a mean test accuracy score of 69.88%.

The representative for RF obtained the best result. This was the model that we selected to use in the the season simulations.

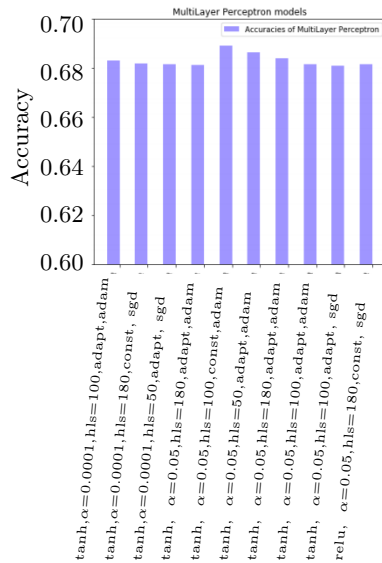


Fig. 4. Top 10 accuracies for the different models of Multilayer Perceptron with different combinations of hyperparameters. (Models used: activation, tanh/relu, alpha, hidden_layer_size (hls), learning_rate, constant/adaptive, solver, adam/sgd).

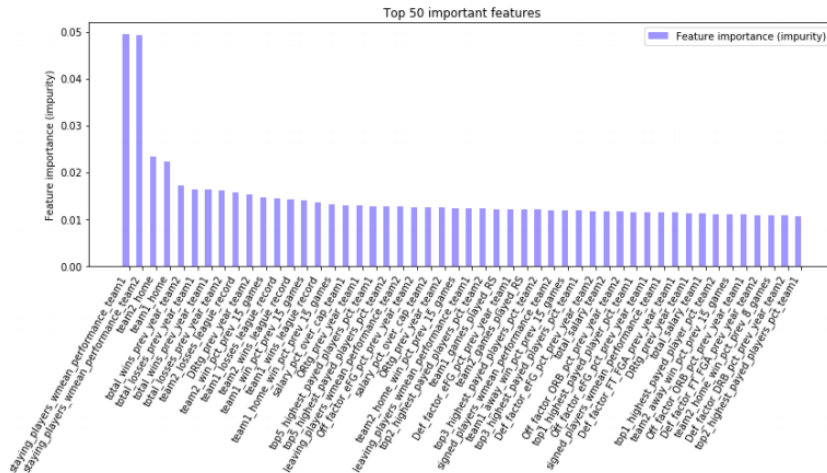


Fig. 5. Feature importance for chosen RF model.

In Fig. 5 we show the 50 most important features with respect to Gini impurity for the chosen model. The most relevant features are the performances during the previous season of the players that stayed with the team. Further, whether a team is the home team in a game is important. This suggests a home team advantage. The rest of the top-50 most important features have relatively similar values. Among these, there are wins and losses in the current and previous seasons. Regarding the last games, it is more important to look at the last-15 games than the other values we looked at (3 and 8). Further, there are some features related to the total salary of a team, the percentage over the salary cap and the salary of the top-2 and top-3 players. Also factors regarding team performance (offense and defense) from the previous years and regarding the performance of leaving and signed players appear in this top-50 list.

4 Season simulation

We simulated 10 complete seasons from 2008-2009 to 2017-2018 using the chosen RF model. Since we had the actual schedule of the regular season from each year, we could simulate the calendar in the same order as it occurred in reality. For every season and every game in the calendar, we predicted the output probabilities of each team to win. During the simulation, we used these probabilities to draw a random number between 0 and 1 uniformly. If the draw landed between 0 and the probability of a team winning, the victory is assigned to the team, otherwise the win went to the opponent. Upon the assignment of the win we updated the values of the dynamic features in order to prepare the input for the upcoming games. Once the whole regular season was simulated, the playoffs started. At this stage, we simulated the playoff series as a means to pick the best team from each playoff matchup until a single team became the NBA champion. This simulation process was repeated 10,000 times in order

Table 2: Predictions for the 2017-2018 season.

