

A Study on Impact of Toss in the results of ICC Men's T20 World Cup

Mr. Hariharan G¹, Miss. Nipuna M², Mr. Thirumalai V³, Dr. Suresh Nandha Kumar⁴

^{1,2,3}II PGDM Student, Xavier Institute of Management & Entrepreneurship, Chennai, India

⁴Senior Assistant Professor, Xavier Institute of Management & Entrepreneurship, Chennai, India.

Corresponding Author: Hariharan G

Date of Submission: 10-10-2022

Date of Acceptance: 21-10-2022

ABSTRACT: Cricket over the years has been seen as a game of entertainment rather than a platform for the players to showcase their skills. With the development of T20 cricket and the biggest league in the name of IPL, the expectation of it is unimaginable. Even though IPL is considered as the greatest revolution of T20 cricket, The biggest stage of them all is the World Cup which is always expected to be a greatest contest among the best players around the world with the precious price of the World Title in their mind. But the last T20 World Cup which happened in 2021 saw this trend going down due to many one-sided games where the team that won the toss had the upper hand in all the cases. So, in this paper, we made an effort to find the Impact of tosses in winning the World T20 matches. For this accord, we have used the past World T20 match results till World T20 2021 and the T20I matches that were played in Australian grounds where the current World Cup matches have been scheduled till 7th Oct, 2022 and has been plugged into a data Analytics tool, 'RapidMiner Studio' where we have used Correlation Analysis using the Correlation matrix model to bring about the relationship between various factors and how it affects the results of the match.

KEYWORDS: Cricket, sports analytics, correlation, Toss analysis and RapidMiner Studio

I. INTRODUCTION

Cricket has evolved so much as the years went by, with the Audience's mindset on the game continuously changing in the way of preference to a shorter format with lots of fireworks happening in a short span of time. The longer format of the game i.e. The Test cricket saw the trend of audiences moving out of the game. Then came the concept of One day matches, where the results of the game are

known on the same day without waiting for a longer time, which brought in lots of spectators and as the time goes that again was forced to evolve as T20 cricket, which has different sets of rules, and the rules being designed specifically to bring in a large set of people in the way of reducing the total time spent for the game, and the shorter boundaries ensured more fireworks, which acts a great entertainer. Now IPL cricket is dominating the viewership, which is arguably one of the biggest leagues in the world. However, the biggest event of them all is the T20 World cup, where the international teams clash between each other for the precious title. The first of its kind was hosted by South Africa in 2007, where team India was crowned as champions. Then on, the tournament is organized by the International Cricket Council (ICC) for every two years till 2016. Then after a break it was again scheduled to happen in 2020 but, due to the Covid-19 pandemic, the event was then moved to UAE, which then happened in 2021, where team Australia came out as winners. But this world cup was nothing short of a disaster, as the world cup was majorly influenced by the luck factor. Trends shows that, in majority of the matches the team won the toss went on to win the match. One can ask what's the big deal in that. Cricket is not just a game of luck, it involves players skills, mindset, etc. The disability of not able to provide matches with fair contests has proved to be disastrous with the danger of audiences moving out of the game. So, we are trying to bring out a model which can be used to predict the impact of the toss decision in the upcoming t20 world cup 2022, which is scheduled to happen in Australia from 16th October to 17th of November 2022. For this model we have used the RapidMiner Studio for analysis. Also, we have used the results of the T20 matches

that have happened in the grounds in which the upcoming matches have been scheduled to play

II. REVIEW OF LITERATURE

Aman Sahu et al. (2021) states that different affecting factors namely, Individual performance, Team performance and environmental factors are needed to be considered in devising a game strategy. So, A machine learning model has been designed to bring out the outcome of the games. Google Colab- a data analytics tool was utilized for analysing two models namely, Training model and Testing model. Results proves the Random Forest Classifier algorithm as the effective method showing 98.14% accuracy for Training model and 89.47% for testing model compared to Multinomial Logistic Regression for which the accuracy rate was mere 29.62% and 27.63% for Training and Testing model respectively and as a last resort, Adaptive Boosting method was tried, which again proved to be less efficient than the Random Forest classifier algorithm.

Vidit Kanungo et al. (2019) identifies the toss impact on performance of batters in IPL cricket using data visualization. Several factors like maximum man of the matches, Maximum hundreds scored, Top 10 run getters and players with high run rate were considered to get about the results which showed Selectors of team Mumbai and Punjab made good decisions by handling the situation correctly.

