Journal of Computer Science and Information Technology
June 2017, Vol. 5, No. 1, pp. 1-14
ISSN: 2334-2366 (Print), 2334-2374 (Online)

Published by American Research Institute for Policy Development
DOI: 10.15640/jcsit.v5n1a1
URL: https://doi.org/10.15640/jcsit.v5n1a1

Predicting National Basketball Association Game Attendance Using Random Forests

Barry E. King[1]

Abstract

Research predicting National Basketball Association game attendance using a random forest approach is presented. Attendance and other data obtained for the 2009 through 2013 basketball seasons are used. Predictor variables include: home team popularity, popularity of opponent, match type (regular season or playoff), day of the week on which the match occurs, home team winning percentage, home city's total personal income, capacity of home venue, conference of the home team, lagged variables on attendance and on winning percentage, and others. A random forest approach, using the R statistical modeling language, was selected in order to use numerous predictor variables without having to first deselect variables and not to over-fit the data. The random forest prediction is compared favorably with that of a multiple linear regression. Additional results indicate that some variables suggested by sports writers do not contribute much to the prediction and that a better measure of a team's popularity is needed.

Keywords: ensemble method, R, regression

Professional basketball team managers need to forecast attendance at matches to plan staff, decide on promotions, and estimate revenues. Numerous predictor variables come to mind from both academic literature and from the popular press. Because there is a large set of predictor variables, it is easy to over fit the data. One way to avoid over fitting is to use a prediction technique that minimizes it. Random forest satisfies this need or requirement. It is employed in this research purposely to avoid over fitting the data. In a random forest, each node is split using the best among a subset of predictors randomly chosen at that node.

[1] Butler University, Barry E. King, Lacy School of Business, Butler University, 4600 Sunset Avenue, Indianapolis, Indiana 46208, USA. Phone: (317) 940-5464, email: king@butler.edu

This counterintuitive strategy turns out to perform very well compared to many other classifiers, including discriminate analysis, support vector machines, and neural networks(Liaw & Wiener, 2002).The first objective of this paper is to demonstrate that using an ensemble technique such as random forest provides better forecasts than traditional multiple linear regression. The second objective is to demonstrate the use of this technique for forecasting attendance at future events. The contributions to the literature are 1) presenting a data mining predictor technique for attendance at future events, and 2) developing an accessible technique using open source software, the statistical language R. This paper is structured as follows: 1) review of literature, 2) method, 3) results, and 4) conclusion.

Review of Literature

Mills and Salaga (2011) report on the use of ensemble methods in sports research. Random forest is an ensemble method. Villar and Guerrero (2009) provide a thorough review of the literature during the 1973-2007 period.

Prediction Factors

Hansen and Gauthier (1989) provide an early assessment of factors affecting sporting event attendance. Borland and MacDonald (2003) review sources and determinants of the demand for professional sporting contests.

Basketball. Factors influencing attendance at basketball matches have been reported by Berri, Schmidt, and Brook, 2004; Leadley and Z ygmont, 2005;Pecha and Crossan, 2009 and Zhang et al. 1995. Deshpande and Jensen (2016) comment on a star player's impact on winning while Jane (2014a) comments on a star player's impact on attendance. Liu (2015) also remarks that a player's star effect and popularity can attract more attendance at a match. Gladden and Funk (2001) point out that winning may not be a significant predictor of attendance among highly committed fans, the ability of a team to entertain is critical. Gladden and Funk (2001) note there is a positive effect on attendance because of fan loyalty. Entertainment may involve cheering squads and crowd participation games. Snipes and Ingram (2007) remark that promotions have a positive impact on attendance. Mongeon and Winfree (2012) report that winning is more important to a television audience than to match attendance. Jane (2014b) comments that "Closer wins by the competing teams within a league and a larger gap in terms of the point spread between two teams in the betting market lead to higher attendance."

Baseball. Beginning with the work of Demmert (1973), many studies have examined the factors believed to impact gate attendance in professional baseball, including promotions (Boyd and Krehbiel, 1999; Hill, Madura, and Zuber, 1982.

