Do Home Ground and Toss Win Make Cricket An Unfair Game? An Empirical Study from the First Fifteen Years of Indian Premier League

Ankita Nagpal*, Gaurav Nagpal**, Nitin Koshta*** and Shivang Singh****

Cricket is the most popular and most-watched game in India. However, the fairness of the game has been questioned many times in the past, on account of the possible advantage of home-ground and toss wins. Therefore, this research study aims to examine the influence of home ground and toss wins on the outcome of Indian Premier League matches over the first fifteen years of the IPL. The chi-square test of association suggests that these effects are statistically significant at a 90 percent confidence level for a few of the matches only, but not for others at the individual match level. However, at the aggregate level, a positive effect has been found by the logistic regression model that was used to predict the probability of winning for any team. The logistic regression model was developed to predict the match outcome based on these variables taken at an aggregate level for the three most prominent IPL teams using logistic regression, since there are other variables that can influence the performance, such as bowling performance, batting performance, etc. The study also mentions the implications of the findings for game players, sponsors, patrons, regulators, etc.

Keywords: Logistic Regression, home ground, toss win, predictive analytics, cricket, Indian Premier League.

1. Introduction

CRICKET is a widely watched sport and is becoming a rapidly growing industry in itself by managing to stay on top in terms of viewership and revenue generation across the world (Asif & McHale, 2016). The sport is growing

*   Ankita Nagpal, Guest Faculty, BITS Pilani.
** Gaurav Nagpal, Assistant Professor, BITS Pilani.
*** Nitin Koshta, Assistant Professor, IIM Amritsar.
****Shivang Singh, Software Engineer, Microsoft.

stronger day by day. For example, there are more than a billion cricket fans worldwide, and the fan base is constantly increasing (Kapadia *et al.*, 2020). Even countries such as China, which have no tradition of Cricket, are promoting the sport (Mani, 2009). In India, cricket has given continuity to the Indian ethnic identity (Nair, 2011; Devan, 2012). It has been a pre-eminent national game in India at the expense of many other sports such as hockey and football (Ray, 2008). The game also represents a spirit of nationalism among the Indian community (Kim Kyunghak, 2010).

The popularity of cricket further increased in India with the advent of the IPL (Indian Premier League) in the year 2008. In IPL, different states of the country are represented by different teams that compete for the title. The IPL, with the participation of international cricketers across the globe, made the game more elite and gave a significant level of financial security to the cricketers. The value of the IPL has risen from US$2 billion in 2009 to US$4.7 billion in 2021 with a peak value of US$6.18 billion in 2018 (*The Times of India*, 2021). The ever-increasing amount of money involved in the IPL is also evident from the fact that an initial bid of just ₹3,000 crore was made to acquire all 8 IPL teams in 2008, the cost shot 17-folds in over one decade where just 2 teams were acquired at a whooping cost of ₹12,715 in 2021 (Mishra, 2021). Over the last 3 years, viewership of each age group has registered significant growth: kids (<15 years) – 52 per cent, that of 15-21 years – 50 per cent, while 22-30 years – 20 per cent (*Broadcast Audience Research Council of India*, 2018). Another staggering statistic that reinforces the relevance and impact of IPL in the country is that "Almost half of the Indian TV viewers watched IPL 2020" (Laghate, 2020). The tenth season (in 2017) of IPL generated a TV unique viewership of about 400 million, an increase from 361 million in 2016 (Keelery, 2021).

Given the increasing popularity and viewership of IPL, about 247 brands were advertised during the 2018 season of the IPL compared to 138 in 2016 (Basuroy, 2021). Large firms such as Coca-Cola, MRF, Castrol, PhonePe, and Maruti Suzuki have invested heavily in the past to reap benefits from the growing popularity of the IPL. Sen and Kayal (2022) report that the stock returns of the sponsors of IPL matches are linked to the performance of the teams. Thus, from a business perspective, firms are willing to sponsor a team that is more likely to win the tournament. While every team may want to win the tournament for obvious reasons, winning also helps teams gain popularity and grab good sponsorship deals. To assist with this decision-making, businesses often look forward to the field of sports analytics for providing valuable insights (Kapadia *et al.,* 2020).

The existing literature on predicting the outcomes of various types of cricket matches has studied various factors such as home ground, experience of the players, record with the opposing team, current form of the players and the individual team, day/night, innings (first or second) and physical fitness of teams (Pathak and Wadhwa, 2016). There is also an effect of the run-scoring

pattern of the team and batting. However, the role of 'coin toss' (also referred to as toss win) and 'home ground' on the outcomes of the IPL match has not been explored. The 'coin toss' was incorporated into the game since its inception in 1877, to be as fair and quick as possible with the decision of which team gets to bat/bowl first. While the coin toss itself is a random event, winning and choosing whether your side bats or bowls are often a game-defining advantage on flat or deteriorating pitches (Sood and Willis, 2021; Sawon, 2021; Cricket Prediction Blog, 2020; Kanungo and Bomatpalli, 2019; and Collins, 2014). Similarly, playing on Home Ground can be an advantage for the hosting team because of two major reasons (1) familiarity with the conditions, as IPL teams play 50 per cent of their games on their home grounds, and (2) having a say in the type of pitch that is prepared so that it could be used to suit their strengths and/or exploit the weakness of the opposition (Krishna, 2021; Oke, 2018; and Karhadkar, 2013).