	Team	1st Round (Model)	Conf. Semifinals (Model)	Conf. Finals (Model)	NBA Finals (Model)	NBA Champion (Model)	Reality (furthest stage)	ELO season start	ELO season end
1	GSW	88.4	73.1	47.9	28.9	20.8	NBA champion	1752	1745
2	CLE	86.5	58.0	37.8	22.3	16.1	NBA finals	1650	1577
3	HOU	83.5	51.6	34.1	20.9	14.6	Conf. finals	1574	1704
4	TOR	85.5	56.1	36.9	20.8	10.4	Conf. semifinals	1532	1600
5	SAS	86.3	64.5	35.7	17.4	8.6	1st round	1617	1551
6	BOS	81.6	50.5	29.8	12.2	7.3	Conf. finals	1532	1580
7	NOP	79.8	57.0	29.8	9.6	4.1	Conf. semifinals	1488	1585
8	UTA	67.5	46.1	19.8	7.3	2.5	Conf. semifinals	1580	1663
9	OKC	65.9	30.8	15.6	6.0	2.2	1st round	1518	1611
10	POR	58.9	30.6	13.1	5.6	1.5	1st round	1531	1579
11	PHI	79.2	33.1	14.0	5.7	1.4	Conf. semifinals	1380	1641
12	WAS	59.0	26.1	12.4	4.8	1.2	1st round	1566	1499
13	MIA	86.6	37.5	13.5	4.4	1.2	1st round	1553	1497
14	IND	54.0	21.7	8.0	3.1	1.0	1st round	1503	1572
15	MIN	60.6	18.9	6.8	2.3	1.0	1st round	1474	1548
16	MIL	39.2	15.4	6.6	2.1	0.7	1st round	1508	1522
17	DET	36.6	12.9	5.3	1.7	0.6	9th East conf.	1457	1488
18	LAC	36.2	13.8	4.7	1.5	0.5	10th West conf.	1591	1506
19	CHO	39.7	10.9	3.7	1.1	0.4	10th East conf.	1473	1501
20	CHI	30.3	14.8	4.0	1.0	0.3	13th East conf.	1497	1317
21	NYK	22.8	8.9	2.8	0.8	0.3	11th East conf.	1407	1378
22	DEN	41.5	13.9	3.7	1.1	0.2	9th West conf.	1540	1587
23	DAL	21.4	6.4	1.7	0.4	0.2	13th West conf.	1441	1357
24	MEM	45.5	10.3	2.1	0.4	0.1	14th West conf.	1489	1322
25	SAC	50.4	10.5	2.4	0.4	0.1	12th West conf.	1421	1360
26	ORL	15.8	4.0	1.2	0.3	0.1	14th East conf.	1390	1335
27	BRK	16.2	4.2	1.2	0.3	0.0	12th East conf.	1405	1408
28	LAL	20.5	4.1	0.9	0.2	0.0	11th West conf.	1401	1486
29	ATL	17.9	4.0	0.8	0.2	0.0	15th East conf.	1486	1349
30	PHO	13.8	5.1	0.9	0.1	0.0	15th West conf	1381	1277

to obtain not only the winning frequencies of each team to become the NBA champion, but also for reaching the different stages of the competition. The whole simulation process was performed for every season 2008-2009 to 2017-2018. To keep consistency in our predictions, we trained our model only on the seasons previous to the one that we were simulating. Table 2 shows the results for the 2017-2018 season. The complete results for the 2008-2009 to 2017-2018 seasons are available at <https://www.ida.liu.se/research/sportsanalytics/projects/conferences/MLSA21-basketball/>. In addition to the predictions of our method, we have also added information about the teams' ELO at the start and end of the season. ELO data was obtained from <https://projects.fivethirtyeight.com/complete-history-of-the-nba>.

Table 3 shows the prediction success of the method over the 10 seasons. We say that a prediction is correct for a team and a season regarding one of the

Table 3: Prediction success. For all stages, the first/second number is the number of correct predictions using our approach (first) and ELO (second) at the start of the season. For the NBA champion, the third number shows the success based on the pre-season odds. (* Two teams with same odds of which one was champion.)