McDonald and Rascher (2000), new stadiums (Clapp and Hakes, 2005), work stoppages (Coates and Harrison, 2005; Schmidt and Berri, 2002), and team success (Baade and Tiehan, 1990; Gitterand Rhoads, 2010).Layson and Rhodes (2011) report that "doubleheaders have a very positive effect on attendance on the day of the doubleheaders but that this is substantially offset by reduced attendance at single games three days surrounding doubleheaders."

Hockey. Leadley and Zygmont, 2006; and Winfree and Fort, 2008, report on factors impacting hockey. In their study, Paul and Weinbach (2011) found that "Fan demand for Canadian Hockey junior league level of hockey is found to be sensitive to the success of the home team and to exhibit normal consumer responses to weekday and monthly effects with weekends being more popular and attendance increasing throughout the season toward the playoffs. On-ice factors such as scoring, a proxy for excitement, and fighting are not shown to have a significant effect on attendance."

Rugby. Using data from 1,226 matches played over 18 seasons, Hogan, Massey, and Massey (2014) analyze match attendances in the group stages of the European Rugby Cup (ERC). They find that short-run (match) uncertainty had little effect on attendances. This finding is significant as the ERC has been replaced by a new competition which may be more unbalanced due to differences in the distribution of revenue between the participating teams. Medium-term uncertainty, i.e. the possibility of the home team reaching the knock-out stages, had a significant impact on attendances. Measures designed to make matches more attractive, e.g. bonus points for high scoring, had little effect.

Soccer. Jewell and Molina, 2005 have studied attendance at soccer matches. Ferreira and Bravo (2007) report "Results regarding team success, team division, population, stadium size and habitual persistence were found to influence professional soccer attendance; other factors such as admission price, age of team, international success, availability of soccer teams in the same vicinity and stadium ownership did not."

Marketing

Coates, Humphreys, and Zhou (2012) develop a consumer choice model of live attendance at a sporting event with reference-dependent preferences. The predictions of the model motivate the "uncertainty of outcome hypothesis" (UOH) as well as a fan's desire to see upsets and to simply see the home team win games.