Therefore, this study considers the two variables- toss win and home ground in conjunction with overall batting, bowling, and fielding performance to predict the likelihood of a win or a loss of an IPL cricket match. The research questions of this study can be described as follows:

RQ1: Is there a statistically significant association between toss win and match outcome or between home ground status and match outcome?

RQ2: How well can the match outcome be predicted using toss win status and home ground status?

To answer these research questions, the authors collect data on match outcome, home ground, and toss win from the official website of the Board of Control for Cricket in India. The Cricket IPL matches commenced in the year 2008. The authors decided to take the first fifteen years of the data on IPL games. Hence, the data was gathered for the period 2008 to 2022. Firstly, a chi-square test is performed to assess the influence of home ground and toss win on the IPL cricket match outcome. Next, we see the impact of home ground and toss win along with the team's overall batting performance, overall bowling performance, and overall fielding performance on the match outcome. The values of overall batting performance, overall bowling performance, and overall fielding performance are calculated as per the methodology outlined in Section 3, and the analysis is performed using Logistic regression analysis. The study results provide valuable implications to the stakeholders to aptly manage the game and can also enable the betters to make informed bets (in countries where it is legal). In addition to this, some of the teams have shares of their own on the stock market. Investment decisions can also be taken more thoughtfully if we have a predictive model for the outcome of any match.

The remainder of the study is structured as follows. Relevant literature is discussed in Section 2. Section 3 details the methodology used to answer the

proposed research questions. Section 4 reports the study findings. Section 5 discusses the study's implications, and the study is concluded in Section 6.

2. Literature Review

Over the past few years, the trend for applying data science and analytics to sports has increased substantially (Bonidia *et al.*, 2018; McCullagh, 2010; Lin, 2013). Not only cricket, but also games such as chess (Brown *et al.*, 2017), baseball (Li *et al.*, 2022), and volleyball (Wenninger *et al.*, 2019) have witnessed the influence of data analytics. Jordan Robertson (2014) and Li *et al.* (2022) discussed the science of predicting baseball match outcomes with objective evidence and developed predictive models aimed at this. Albert (2010) discussed how analytics could shape the future of baseball as a sport. Manoj *et al.* (2019) used AHP to predict the outcome of baseball matches. Wenninger *et al.* (2019) performed a very interesting work of applying market basket analysis to detect tactical patterns for predicting the outcome of volleyball championships. Schumaker *et al.* (2010) wrote a textbook on sports data mining, which presented data mining techniques in several sports such as cricket, football, basketball, soccer, baseball, etc.

In the context of cricket, Uddin *et al.* (2019) proposed a framework for conducting a cricket match and for evaluating the players. Precisely, the framework aims to assist the selection board in selecting the right players for the international match and predicting the man of the match. Pathak and Wadhwa (2016) used classification techniques to predict the outcome of one-day international (ODI) matches. They show that Naïve Bayesian Model is helpful in predicting the outcomes of ODI matches. Sankaranarayanan *et al.* (2014) applied data mining techniques to predict the outcomes in ODI matches. Bhattacherjee *et al.* (2015), and Das *et al.* (2016) predicted the outcomes of the ICC World Cup. Ul Mustafa *et al.* (2017) predicted the match outcomes using social media opinions. Precisely, they show that fans' score prediction and sentiments, and a total number of tweets are significant predictors of cricket match results. Asif & McHale (2016) proposed a dynamic logistics regression model to predict the outcomes of ODI matches. Akhtar and Scarf (2012) forecasted the outcomes of test cricket matches based on the pre-match effects and in-play effects. They show that wicket resources used, pre-match team, and ground effect are the most important predictors of a test match win. The focus of these studies is on predicting the outcomes of either ODI or test match, whereas the focus of this study is on IPL which follows a T20 format. T20 is a short format and hence the results of ODI or test cricket may not be applicable.

Recently, scholars have also investigated the impact of various factors on T20 and IPL match outcomes. For example, Dubey *et al.* (2021) report that the Naïve Bayesian Classification Technique can predict the T20 cricket match outcome with 71 per cent accuracy, closely followed by KNN and logistic

regression with 66 and 64 per cent accuracy. Sacheti *et al.* (2014) find partial support for the impact of choice to bat first and winning the toss on the likelihood of T20 International (T20I) match winning. Scholes & Shafizadeh (2014) showed that fielding indicators are significant predictors of T20I matches. Kanungo and Tulasi (2019) did an exploratroy analysis of batting performance, bowling performance, and cricket score to help the IPL franchise in choosing players for the match. Agrawal *et al.* (2018) attempted to predict the IPL match outcomes using machine learning. Their models predicted the IPL match outcomes with more than 95 per cent accuracy. Kapadia *et al.* (2020) show that the random forest algorithm better predicts the IPL match outcome compared to other machine learning algorithms such as Naïve Bayes, K-nearest Neighbour, and Model Tree.