Kumash Kapadia et al. (2019) have addressed the difficulties in predicting results of the Indian Premier League (IPL) games with past match data. Instrumental factors of the data have been found using Filter-based methods. Two subsets were created in which one was with respect to home team advantage and other was on Toss' decision. Selective Machine Learning (ML) techniques have been used to create a predictive model. Results implies that the Random Forest model showed better results than Probabilistic and statistical models in the grounds of accuracy and recall metrics. But No model was supported in deriving a precise predictive model for the toss decision subset.

Basil M. de Silva et al. (2015) has taken 427 ODI matches that happened during the 1990s to see if winning the toss and home team advantage gives a competitive advantage in the One-day International Matches. Results suggest that winning the toss doesn't provide an edge over the other team in winning the match however home field advantage proves to be an influential factor.

Hemanta Saikia et al. (2010) says the toss has proved to be an influential factor in Tests and One day matches and this paper tries to prove the impact of Toss on the results of the shorter format of the game i.e., T20 cricket using the statistical tool, the binomial logistic regression. The Home Team advantage and the time at which the game takes place viz day or day-night is also considered for the model. Results suggest that there is very less home team advantage but the toss proves to be an impactful factor but, when factors like 'ground*toss' and 'Play under*toss' are brought into the model, the Toss doesn't impact the result of a T20I match.

III. DATA COLLECTION

The data has been collected from the website 'EspnCricinfo'. Data contains details for a total of 222 T20 world cup matches from 2007 to 2021 where seven tournaments so far have been played. It includes 4 matches with no result, 2 matches with tied result and 1 abandoned match. The dataset file "T20 world cup toss data.csv" gives the details of all the world cup T20 matches which includes different attributes such as date, venue, Team details, Toss details and results of each match. The data considered here is populated with the matches played between 20 nations belonging to the International Cricket Council (ICC). These teams are Afghanistan, Australia, Bangladesh, England, Hong Kong, India, Ireland, Kenya, Namibia, Nepal, Netherland, New Zealand, Oman, Pakistan, Papua New Guinea, Scotland, South Africa, Sri Lanka, West Indies and Zimbabwe.

Table 1 describes the attributes used in the dataset for toss related analysis. The key attributes are Toss won and Team won for all the 222 T20 world cup matches from 2007 to 2021.

Table 1 Description of attributes used

Attributes	Description
Date	The date of commencement of the match.
Team 1	Name of one of the two teams in the match.
Team 2	Name of the opposition team played in the match against Team 1.
Toss won	The team won the toss.
Elected to	Decision made by the toss winning team to bat or field first.
Team won	The team won the match.

IV. ANALYSIS AND INTERPRETATION

After the secondary data collection of past World cup T20 matches, Dummy values have been assigned for all the qualitative data. The data is derived in the .csv format and imported to the RapidMiner Studio for further analysis.