Method

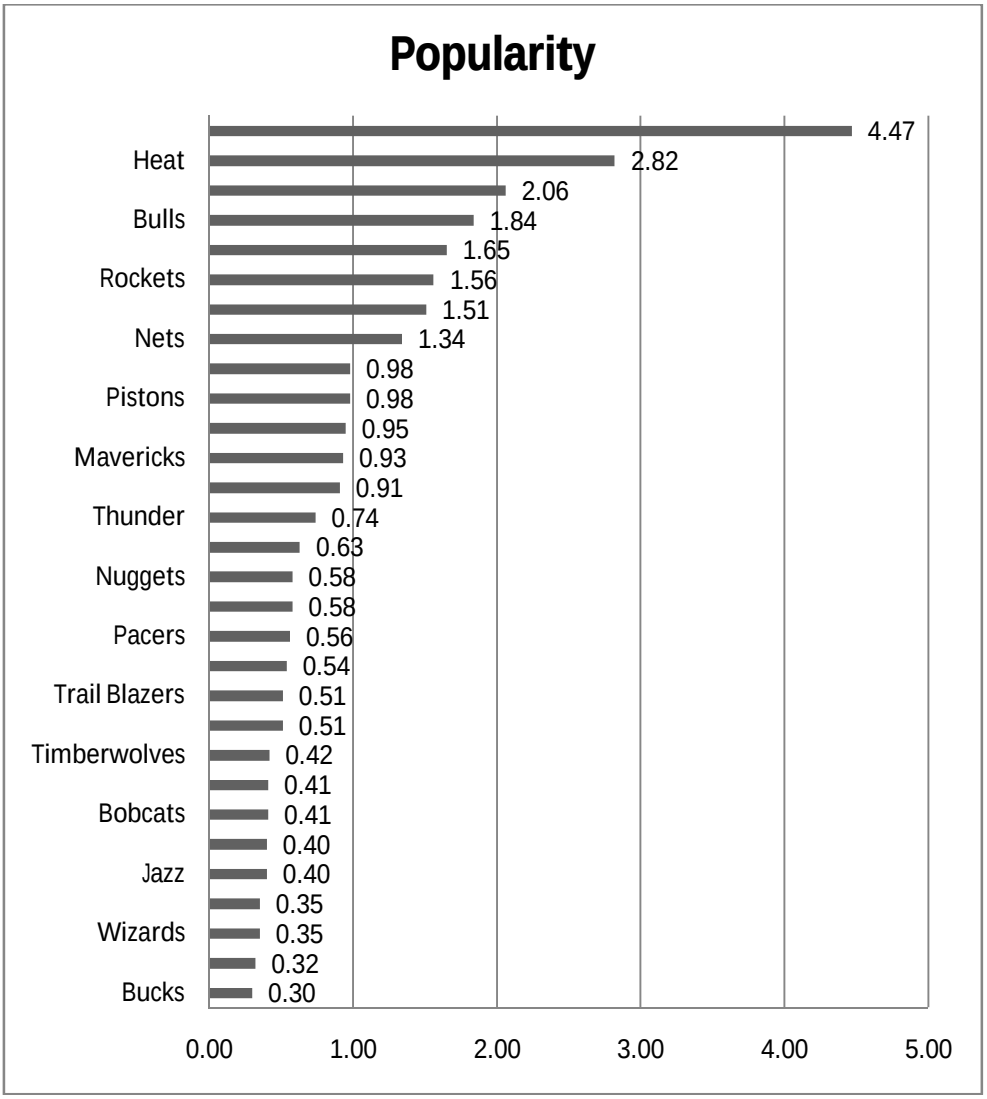
Research Question

Can attendance at National Basketball Association (NBA) games be forecast with lower error by using the random forest technique than by using multiple linear regression? Traditional linear regression necessitates tuning the model to have the proper repressors. This requires skill on the part of the analyst to avoid multicollinearity and over fitting the data. The data mining technique random forest is robust in that it does not require preselecting the covariates. This is the main reason random forest was selected as the regression technique. It does, however, require some tuning. The optimal number of predictors to try at each split needs to be decided and the number of trees to build requires input.

Data

Approximately 6300 records were obtained for all NBA matches from the 2009 season through the 2013 season. These records were obtained from Qualex Consulting Services, Inc. of North Miami, Florida. A record includes home and away team names, home and away scores, venue, match type (regular or playoff), date, and attendance. Additional data were constructed for Capacity of a venue, from Wikipedia. Team's conference, obtained by searching the NBA Internet site. Total personal income of a team's city, calculated by multiplying per capita income by population as found on the U.S. Census Bureau's quick facts site. Relative popularity surrogate score of a team, based on the number of searches for the team on Google obtained from Google Trends (see Figure 1).

Figure 1: Relative popularity of teams based on searches on Google.



Prediction Variables

Twelve repressor were selected for this attendance prediction model:

- Home team popularity as measured by the relative number of searches on Google. Some home teams are more popular than others as shown in Figure 1.

- The popularity of the opponent influences a patron's desire to attend a match. Popular opponents bring more patrons to a home game (Neuteufel, 2014).
- The day of the week affects attendance. Patrons may be more likely to attend a weekend match than a weekday match.
- Match type, whether regular season or playoff. The data suggests that playoff matches are almost always sellouts.
- The season's winning percentage. It is thought that patrons are more likely to attend a winning team's match than a losing team's match (Neuteufel, 2014).
- Last season's winning percentage. This is a one season lagged variable on the home team's winning percentage.
- Attendance last home game. This is a one game lagged variable.
- Attendance two home games ago. This is a two-game lagged variable.
- Attendance last time the home team played this visitor. A popular opponent may draw more customers than an unpopular opponent.
- Capacity of the venue. Attendance cannot be greater than the capacity during playoffs and during other matches expected to attract large audiences.
- A city's total personal income. This can affect the number of patrons. Wealthier cities will have greater attendance than those cities that are not so wealthy (The worst NBA markets of 2012-2013, 2013).
- Conference. Eastern conference matches draw slightly more patrons than do western conference games (Neuteufel, 2014).