Season	1st Round	Conf. Semifinal	Conf. Final	NBA Final	NBA Champion
2008-2009	13/12	6/3	3/1	1/1	0/0/0.5*
2009-2010	13/13	6/4	3/2	1/1	1/1/1
2010-2011	12/10	3/2	1/0	0/0	0/0/0
2011-2012	13/11	5/5	3/2	1/1	1/0/1
2012-2013	14/12	5/6	3/2	2/2	1/0/1
2013-2014	12/11	5/5	4/3	2/2	0/0/0
2014-2015	12/11	4/4	0/1	0/0	0/0/0
2015-2016	12/12	5/4	2/2	1/1	0/0/0
2016-2017	13/13	5/3	3/3	2/1	1/0/1
2017-2018	15/13	7/4	2/2	2/2	1/1/1
Total	129/118	51/40	24/18	12/11	5/2/5.5
out of	160	80	40	20	10

stages NBA Champion, NBA Final, Conference Final, Conference Semifinal and 1st Round, if the prediction score for the team reaching the stage is among the 1, 2, 4, 8, 16 highest, respectively, for the season. Further, we compare with the ELO at the start of the season and for the NBA Champions also with the pre-season odds at <https://www.basketball-reference.com/>. The Spearman correlation of our prediction scores and ELO at the start of the season for NBA Champion ranges from 0.71 to 0.96. For the other stages NBA Final, Conference Finals, Conference Semifinals and 1st Round, these ranges are 0.72 to 0.95, 0.71 to 0.95, 0.73 to 0.92 and 0.69 to 0.92, respectively (Table 4). The highest correlation for each stage is for the 2016-2017 season, while the lowest is for the 2017-2018 season. Note that for all stages our approach outperforms the ELO approach. We obtain the same predictions as the odds-based approach for all seasons except 2008-2009 where two teams had the same lowest odds.

Table 4: Spearman correlation between prediction score and ELO at start of season.

Season	1st Round	Conf. Semifinal	Conf. Final	NBA Final	NBA Champion
2008-2009	0.8432529	0.8525648	0.9086255	0.9178359	0.9181914
2009-2010	0.8868365	0.9048626	0.9370062	0.9313884	0.9361926
2010-2011	0.8196685	0.8908666	0.9139869	0.9081600	0.9253556
2011-2012	0.8265658	0.8661698	0.8985088	0.8830014	0.9045688
2012-2013	0.8792701	0.8787002	0.8971390	0.8662733	0.8859901
2013-2014	0.9031038	0.9065421	0.9143112	0.9048204	0.9203673
2014-2015	0.7259177	0.7872719	0.8261741	0.8330925	0.8301203
2015-2016	0.8700490	0.8929446	0.8981637	0.8956920	0.9083453
2016-2017	0.9209033	0.9202181	0.9565992	0.9540246	0.9622899
2017-2018	0.6908444	0.7329773	0.7199778	0.7292364	0.7118506

5 Conclusion

In this paper, we first proposed an approach for game outcome prediction that reached a mean accuracy of 69.88%. The most relevant features in the model are found to be the performances during the previous season of the players that stayed on the team as well as whether a team plays at home. Other important features are wins and losses in the current (last 15 games) and previous seasons, offensive and defensive performance from previous years, performance of signed and leaving players, and salary features. Second, we then used this approach to simulate 10 NBA seasons 10,000 times and computed frequencies for teams reaching different stages in the playoffs. We showed that the approach was equally successful in picking a Champion as the odds makers and consistently outperformed ELO for all playoff rounds (except one 2014-2015 round). Future work will investigate whether the approach is equally successful for other sports.