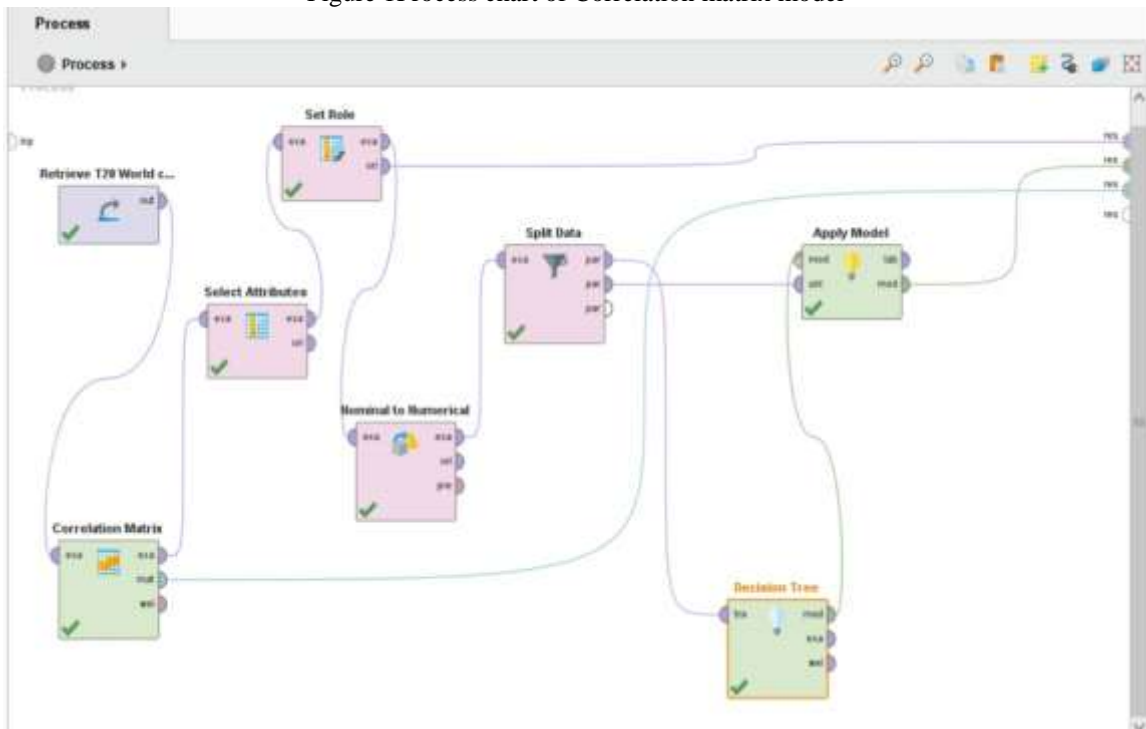
Figure 1 is the process flow of this model of our research. The process starts by importing data from the “T20 world cup toss data.csv” file by using Retrieve operator in RapidMiner Studio. Then to find the relationship between our proposed attributes the correlation matrix is made with all the attributes from the dataset. For creating the correlation matrix, Correlation Matrix operator is used in the model. All the attributes of the dataset are selected for the

analysis by using Select Attributes operator. After selecting all the attributes from the dataset, we used

Set Role operator. Next, all nominal data are converted to numerical by using the Nominal to Numerical operator. Dummy coding has been done for changing nominal values. Then the Split Data operator is used for splitting the data for both training and testing. Decision Tree operator is used for the accurate results, Finally, Apply Model operator is used for getting the result using the proposed process model.

While interpreting based on correlation operator in the correlation matrix model, the correlation values of the two attributes Toss won and Team won gave the linear positive relationship between each of them. This also says that both of them have a relationship in the same direction.

Figure 1 Process chart of Correlation matrix model



V. FINDINGS

In this paper finding out the real advantage of winning the toss in T20 world cup matches is the main objective. The correlation matrix model made using the RapidMiner Studio indicated that there is a strong correlation between winning the toss and winning the match.

Table 2 gives the correlation value between Toss won and Team won. The correlation co-efficient value 0.694 indicates a strong positive

relation between winning the toss and winning the match. Hence, it reveals that in T20 world cup tournaments, the team who wins the toss has a 69.4% chance of winning the match.

Figure 2 gives the result of the analysis using the correlation matrix model in heatmap where the red and blue colors are used to the strong positive and strong negative correlation among the attributes. The shade of each color represents the

correlation value from -1 to 1 where the shades start with blue and end with red color respectively.

Table 3 gives the ground analysis of Australian grounds where the upcoming World cup T20 tournament will take place from 16th October to 17th of November 2022. This data holds the attributes of past T20 matches played in the respective grounds. The data in the table reveal

that deciding batting first while winning the toss will be favorable for winning the match in Sydney and Brisbane Cricket grounds. Meanwhile, deciding bowling first will be favorable for winning the match in Melbourne Cricket ground. In the rest of the grounds favorable toss decisions cannot be analyzed because of the smaller number of T20 matches played in that grounds.

Table 2 Correlation matrix

Attributes	Date	Team 1	Team 2	Toss won	Elected to	Team won
Date	1	-	-	-	-	-
Team 1	-	1	0.311	0.579	0.029	0.720
Team 2	-	0.311	1	0.703	0.001	0.579
Toss won	-	0.579	0.703	1	0.005	0.694
Elected to	-	0.029	0.001	0.005	1	0.051
Team won	-	0.720	0.579	0.694	0.051	1