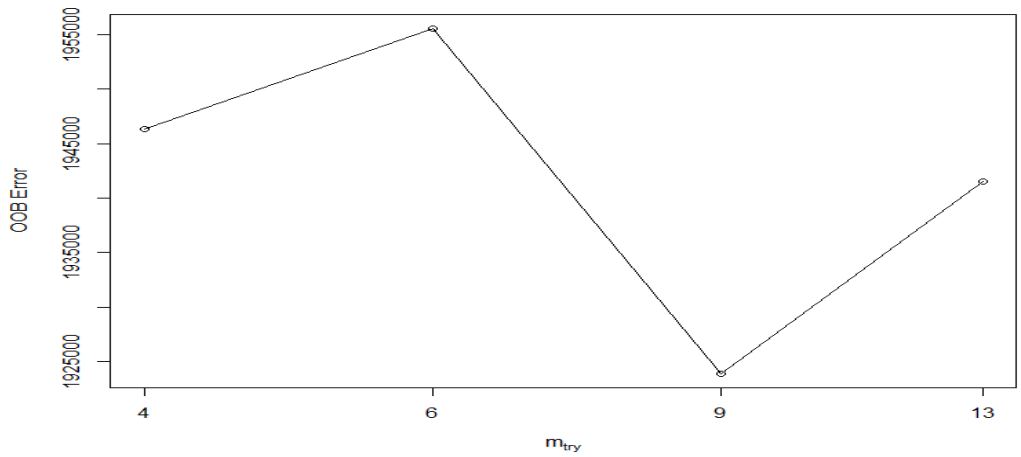
Variables not represented in the model include:

- A city's number of major league teams – baseball, football, hockey, and major league soccer – is not included. Previous research suggests the number of major league teams in a city is not an important factor (Neuteufel, 2014).
- Weather is not included. Weather has a greater impact on outdoor events than it does on an indoor event such as basketball.

Procedure

The procedure was to train the random forest on 75 percent of the data set and then test it on the remaining 25 percent. Random forest was chosen as the regression technique since the originators of random forests, Brieman and Cutler (n.d.), mention it can handle thousands of input variables without variable deletion. Cross-validation is not necessary with random forests. As pointed out by Topchef (n.d.)"…cross-validation isn't necessary as a guard against over-fitting. This is a nice feature of the random forest algorithm." The model was tuned to select the number of variables to try at each split in order to minimize out-of-bag error. As seen in Figure 2 the best number of variables is nine.

Figure 2: Out-of-bag error associated with the number of variables to try at each split. Nine produces the lowest error.



Results

Training Data Set

Running random forest against the training data set produced the results of Figure 3. Seventy-two percent of the variation is explained by the model and the Pearson correlation between observed training values and out-of-bag predictions is 0.85.

Figure 3: Results of training the random forest regression. Percent variation explained is 71.84.

```
               Type of random forest: regression
                     Number of trees: 1000
No. of variables tried at each split: 9

          Mean of squared residuals: 1995532
                    % Var explained: 71.84
```