References

1. Baghal, T.: Are the "four factors" indicators of one factor? an application of structural equation modeling methodology to NBA data in prediction of winning percentage. *Journal of Quantitative Analysis in Sports* **8**(1) (2012). <https://doi.org/doi:10.1515/1559-0410.1355>
2. Cheng, G., Zhang, Z., Kyebambe, M., Kimbugwe, N.: Predicting the outcome of NBA playoffs based on the maximum entropy principle. *Entropy* **18**, 450:1–15 (2016). <https://doi.org/10.3390/e18120450>
3. Kubatko, J., Oliver, D., Pelton, K., Rosenbaum, D.T.: A starting point for analyzing basketball statistics. *Journal of Quantitative Analysis in Sports* **3**(3) (2007). <https://doi.org/10.2202/1559-0410.107>
4. Loeffelholz, B., Bednar, E., Bauer, K.W.: Predicting NBA games using neural networks. *Journal of Quantitative Analysis in Sports* **5**(1) (2009). <https://doi.org/doi:10.2202/1559-0410.1156>
5. Manner, H.: Modeling and forecasting the outcomes of NBA basketball games. *Journal of Quantitative Analysis in Sports* **12**(1), 31–41 (2016). <https://doi.org/doi:10.1515/jqas-2015-0088>
6. Miljković, D., Gajić, L., Kovačević, A., Konjović, Z.: The use of data mining for basketball matches outcomes prediction. In: *IEEE 8th International Symposium on Intelligent Systems and Informatics*. pp. 309–312 (2010). <https://doi.org/10.1109/SISY.2010.5647440>
7. Oh, M., Keshri, S., Iyengar, G.: Graphical model for basketball match simulation. In: *MIT Sloan Sports Analytics Conference* (2015)
8. Oliver, D.: *Basketball on Paper*. Brassey's (2004)
9. Teramoto, M., Cross, C.L.: Relative importance of performance factors in winning NBA games in regular season versus playoffs. *Journal of Quantitative Analysis in Sports* **6**(3) (2010). <https://doi.org/doi:10.2202/1559-0410.1260>
10. Thabtah, F., Zhang, L., Abdelhamid, N.: NBA game result prediction using feature analysis and machine learning. *Annals of Data Science* **6**, 103–116 (2019). <https://doi.org/10.1007/s40745-018-00189-x>
11. Zimmermann, A.: Basketball predictions in the NCAA and NBA: Similarities and differences. *Statistical Analysis and Data Mining* **9**(5), 350–364 (2016). <https://doi.org/10.1002/sam.11319>

12. Štrumbelj, E., Vračar, P.: Simulating a basketball match with a homogeneous markov model and forecasting the outcome. *International Journal of Forecasting* **28**(2), 532–542 (2012). <https://doi.org/10.1016/j.ijforecast.2011.01.004>

Appendix - Features

Table 5: Features - 1.

Box score data per game	
Home team	1 if team is home team, 0 if team is away team.
Season stage	One of: regular season, 1st round, conference semi-finals, conference finals, NBA final.
Games played in Regular Season	Amount of games played by the team up to, but not including the current game during regular season. The value is set to 82 during play-offs.
Wins in League Record	Number of wins up to current game during the regular season. Not updated during playoffs.
Losses in League Record	Number of losses up to current game during the regular season. Not updated during playoffs.
Games played in current play-offs round	Number of games played by the team up to, but not including the current game during each play-off round. The value is reset to 0 at the beginning of each playoff round. The value is set to 0 during the regular season.
Wins in current play-offs round	Number of wins by the team up to, but not including the current game during each play-off round. The value is reset to 0 at the beginning of each playoff round. The value is set to 0 during the regular season.
Losses in current play-offs round	Number of losses by the team up to, but not including the current game during each play-off round. The value is reset to 0 at the beginning of each playoff round. The value is set to 0 during the regular season.
Won Last game	1 if team won the last game; 0 otherwise.
Won Last Home game	1 if team won the last home game; 0 otherwise.
Won Last Away game	1 if team won the last away game; 0 otherwise.
Wins in previous 3, 8 and 15 games	Number of wins during the previous 3, 8 and 15 played games by the team
Wins in previous 3, 8 and 15 home games	Number of wins during the previous 3, 8 and 15 played home games by the team
Wins in previous 3, 8 and 15 away games	Number of wins during the previous 3, 8 and 15 played away games by the team
Back-to-back game	1 if the team has played a game within the last 36 hours; 0 otherwise.