While the literature on T20 and IPL cricket matches is vast and constantly increasing, we find that the studies have majorly focused on comparing various machine learning algorithms to find the best algorithm. However, the extant body of literature on identifying the relevant factors that contribute to the match outcomes is still at a nascent stage. To this end, we explore how various factors like toss win, home ground, and team performance (batting, bowling, and fielding) affect the IPL match outcomes. As highlighted earlier, the IPL teams are composed of various national and international cricketers. While a particular IPL team may belong to a state (or city), all the players are not from the same state (or city). Hence, it is important to assess if playing at a home ground is advantageous or not. Toss winning is generally considered important by cricket professionals because the winner gets to decide between bowling and batting based on the pitch conditions. While the importance of toss win is proven in other formats of cricket (ODI and Test), its impact on the outcomes of the T20 format is still unexplored. Additionally, we also consider the team performance parameters namely, batting, bowling, and fielding performance which are vital to any game/sport.

3. Research Methodology

The research methodology used in this study consists of five broad steps as shown in Figure 3.1. First, the raw data was collected from the website of BCCI. It contained the details of each IPL match from 2008 till 2022, regarding the playing teams, the toss-winning team, the home ground team, the winning team, and the variables related to runs and wickets that could help us compute the bowling performance, batting performance, and fielding performance. The Cricket IPL matches commenced in the year 2008. The authors decided to take the first fifteen years of the data on IPL games. Hence, the data was gathered for the period 2008 to 2022.

In the second step, a cross-tabulation was done between the match outcome on one side and the home-ground advantage and toss advantage on the other

FIGURE 3.1

THE FIVE-STEP RESEARCH METHODOLOGY USED IN THE STUDY
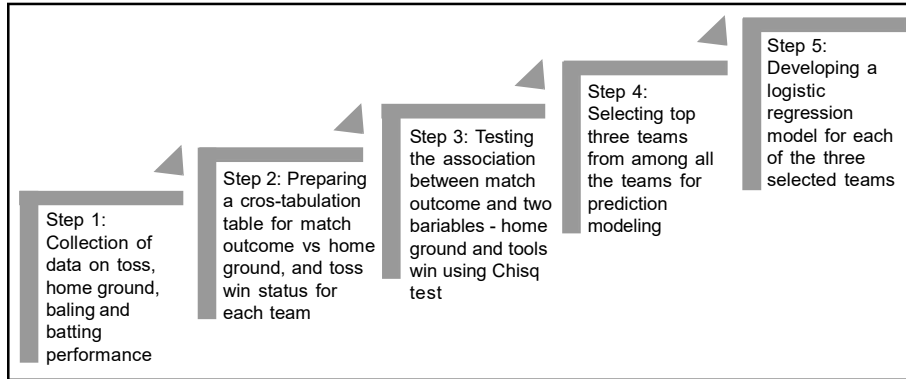
side. Such cross-tabulation was necessary to get the data in a format where the target variable and features can be read as observations of the study arranged in rows of the table.

In the third step, the Chi-square test was done to check the association between the homeground status and match outcome, and the association between the tosswin advantage on the match outcome. This test helped the authors identify these associations on a stand-alone basis while ignoring the other variables.

In the fourth step, an exhaustive analysis was carried out to select the top three players in the IPL. Three parameters were decided to select the top teams- per cent Wins in Matches Played, Number of Seasons Played, and Number of Matches Played Per Season - to indicate how long the team remained in the competition. Only 3 teams scored more than the Median Value for all three parameters: Mumbai Indians, Chennai Super Kings, and Royal Challenger Bangalore as shown in Table 3.1.

TABLE 3.1

SELECTION OF THREE TEAMS FOR DEVELOPMENT OF LOGISTIC REGRESSION MODEL

Team Name	Team 1	Team 2	Total	Wins	% Wins	No of Seasons	Matches per Season	% Win Crit	No of Season Crit	Matches per Season Crit	Selection
Chennai Super Kings	77	54	131	79	60.31%	8	16.38	1	1	1	1
Deccan Chargers	43	32	75	29	38.67%		15.00	0			0
Delhi Daredevils	63	84	147	62	42.18%	10	14.70	0			0
Gujarat Lions	14	16	30	13	43.33%	2	15.00	0		0	0
Kings XI Punjab	77	71	148	70	47.30%	10	14.80	0			0
Kochi Tuskers Kerala	7	7	14	6	42.85%	1	14.00	0		0	0
Kolkata Knight Riders	68	80	148	77	52.03%	10	14.80	1			0
Mumbai Indians	82	75	157	91	58.60%	10	15.70	1	1	1	1
Pune Warriors	20	26	46	12	26.09%	3	15.33	0			0
Rajasthan Royals	54	64	118	63	53.39%	8	14.75	1	1	0	0
Rising Pune Supergiant	8	8	16	10	62.50%	1	16.00	1			0
Rising Pune Supergiants	7	7	14	5	35.71%	1	14.00	0		0	0
Royal Challengers Bangalore	70	82	152	73	48.03%	10	15.20	1	1	1	1
Sunrisers Hyderabad	45	30	75	42	55.26%	5	15.20	1		0	0
			Median Values		47.7%	6.5	15.00	IF % WINS > MEDIAN, VALUE=1, ELSE 0	IF NO OF SEASONS > MEDIAN, VALUE=1, ELSE 0	IF MATCHES PER SEASON > MEDIAN, VALUE=1, ELSE 0	Product of all Criteria. If 1, Team is selected for analysis