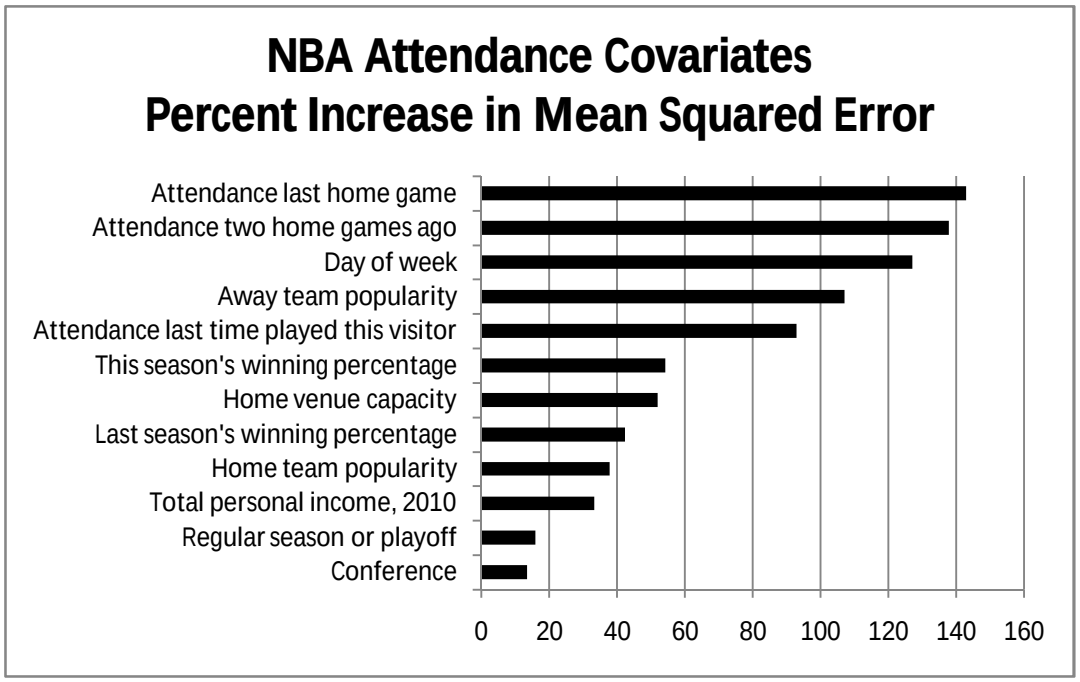
Test Data Set

The fitted model was then applied to the test data set yielding a mean absolute percent error (MAPE) of 5.77 with an over-predicting bias of 30 seats. It is desirable to have a near-zero bias. A thirty-seat bias for a 20,000-seat arena is a near-zero bias.

Full Data Set

Once the model was fitted and tested, all observations were then presented to the random forest algorithm. The results of re-building the model with all data resulted in the importance-of-covariates figure, Figure 4.

Figure 4: Results of random forest regression, importance of covariates, on the full data set.



%Inc MSE is the most robust and informative measure. It is the increase in mse of predictions (estimated with out-of-bag-CV) as a result of variable j being permuted (values randomly shuffled) the higher number, the more important. (Stackoverflow, n.d.) It is worth pointing out that total personal income of the home city and the conference, factors suggested by sports writers, turned out not to be very important in creating attendance predictions. A better measure of an away team's popularity may be worth developing since it is such an important covariate. As evidenced in Figure 4, the strongest drivers of attendance prediction are attendance last home game and attendance two home games ago.

Multiple Linear Regression

A multiple linear regression model was constructed using the same variables as those of the random forest model. This model can be examined in Figure 5.

Figure 5: Multiple linear regression model of NBA attendance

```
Coefficients:
                                      Estimate Std. Error t value Pr(>|t|)
(Intercept)                         -2.080e+03  4.877e+02  -4.265 2.04e-05 ***
home_team_popularity                 1.193e+02  3.001e+01   3.975 7.14e-05 ***
regular_season_or_playoffRegular Season  -1.614e+02  9.705e+01  -1.663   0.0964 .
lagged_attendance_1_period           3.108e-01  1.261e-02  24.644  < 2e-16 ***
lagged_attendance_2_periods          2.661e-01  1.261e-02  21.094  < 2e-16 ***
win_percentage                       2.209e+03  1.889e+02  11.694  < 2e-16 ***
day_of_weekMonday                   -9.594e+02  8.256e+01 -11.621  < 2e-16 ***
day_of_weekSaturday                  3.235e+02  7.817e+01   4.139 3.55e-05 ***
day_of_weekSunday                   -4.745e+02  8.388e+01  -5.657 1.63e-08 ***
day_of_weekThursday                 -5.329e+02  1.038e+02  -5.134 2.95e-07 ***
day_of_weekTuesday                  -1.129e+03  8.308e+01 -13.591  < 2e-16 ***
day_of_weekWednesday                -8.844e+02  7.063e+01 -12.523  < 2e-16 ***
lagged_win_percentage               -1.905e+02  1.864e+02  -1.022   0.3068
conferenceWest                       2.919e+02  5.178e+01   5.637 1.83e-08 ***
away_team_popularity                 4.514e+02  2.626e+01  17.187  < 2e-16 ***
total_personal_income_2010           4.570e-01  1.022e-01   4.470 8.00e-06 ***
capacity                             2.421e-01  2.627e-02   9.216  < 2e-16 ***
attendance_last_time_played_this_visitor  2.051e-01  1.159e-02  17.692  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1559 on 4728 degrees of freedom
Multiple |Adjusted R-squared:  0.6617
F-statistic: 546.9 on 17 and 4728 DF,  p-value: < 2.2e-16
```

Note that the adjusted R^2 is 0.667.