Table 6: Features - 2.

Team performance in previous season	
Previous season furthest stage	One of: not qualified for play-offs, 1st round loss, conference semi-finals loss, conference finals loss, NBA final loss or NBA champion.
Previous season regular season record	Number of wins and losses during the previous regular season.
Previous season offensive rating	Estimated amount of points scored in 100 possessions in the previous season.
Previous season defensive rating	Estimated amount of points allowed in 100 possessions in the previous season.
Offense Four Factors: eFG%, TOV%, ORB%, FT/FGA	Effective Field Goals percentage, Turnovers committed per 100 plays, Percentage of available Offensive Rebounds, Free Throws per Field Goal attempt.
Defense Four Factors: eFG%, TOV%, DRB%, FT/FGA	Opponent effective Field Goals percentage, Turnovers caused on the opponent per 100 plays, Percentage of available Defensive Rebounds, Opponent Free Throws per Field Goal attempt.

Table 7: Features - 3.

Player performance in previous season	
Staying players weighted mean eWS48	Weighted mean performance of the players that remained in the team from the previous season.
Signed players weighted mean eWS48	Weighted mean performance of the players that joined in the team after the previous season.
Leaving players weighted mean eWS48	Weighted mean performance of the players that left the team after the previous season.

Table 8: Features - 4.

Player Salaries	
Total Salary	Sum of the salaries of all the players in the team.
Total Salary / Salary Cap Ratio	Ratio between total salary payed by a team and the salary limit established by the league.
Top-1 player salary ratio	Ratio between the salary of the top player and the total salary of the team.
Top-2 players salary ratio	Ratio between the sum of the salaries of the top 2 players and the total salary of the team.
Top-3 players salary ratio	Ratio between the sum of the salaries of the top 3 players and the total salary of the team.
Top-5 players salary ratio	Ratio between the sum of the salaries of the top 5 players and the total salary of the team.

Table 9: Features - 5.

NBA draft picks	
Previous season draft picks in positions 1 to 3	Number of draft picks in positions 1 to 3 during previous season
Previous season draft picks in positions 4 to 10	Number of draft picks in positions 4 to 10 during previous season
Previous season draft picks in positions 11 to 20	Number of draft picks in positions 11 to 20 during previous season
Previous season draft picks in positions 21 to end of 1st round	Number of draft picks in positions 21 to end of 1st round during previous season
Previous season draft picks in 2nd round	Number of draft picks in 2nd round during previous season
Previous 3 seasons draft picks in positions 1 to 3	Number of draft picks in positions 1 to 3 during previous 3 seasons
Previous 3 seasons draft picks in positions 4 to 10	Number of draft picks in positions 4 to 10 during previous 3 seasons
Previous 3 seasons draft picks in positions 11 to 20	Number of draft picks in positions 11 to 20 during previous 3 seasons
Previous 3 seasons draft picks in positions 21 to end of 1st round	Number of draft picks in positions 21 to end of 1st round during previous 3 seasons
Previous 3 seasons draft picks in 2nd round	Number of draft picks in 2nd round during previous 3 seasons
Previous 5 seasons draft picks in positions 1 to 3	Number of draft picks in positions 1 to 3 during previous 5 seasons
Previous 5 seasons draft picks in positions 4 to 10	Number of draft picks in positions 4 to 10 during previous 5 seasons
Previous 5 seasons draft picks in positions 11 to 20	Number of draft picks in positions 11 to 20 during previous 5 seasons
Previous 5 seasons draft picks in positions 21 to end of 1st round	Number of draft picks in positions 21 to end of 1st round during previous 5 seasons
Previous 5 seasons draft picks in 2nd round	Number of draft picks in 2nd round during previous 5 seasons