In the fifth step, home ground, toss win, overall batting performance, overall bowling performance, and overall fielding performance were computed for the three teams selected in step 3. Overall Batting Performance was calculated by adding ScRD + PD, where ScRD is the Score Rate (Runs Scored per Ball) Difference in Every Match for Selected Team & Opponent Team; and PD is the Partnerships (Runs Scored per Wicket) Difference in Every Match for Selected Team & Opponent Team. Then, the Normalized Overall Batting Performance was calculated using the mean and standard deviation of Overall Batting Performance. Overall Bowling Performance was calculated by adding SRD + ERD, where SRD is the Strike Rate (Balls per Wicket) Difference in Every Match for Selected Team & Opponent Team; and ERD is the Extra Runs Difference (Runs Scored per Wicket) in Every Match for Selected Team & Opponent Team. Then, Normalized Overall Bowling Performance was calculated using the mean and standard deviation of Overall Bowling Performance. Then, the difference in counts for the Selected Team and the Opponent Team was calculated as Overall Fielding Performance. The instances where the fielder was involved in a dismissal of a batsman were calculated for the Selected Team & Opponent Team. After this, the logical Regression analysis was performed on a subset of teams to predict the likelihood of a particular team winning based on some chief variables: Home Ground, Toss Win and Overall Batting Performance, Overall Bowling Performance and Overall Fielding Performance. When checked for a correlation between attributes, a high degree of correlation (70% to 98%) was observed between Overall Batting Performance, Overall Bowling Performance & Overall Fielding Performance. Therefore, Overall Bowling performance and Overall Fielding Performance were dropped from the final regression model. Thus, only 3 Attributes for a selected team were considered for Logistic Regression viz- Home Ground, Toss Win, Overall Batting Performance. For each model, Cox & Snell R^2 and Nagelkerke R^2 were calculated. To check the fitness of the model Confusion Matrix was prepared for all three teams to calculate Accuracy, Precision & Recall.

It is worthwhile to note that while examining the home ground advantage, the authors excluded those instances in the past where different circumstances had led to the IPL being conducted outside the country. Since neither of the playing teams in these matches had home-ground advantage, such cases were excluded.

4. Findings of the Research

This section illustrates the findings of the research. Table 4.1 shows the findings from the chi-square test done to find the association of toss win and home ground with the team winning.

TABLE 4.1

CHISQUARE STATISTIC FOR THE TOSS-WIN ASSOCIATION AND HOME GROUND ASSOCIATION

Team	Toss-win	Home ground
CSK	0.715	6.941
RCB	0.378	0.631
MI	0.356	0.254
KKR	2.398	3.509
RR	0.025	11.648
DD	0.204	1.646
DC	0.062	0.311
SRH	0.736	3.526
DCh	0.186	1.026
KXIP	1.121	2.572
PWI	1.739	0.069
RPS	0.909	0.032
GL	5.992	0.909
KTK	0.389	0.000

Since the critical value of chi-square at a 90 per cent confidence level for 1 degree of freedom is 2.706, we can infer that the toss winning has a statistically significant association with winning the game only for Gujarat Lions at a 90 per cent confidence level. This is because the computed chi-square value is greater than the critical value of 2.706.

Similarly, it can be observed that at a 90 per cent confidence level, the foreground has a statistically significant association with winning for four teams, namely CSK, KKR, RR, and SRH. Table 4.2 shows the coefficients of the predictor variables in the logistic regression model that models the chances of winning of the teams.

TABLE 4.2

COEFFICIENTS OF THE PREDICTOR VARIABLES IN THE LOGISTIC REGRESSION MODEL FOR PREDICTING THE WINNING OF THE TEAMS

Team	Intercept	Homeground$_{overall}$	Tosswin$_{overall}$ (%)	Batting Per$_{overall}$ (%)
MI	0.5457	0.7349	89	89
CSK	0.4199	0.5717	82	83
RCB	0.4867	0.6611	85	88

TABLE 4.3

PSEUDO-R-SQAURE, ACCURACY, PRECISION AND RECALL VALUES FOR THREE TEAMS

Team	Cox & Snell Rsq	Nagelkerke Rsq	Accuracy (%)	Precision (%)	Recall (%)
MI	0.5457	0.7349	89	89	91
CSK	0.4199	0.5717	82	83	89
RCB	0.4867	0.6611	85	88	81

The following regression equations were noted for three teams from the overall aggregate observations given in Table 4.2.

In(Odds of MI Winning)

$$= 0.19 + 0.73 * Homeground_{overall} + 0.12 * Tosswin_{overall} + 7.13 * Batting\ Performance_{overall}$$

In(Odds of CSK Winning)

$$= 0.44 + 0.87 * Homeground_{overall} + 0.82 * Tosswin_{overall} + 3.99 * Batting\ Performance_{overall}$$

In(Odds of RCB Winning)

$$= -0.10 + 0.33 * Homeground_{overall} + 0.29 * Tosswin_{overall} + 5.63 * Batting\ Performance_{overall}$$

Since the coefficients of the three variables on the chances of winning the game are positive in the case of these three teams, it can be observed that there is a positive effect of these variables on the winning outcome when the average values of these variables are considered.