Random Forest versus Multiple Linear Regression

Both the random forest model and the multiple linear regression model were run with the test data set. Statistics from these runs are reported in Table 1.

Table 1 Comparison of Random Forest and Multiple Linear Regression Predictions

Statistic	Random Forest	Multiple Linear Regression
Mean Error (ME)	88.27	77.19
Root Mean Squared Error (RMSE)	1334.29	1500.85
Mean Absolute Percent Error (MAPE)	5.47	6.81

The two error vectors were supplied to the Diebold-Mariano test. The results appear in Figure 6.

Figure 6: Results of Diebold-Mariano test on the errors from the random forest technique and those from multiple linear regression.

```
        Diebold-Mariano Test

data:  diff.rf diff.lm
DM = -13.533, Forecast horizon = 1,
Loss function power = 1, p-value < 2.2e-16
alternative hypothesis: two.sided
```

The p-value is near zero so the null hypothesis of equal accuracy is rejected. Based on the root mean squared error in Table 1 and the result of the Diebold-Mariano test, we conclude that the random forest technique proves to be a more accurate prediction than the multiple linear regression technique.

Application

Predicting Future Attendance

The fitted random forest regression model can be saved and applied to a team's future data. This can be a useful tool to team and venue management for planning staffing, for arranging for inventory, and for scheduling promotions.

Example of Predicting Future Attendance

We use the Indiana Pacers 2013-2014 season as an illustration. In this example the Pacers have completed the season through December 31, 2013 and wish to predict the home attendance for the first six home matches in January 2014. Table 2 shows the data for December 2013. All the data from the 2009 season through December 2013, not just the December 2013 data, are used to build the random forest prediction model. The attendance column is the actual home attendance for the Pacers. The Pacers home venue, Bankers Life Field house, has a capacity of 18,165. Six of the seven home games in December 2013 reached this capacity.

Table 2 December 2013 Matches and Home Attendance for the Indiana Pacers

Date	Match Type	Home Team	Away Team	Attendance
12/1/2013	Regular Season	LAC	IND	
12/2/2013	Regular Season	POR	IND	
12/4/2013	Regular Season	UTA	IND	
12/7/2013	Regular Season	SAS	IND	
12/8/2013	Regular Season	OKC	IND	
12/10/2013	Regular Season	IND	MIA	18165
12/13/2013	Regular Season	IND	CAH	18165
2/16/2013	Regular Season	IND	DET	15443
12/18/2013	Regular Season	MIA	IND	
12/20/2013	Regular Season	IND	HOU	18165
12/22/2013	Regular Season	IND	BOS	18165
12/23/2013	Regular Season	BRK	IND	
12/28/2013	Regular Season	IND	BRK	18165
12/31/2013	Regular Season	IND	CLE	18165

Table 3 shows the input data for the first six Indiana Pacer home games of 2014. The date column allows derivation of day of week the game is to occur. Away team abbreviation allows looking up the visitor's popularity surrogate and deriving the attendance the last time the Pacer's played them.

Table 3 The Input Data for the First Six Pacer Home Games of 2014

Date	Match Type	Home Team	Away Team
1/4/2014	Regular Season	IND	NOP
1/7/2014	Regular Season	IND	TOR
1/10/2014	Regular Season	IND	WAS
1/14/2014	Regular Season	IND	SAC
1/16/2014	Regular Season	IND	NYK
1/8/2014	Regular Season	IND	LAC

Table 4 shows the result of the prediction.

Table 4 The Result of Applying the Model to the First Six Home Games of 2014 Along with Actual Attendance

Date	Match Type	Home Team	Away Team	Actual Attendance	Predicted Attendance	Percent Error
1/4/2014	Regular Season	IND	NOP	18165	16883	7.1%
1/7/2014	Regular Season	IND	TOR	16147	16086	0.4%
1/10/2014	Regular Season	IND	WAS	18165	17076	6.0%
1/14/2014	Regular Season	IND	SAC	17530	16257	7.3%
1/16/2014	Regular Season	IND	NYK	18165	17126	5.7%
1/8/2014	Regular Season	IND	LAC	18165	17412	4.1%
					MAPE	5.1%

The model tended to slightly under-predict attendance resulting in a MAPE of 5.1 percent.