Table 4.3 shows the metrics of the logistic regression model such as pseudo-r-squared measures, accuracy, precision and recall. Since the Nagelkerke Rsq measures are more than 0.5 and the accuracy, precision and recall metrics are also more than 80 per cent, it can be inferred that the logistic regression models developed have a good predictive ability.

5. Study Implications

Any data mining exercise on sports data can benefit multiple stakeholders. The important ones among them are the professionals working with sports organizations, the sports coaches, the sports administrators, sports associations and government departments, the players themselves, the sports fans and the academicians, who are teachers or researchers in the area of sports sciences or data mining. This Section discusses what implications this research carries for academics, teams, policy-makers and cricket followers.

5.1 Implications for the Teams

This study can be quite helpful for the IPL Teams in curating strategies and plans for the team, developing a roadmap for the tournament, setting expectations and goals for the team for individual matches, and preparing for individual matchups and scenarios. It also promotes marketing and advertising campaigns for teams in their home grounds as they witness a disproportionate representation in terms of their fans in their home ground as compared to the away team.

It was found by the study that the Mumbai Indians, as a team, take benefit of playing matches on its home ground. The batting performance of a team has a huge influence on its win. A batting collapse in the case of this team was mostly found to result in a match loss. The model results also indicate that probably their bowling performance needs improvement. However, the study also shows that this team has the potential to win matches even after losing the toss, which indicates that the team has a flexible game strategy and can adapt to a situation better than other teams. The positive value of the intercept in the logistic regression equation indicates that the team can win matches even if all attributes are not in their favour.

The study finds that Chennai Super Kings also takes maximum benefit of playing matches on its home ground. The batting performance of this team was found to have a relatively lower influence on its win. The study also found that its bowling was best among the three teams. It was also found that losing a toss has the maximum adverse effect on match results as compared to the other two teams indicating that the team needs to be more adaptive to be able to change its game plans as per toss result. The positive value of the intercept in the logistic regression equation indicates that the team can win matches even if all attributes are not in its favour.

It was also found from the study that the Royal Challenger Bangalore takes the least benefit of playing matches on their home ground. The batting performance of this team has a moderate influence on their win. It also indicates that probably its bowling department is good among the three teams. Losing a toss was found to have a moderate adverse effect on the match result. The negative value of the intercept in the logistic regression equation indicates that the team is most likely to lose matches if all attributes are not in their favour. This is a major sign of a team that probably cracks under pressure and cannot win important matches when everything seems to go in favour of the opponent team. This may be an indication that this team needs a management and leadership change.

5.2 Implications of this Study for the Academics

The professors or researchers in the area of sports can appreciate from this study that both the factors- 'Toss Advantage' and 'Home Ground Advantage' have a significant impact in predicting the winner of any given IPL match.

There is also a very important role being played by the other three factors of batting performance, bowling performance and fielding performance. There are some teams that are able to convert home ground and toss into a win, depending on how well they make a decision in case of a toss win. Similarly, some teams are more comfortable with home ground as compared to other teams. Also, home ground and toss win are two of the factors that influence the match outcome. The other important factors are the capability of the team in terms of batting, bowling and fielding. The professors or researchers working in the field of data science or data analytics can also take a cue from this study and apply similar research methodologies and techniques to generate insights in any domain.

5.3 Implications of the Study for the Policy-makers

First, this study can benefit the BCCI (Board of Control for Cricket in India) in scheduling the IPL matches particularly the final leg of the tournament in such a way that there is no home ground advantage for any participating team. One way to do it could be to organize these matches on neutral venues as having a way to predict the winner hampers the quality of such important matches.

Second, the outcome of this study suggests that the ICC (International Cricket Council) should think of devising some better solutions in terms of fairness and skill-dependency, that could replace 'toss'. This is not only applicable to the IPL but also to Cricket in general since having fewer 'uncontrollable' in the sport directly leads to a more exciting contest between the participants.

Third, sports organizations, generally, do not understand the importance of data mining in sports. They need to realize that the use of data can help them perform better. And some of them who have started using data mining for sports, need to share their learnings with their peers so that other organizations can also start using analytics in sports. Also, if there can be digital repositories of sports data, that would help the organizations to access any pertinent piece of data.

5.4 Implications for the Sports Lovers

This study incentivizes the fans to go out to the stadiums to support their teams in their respective home ground as there is a higher chance for them to win. However, unlike toss advantage, home-ground advantage may not be seen as a form of prejudice within the game but it is seen as a factor that makes the game much more interesting and exhilarating. It serves as an integral part of team strategies both during the auctions (teams choose personnel whose styles are more suited to their home pitches) and during the tournament as well (teams make regular changes in their team compositions to ensure that they are best suited for specific playing conditions).

6. Conclusions, Limitations and Future Extensions

Over the past few years, IPL has seen tremendous growth both in terms of popularity as well as relevance in the Cricket world. It is not only an extremely unique opportunity for players from around the world to prove their mettle and showcase their talent but also an equally lucrative construct for the fans to see their favourite players from around the world together, in a highly competitive and fast-paced tournament. When the stakes are this high, the prediction of the matches is also something, which not only helps the decision-makers of the game to make better choices and policies but also provides incentives to evaluate and inspect various aspects of the game to remove any biases that may have crept into the game to ensure a level playing field for the teams.

The model developed in this research study attempts to predict the outcome of a match for each of the teams based on venue, toss result and batting performance. As a result, the teams can change their strategies as the game progresses to maximize the probability of winning a game. The high values of Accuracy, Precision and Recall for the prediction model developed indicate that the model is robust.

Despite its valuable contributions the present study has certain limitations. In the case of a complex sport such as Indian Cricket and a colossal tournament like the IPL, there may be many other parameters that can introduce bias to the game and can have a direct impact on deciding the winner of the match. Some of such variables can be the amount of dew on the field, the proportion of left-handers in the team, the nationality of the team captain, weather, the competency level of the coach, etc. There can be many intangible factors also such as umpiring errors, injury to the key players, collective psychology of the team players, interpersonal relations among the top players, team politics, etc. This model does not consider these factors. Capturing all these and other such factors in a single model may be quite perplexing, but can be explored as possible extensions of this research in the future.

Conflict of Interest

The author declares that they have no conflict of interest.

REFERENCES

1. Agrawal, S., Singh, S.P., and Sharma, J.K. (2018), Predicting Results of Indian Premier League T-20 Matches Using Machine Learning, Proceedings - 2018 8th International Conference on Communication Systems and Network Technologies, CSNT 2018, 8820235, pp. 67-71.

2. Akhtar, S., and Scarf, P. (2012), Forecasting Test Cricket Match Outcomes in Play, *International Journal of Forecasting, 28*(3), pp. 632-643.

3. Albert, J. (2010), Sabermetrics: The Past, the Present, and the Future, *Mathematics and Sports*, pp. 3-14.

4. Aneem-Al-Ahsan, R., Saddam Hossain Mukta, Md., and Najmul Islam, A.K.M. (2020), Predicting Bowling Performance in Cricket from Publicly Available Data, *ACM International Conference Proceeding Series.*

5. Asif, M., and McHale, I.G. (2016), In-play Forecasting of Win Probability in One-Day International Cricket: A Dynamic Logistic Regression Model, *International Journal of Forecasting*, *32*(1), pp. 34-43.

6. Basuroy, T. (2021), Number of Brands Advertised in the IPL 2016-2018. https://www.statista.com/statistics/947651/india-number-of-brands-advertised-in-ipl/#:~:text=About% 20247%20brands%20were%20 advertised,from%20about%20138%20in%202016. Last accessed on 14 August 2022.

7. Bhattacherjee, S., Sahoo, J., and Goswami, A. (2015), Association Rule Mining Approach in Strategy Planning for Team India in ICC World Cup 2015, Proceedings - 2015 2nd IEEE International Conference on Advances in Computing and Communication Engineering, *ICACCE 2015*, 7306757, pp. 616-621.

8. Bonidia, R.P., Brancher, J.D., and Busto, R.M. (2018), Data Mining in Sports: A Systematic Review, IEEE Latin America Transactions.

9. Bonidia, R.P., Rodrigues, L.A.L., Avila-Santos, A.P., Sanches, D.S., and Brancher, J.D. (2018), Computational Intelligence in Sports: A Systematic Literature Review, *Advances in Human-Computer Interaction*, 2018, 3426178.

10. Borooah, V.K., and Mangan, J. (2012), Mistaking Style for Substance: Investor Exuberance in the 2008 Indian Premier League Auction, *Journal of Sports Economics*, 13(3), pp. 266-287.

11. Bray, S.R. (1999), The Home Advantage from an Individual Team Perspective, *Journal of Applied Sport Psychology*, 11(1), pp. 116-125. DOI: 10.1080/ 10413209908402954

12. Broadcast Audience Research Council, India (2018), Think IPL. https:// www.barcindia.co.in/newsletter/ipl-over-the-years.pdf. Last accessed on 14 August 2022.

13. Brown, J.A., Cuzzocrea, A., Kresta, M., Kristjanson, K.D.L., Leung, C.K., and Tebinka, T.W. (2017), A Machine Learning Tool for Supporting Advanced Knowledge Discovery from Chess Game Data, *Proceedings - 16th IEEE International Conference on Machine Learning and Applications, ICMLA 2017*, December, pp. 649-654.

14. Chaudhary, K., Gupta, M., and Kaur, P. (2019), Analyzing IPL Dataset with MongoDB, *Proceedings of the 9th International Conference on Cloud Computing, Data Science and Engineering, Confluence 2019*, 8776979, pp. 212-216.

15. Collins, T. (2014), Numbers Suggest the Toss Has Never Been More Important in Test Cricket. https://bleacherreport.com/articles/1969898-numbers-suggest-the-toss-has-never-been-more-important-in-test-cricket. Retrieved on 10 April'2022.

16. Cricket Prediction Blog (2020), Toss Prediction - Can we Predict the Toss in IPL? https://cricketprediction.com/betting/toss-prediction-in-ipl/ Retrieved on 10 April 2022.

17. Das, A., Parida, A.R., and Srivastava, P.R. (2016), ICC Cricket World Cup Prediction Model, in Information Systems Design and Intelligent Applications, Springer, New Delhi, pp. 529-539.