Conclusion

Some of the prediction variables suggested by other writers, such as home city total personal income and conference, are not important to predicting attendance and that a better measure of a team's popularity is needed. Random forest regression can provide more accurate predictions than multiple linear regression. The use of random forest to predict attendance at future matches is a better technique. Random forest is an attractive data mining technique since it does not require preselecting the covariates and it naturally avoids the hazard of over fitting. Random forest does not require careful tuning as does multiple linear regression. The number of covariates to try at each split of the random forest algorithm is a tuning parameter. The optimal value can be obtained through an auxiliary routine of the random Forest package provided to the R statistical language.

Implications for Sport Managers

R is open source software. Team managers can acquire R and the R Studio integrated development environment or IDE at no financial cost. The random Forest package is freely available to R users. This makes applying the prediction technique presented in this article accessible to sport managers. Although this article presents attendance predictions for NBA teams, the general technique can be applied to other sport leagues such as collegiate basketball or Major League Soccer.

References

Baade, R. A., & Tiehen, L. J. (1990). An analysis of major league baseball attendance, 1969-1987. Journal of Sport & Social Issues, 14(1), 14-32.

Berri, D. J., Schmidt, M. B., & Brook, S. L. (2004). Stars at the gate the impact of star power on nba gate revenues. Journal of Sports Economics, 5(1), 33-50.

Borland, J., & MacDonald, R. (2003). Demand for sport. Oxford review of economic policy, 19(4), 478-502.

Boyd, T. C., & Krehbiel, T. C. (1999). The effect of promotion timing on major league baseball attendance. Sport Marketing Quarterly, 8(4), 23-34. (n.d.).www.stat.berkeley. edu /~breiman/RandomForests/cc_home.htm

Clapp, C. M., & Hakes, J. K. (2005). How long a honeymoon? The effect of new stadiums on attendance in Major League Baseball. Journal of Sports Economics, 6(3), 237-263.

Coates, D., & Harrison, T. (2005). Baseball strikes and the demand for attendance. Journal of Sports Economics, 6(3), 282-302.

Coates, D., Humphreys, B., & Zhou, L. (n.d.).Outcome uncertainty, reference-dependent preferences and live game attendance. Retrieved fromhttp://www.economics. ualb erta.ca/~/media/economics/FacultyAndStaff/WPs/WP2012-07-Coates-umphreys-Zhou

Demmert, H. G. (1973). The economics of professional team sports. Lexington, Mass: Lexington Books.

Deshpande, S. K., & Jensen, S. T. (2016). Estimating an NBA player's impact on his team's chances of winning. Journal of Quantitative Analysis in Sports, 12(2), 51-72.

Ferreira, M., & Bravo, G. (2007). A multilevel model analysis of professional soccer attendance in Chile 1990-2002. International Journal of Sports Marketing and Sponsorship, 8(3), 49-66.

Gitter, S. R., & Rhoads, T. A. (2010). Determinants of minor league baseball attendance. Journal of Sports Economics, 11(6), 614-628.

Gladden, J. M., & Funk, D. C. (2001). Understanding brand loyalty in professional sport: Examining the link between brand associations and brand loyalty. International Journal of Sports Marketing and Sponsorship, 3(1), 54-81.

Goleman, A. & Taylor, Z. (2012). The effects and consequences of advancing technology in the sports industry: the declining attendance rates at NFL games, Temple University School of Tourism and Hospitality Management, 2012.

Hansen, H. & Gauthier, R. (1989). Factors affecting attendance at professional sports events. Journal of Sport Management, 3(1), 15-32.

Hill, J. R., Madura, J., & Zuber, R. A. (1982). The short run demand for major league baseball. Atlantic Economic Journal, 10(2), 31-35.

Hogan, V., Massey, P., &Massey, S. (2014).Analys ingmatch attendance in the European rugby cup, University College Dublin Centre for Economic Research Working Paper Series.