18. Devan, P. (2012), Cricket and the Global Indian Identity, *Sport in Society, 15*(10), pp. 1413-1425. DOI: 10.1080/17430437.2012.744210

19. Dubey, P.K., Suri, H., and Gupta, S. (2021), Naïve Bayes Algorithm Based Match Winner Prediction Model for T20 Cricket, *Advances in Intelligent Systems and Computing*, 1172, pp. 435-446.

20. Fujak, H., Frawley, S., Lock, D., and Adair, D. (2019), Consumer Behaviour toward a New League and Teams: Television Audiences as a Measure of Market Acceptance, *European Sport Management Quarterly,* 22:2, pp. 161-181. DOI: 10.1080/16184742.2020.1770310

21. Jordan Robertson, T. (2014), Building a Predictive Model for Baseball Games, December. Retrieved from http://cornerstone.lib.mnsu.edu/cgi/viewcontent.cgiarticle=1381&context=etds

22. Kanungo, V. and Bomatpalli, T. (2019), Data Visualization and Toss Related Analysis of IPL Teams and Batsmen Performances, *International Journal of Electronic and Computer Engineering, 9*(5), pp. 4423-4432.

23. Kanungo, V., and Tulasi, B. (2019), Data Visualization and Toss Related Analysis of IPL Teams and Batsmen Performances, *International Journal of Electrical and Computer Engineering*, Accepted for publication.

24. Kapadia, K., Abdel-Jaber, H., Thabtah, F., and Hadi, W. (2020), Sport Analytics for Cricket Game Results Using Machine Learning: An Experimental Study, *Applied Computing and Informatics*, (ahead-of-print).

25. Karhadkar, A. (2013), Why Home Advantage Makes a Difference? https://www.espncricinfo.com/story/ipl-2013-why-home-advantage-makes-a-difference-in-the-ipl-633932. Retrieved on 10 April 2022.

26. Keelery, S. (2021), Unique TV Viewers of IPL Across India 2017, by Season. https://www.statista.com/statistics/825753/india-unique-tv-viewers-of-

ipl-by season/#:~:text=The%20tenth%20season%20(in% 202017,IPL%20 season%20in%20the%20country. Last accessed on 14 August 2022.

27. Kim Kyunghak (2010), Cricket and Indian Nationalism in Post-Colonial Age, *Journal of South Asian Studies* 15:3, pp. 85-112.

28. Kitchin, P. (2014), Global Sport Business: Community Impacts of Commercial Sport, Managing Liesure, 19(6), pp. 461-463. DOI: 10.1080/16184742. 2014.964576

29. Krishna, S. (2021), How Important is Home Advantage in IPL? https:// www.linkedin.com/pulse/how-important-home-advantage-ipl-siddharth-krishna/?trk=public_profile_article_view Retrieved on 10 April 2022.

30. Laghate, G. (2020), Almost Half of Indian TV Viewers Watched IPL 2020, *The Economic Times*. https://m.economictimes.com/industry/media/ entertainment/almost-half-of-indian-tv-viewers-watched-ipl-2020/ articleshow/79323814.cmsLast accessed on 14 August 2022.

31. Li, S.F., Huang, M.L., and Li, Y.Z. (2022), Exploring and Selecting Features to Predict the Next Outcomes of MLB Games, *Eltropy*.

32. Lin, J. (2013), Application of Computer Technology in Sport Data Statistics, *ACM International Conference Proceeding Series*, pp. 248-250.

33. Mani, E. (2009), A Strong Sport Growing Stronger: A Perspective on the Growth, Development and Future of International Cricket, *Sport in Society*, 12(4-5), pp. 681-693. DOI: 10.1080/17430430802702970

34. Manoj, M., Prashant, R., and Parikh, V. (2019), American League Baseball Championship 2017 Prediction using AHP, Proceedings of the 2018 International Conference on Communication, *Computing and Internet of Things*, IC3IoT 2018.

35. Matondkar, A., and Bagchi, A. (2020), IPL as a Catalyst for Sports Tourism in India a Learning Curve for other Indian Leagues, *Annals of Tropical Medicine and Public Health*, 23(17), SP231705.

36. Mauws, M.K., Mason, D.S. and Foster, W.M. (2003), Thinking Strategically about Professional Sports, *European Sport Management Quarterly, 3*(3). doi: 10.1080/16184740308721948

37. McCullagh, J. (2010), Data Mining in Sport: A Neural Network Approach, *International Journal of Sports Science and Engineering*, 4 (3), pp. 131-138.

38. Mishra, M. (2021), 17-Fold Jump in Franchise Value, Thrice the Revenue & Counting: Why IPL is a Money Pit, *India Today Blog*. https://www. indiatoday.in/diu/story/17-fold-jump-in-franchise-value-thrice-the-revenue-counting-why-ipl-is-a-money-pit-1870164-2021-10-27. Last accessed on 14 August 2022.

39. Mitra, S. (2010), The IPL: India's Foray into World Sports Business, *Sport in Society*, 13(9), pp. 1314-1333.

40. Modani, N., Kilaru, M., Kaur, A., Sinha, R., and Khetan, H. (2020), Predicting Outcomes in Limited-overs Cricket Matches, *ACM International Conference Proceeding Series*, pp. 65-72.

41. Morley, B. and Thomas, D. (2005), An Investigation of Home Advantage and other Factors Affecting Outcomes in English One-day Cricket Matches, *Journal of Sports Sciences, 23*(3), pp. 261-268. DOI: 10.1080/02640410410001730133

42. Moujahid, A. (2014), Baseball Analytics: An Introduction to Sabermetrics Using Python, 13 July. Retrieved from http://adilmoujahid.com/posts/2014/07/baseball-analytics

43. Nair, N. (2011), Cricket Obsession in India: Through the Lens of Identity Theory, *Sport in Society*, 14(5), pp. 569-580. DOI: 10.1080/17430437.2011.574351

44. Oke, A. (2018), Home Advantage in the IPL. http://www.holdingwilley.com/article/musings/home-advantage-in-the-ipl Retrieved on 10 April 2022.

45. Pathak, N., and Wadhwa, H. (2016), Applications of Modern Classification Techniques to Predict the Outcome of ODI Cricket, *Procedia Computer Science*, 87, pp. 55-60.

46. Raj, K.A.A.D., and Padma, P. (2013), Application of Association Rule Mining: A Case Study on Team India, *2013 International Conference on Computer Communication and Informatics, ICCCI2013*, 6466294.

47. Rana, A.S., and Bagchi, A. (2020), Impact of Indian Premier League on Test Cricket in India, *Annals of Tropical Medicine and Public Health*, 23(17), p. SP231709.

48. Ray, S. (2008), 'The Wood Magic': Cricket in India – A Postcolonial Benediction, *The International Journal of the History of Sport*, 25(12), pp. 1637-1653. DOI: 10.1080/09523360802327483

49. Sacheti, A., Gregory-Smith, I., and Paton, D. (2016), Managerial Decision Making Under Uncertainty: The Case of Twenty20 Cricket, *Journal of Sports Economics*, *17*(1), pp. 44-63.

50. Sankaranarayanan, V.V., Sattar, J., and Lakshmanan, L.V.S. (2014), Auto-play: A Data Mining Approach to ODI Cricket Simulation and Prediction, *SIAM International Conference on Data Mining 2014*, SDM 20142, pp. 1064-1072.

51. Sawon (2021), Toss Factor in IPL. https://medium.com/top-level-sports/toss-factor-in-ipl-d3f051b9c29d#:~:text=Though%20overall%20the%20

toss%20factor,won%20the%20toss%20or%20not. Retrieved on 10 April 2022.

52. Schumaker, R.P., Solieman, O.K., and Chen, H. (2010), Sports Data Mining: The Field, *Sports Data Mining*, pp. 1-13.

53. Sen, S., and Kayal, P. (2022), Indian Premier League (IPL) Match Victory and Sponsors' Stock Returns, *Journal of Public Affairs*, 22(1), p. e2352.

54. Singh, T., Singla, V., and Bhatia, P. (2016), Score and Winning Prediction in Cricket Through Data Mining, International Conference on Soft Computing Techniques and Implementations, *ICSCTI 2015*, 7489605, pp. 60-66.

55. Sood, G. and Willis, D. (2021), Why Replacing the Toss with an Auction is the Fair Thing To Do? https://www.espncricinfo.com/story/why-replacing-the-toss-with-an-auction-is-the-fair-thing-to-do-1277744 Retrieved on 3 March 2022.

56. Srinivas, S., Bhat, N.N., and Revanasiddappa, M. (2021), Data Analysis of Cricket Score Prediction, *Advances in Intelligent Systems and Computing*, 1245, pp. 465-472.

57. *The Times of India* (2021), BCCI could Earn up to US$5 billion from IPL Broadcasting Rights Only. https://timesofindia.indiatimes.com/sports/cricket/ipl/top-stories/bcci-could-earn-up-to-usd-5-billion-from-ipl-broadcasting-rights-only/articleshow/87182989.cms.

58. Uddin, M.A., Hasan, M., Halder, S., Ahamed, S., and Acharjee, U.K. (2019), CRICRATE: A Cricket Match Conduction and Player Evaluation Framework, *Advances in Intelligent Systems and Computing*, 755, pp. 491-500.

59. Ul Mustafa, R., Nawaz, M.S., Lali, M.I.U., Zia, T., and Mehmood, W. (2017), Predicting the Cricket Match Outcome Using Crowd Opinions on Social Networks: A Comparative Study of Machine Learning Methods, *Malaysian Journal of Computer Science*, 30(1), pp. 63-76.

60. Wenninger, S., Link, D., and Lames, M. (2019), Data Mining in Elite Beach Volleyball-Detecting Tactical Patterns Using Market Basket Analysis, *International Journal of Computer Science in Sport.*

61. Weeraddana, N., and Premaratne, S. (2021), Unique Approach for Cricket Match Outcome Prediction Using Xgboost Algorithms, *Journal of Theoretical and Applied Information Technology*, 99(9), pp. 2162-2173.

62. Yousaf, A., Gupta, A., and Mishra, A. (2017), Sport Team Brand-Equity Index: A New Measurement, *Journal of Indian Business Research*, 9(2), pp. 169-188.