Jane, W. J. (2014a). The Effect of Star Quality on Attendance Demand The Case of the National Basketball Association. Journal of Sports Economics, 1527002514530405.

Jane, W. J. (2014b). The relationship between outcome uncertainties and match attendance: New evidence in the National Basketball Association. Review of Industrial Organization, 45(2), 177-200.

Jewell, R. T., & Molina, D. J. (2005). An evaluation of the relationship between Hispanics and Major League Soccer. Journal of Sports Economics,6(2), 160-177.

Layson, S. & Rhodes, M. T. (2011).Were major league baseball doubleheaders a mistake?. Retrievedromhttp://bae.uncg.edu/assets/research/econwp/2011/layson_ Rhodes .pdf

Leadley, J. C., & Zygmont, Z. X. (2005). When is the honeymoon over? National Basketball Association attendance 1971-2000. Journal of Sports Economics, 6(2), 203-221.

Leadley, J. C., & Zygmont, Z. X. (2006). When is the honeymoon over? National Hockey League attendance, 1970-2003. Canadian Public Policy/Analyse de Politiques, 213-232.

Liaw, A. & Wiener, M. (2002) Classification and regression by random Forest," R News, 2/3.

Liu, Z. H. (2015). Star Effect on attendance in National Basketball Association.

McDonald, M., & Rascher, D. A. (2000). Does bat day make cents?: The effect of promotions on the demand for baseball. Journal of Sport Management, 14.

Mills, B. M., & Salaga, S. (2011). Using tree ensembles to analyze National Baseball Hall of Fame voting patterns: an application to discrimination in BBWAA voting. Journal of Quantitative Analysis in Sports, 7(4).

Mongeon, K., & Winfree, J. (2012). Comparison of television and gate demand in the National Basketball Association. Sport Management Review, 15(1), 72-79.

Neuteufel, N. (2014) Expected attendance in the NBA, Sports & Illumination. Retrieved fromhttp://sportsandillumination.com/2014/08/13/expected-attendance-nba/

Paul, R. & Weinback, A. (n.d.) Determinants of attendance in the Quebec major junior hockey league: role of winning, scoring, and fighting, Atlantic Economics Journal, 39(3) pp. 303-311.

Pecha, O., &Crossan, W. (2009). Attendance at basketball matches: a multilevel analysis with longitudinal data. Acta Kinesiologica, 3(1), 68-76.

Schmidt, M. B., & Berri, D. J. (2002). The impact of the 1981 and 1994-1995 strikes on Major League Baseball attendance: A time-series analysis. Applied Economics, 34(4), 471-478.

Snipes, R. L., & Ingram, R. (2007). Motivators of collegiate sport attendance: A comparison across demographic groups. Innovative Marketing, 3(2), 61.

Stack overflow (n.d.)http://stats.stackexchange.com/questions/162465/in-a-random-forest-is-larger-incmse-better-or-worse

Topchef (n.d.)http://stackoverflow.com/questions/19760169/how-to-perform-random-forest-cross-validation-in-r. para 12.

Villar, J. G., & Guerrero, P. R. (2009). Sports attendance: a survey of the literature 1973-2007. Rivista di Diritto e di Economia dello Sport, 5(2), 112-151.

Watts, T. L., Jr. & Bass, E.J. (2004) A regression-based predictive model of student attendance at UVA men's basketball games, Proceedings of the 2004 Systems and Information Engineering Design Symposium.

Winfree, J. A., & Fort, R. (2008). Fan substitution and the 2004-05 NHL lockout. Journal of Sports Economics.

The worst NBA markets of 2012-2013 (2013)The Wages of Wins Journal. Retrieved from http://wagesofwins.com/2013/04/02/the-worst-nba-markets-of-2012-13/

Zhang, J. J., Lam, E. T., & Connaughton, D. P. (2003). General market demand variables associated with professional sport consumption. International Journal of Sports Marketing and Sponsorship, 5(1), 24-46.

Zhang, J. J., Pease, D. G., Hui, S. C., & Michaud, T. J. (1995). Variables affecting the spectator decision to attend NBA games. Sport Marketing Quarterly, 4, 29-40.