Figure 2 Heatmap

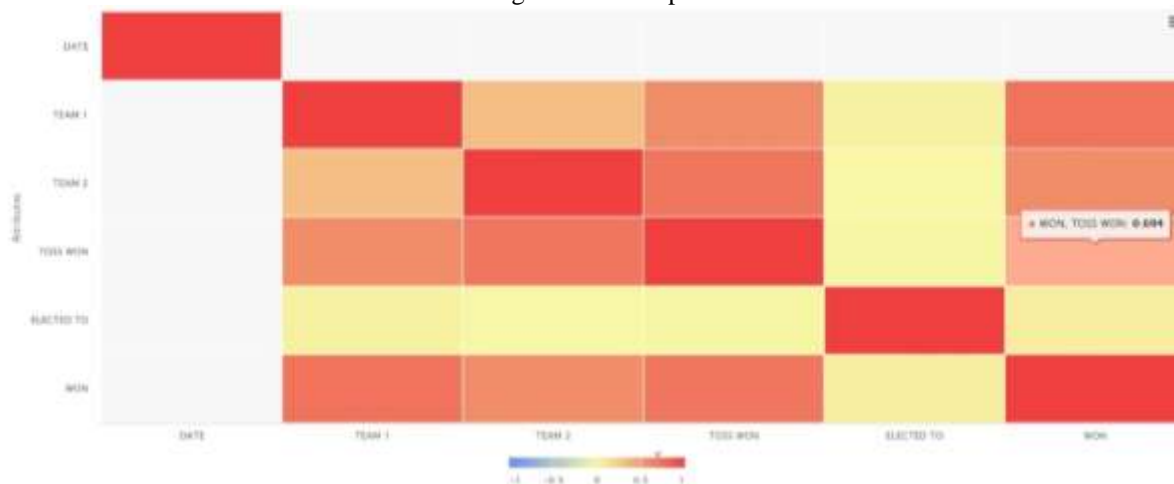


Table 3 Australian ground data

Ground	Total Matches	Matches won Batting first	Matches won Bowling First
Sydney Cricket Ground (SCG)	12	6	5
Kardinia Park in Geelong	1	0	1
Bellerive Oval or Blundstone Arena, Hobart	3	2	1

Melbourne Cricket Ground (MCG)	18	7	10
Perth Stadium or Optus Stadium	1	0	1
Adelaide Oval	5	3	2
Brisbane Cricket Ground or The Gabba	6	4	2

(Source: Cricbuzz)

VI. RECOMMENDATION AND CONCLUSION

The analysis of the results achieved by using the tool “RapidMiner Studio” based on data of all seven WT20 tournaments indicates that the impact of winning the toss generated a positive result of winning the match in T20 world cup tournaments. Even though there are other factors like pitch, weather conditions and dew in the limited over world cup T20 matches, winning toss gives greater advantage for the team to select their favourable decision whether to bowl or bat first.

Future Prediction and Analysis of cricket sports by considering past game data, players performance, natural parameters, pitch and other factors can be done by applying machine learning technology for the possibility of improving the accuracy rate of the prediction on the results of matches.

Using Machine learning might slightly improve the prediction of the result. The machine learning Technique to be more accurate the requirement of data will include the live streaming data and statistics of players. This would also be advantageous to predict the final scores of the innings by the analysis of run rate per over where run rate can make major impact in first innings and the second innings. Through this we can check the probability of winning of each team depending on the actual run rate and the required run rate on the second innings. Finally in the near future, we aspire to implement machine learning using python to make the prediction of the result more accurate and with several factors involved.

REFERENCES

- [1]. Hemanta Saikia, Dibyojyoti Bhattacharjee (2010), On the Effect of Home Team Advantage and Winning the Toss in the Outcome of T20 International Cricket Matches, Assam University Journal of Science and Technology, Vol.6, No.2, pp.88-93
- [2]. Basil M. de Silva, Tim B. Swartz (2015), Winning the Coin Toss and Home Team

Advantage in One-Day International Cricket Matches, The New Zealand Statistician,

- [3]. Kumash Kapadia, Hussein Abdel, Fadi Thabtah, Wael Hadi (2019), Sport analytics for cricket game results using machine learning: An experimental study, Applied Computing and Informatics, Vol 18 No. 3/4, pp. 256-266, DOI: 10.1016/j.aci.2019.11.006
- [4]. Vidit Kanungo, Tulasi B (2019), Data visualization and toss related analysis of IPL teams and batsmen performances, International Journal of Electrical and Computer Engineering (IJECE), Vol.9, No.5, pp.4423-4432, DOI: 10.11591/ijece.v9i5.pp4423-4432
- [5]. Aman Sahu, Devang Kaushik, A. Meena Priyadarshini (2021), Predictive Analysis of Cricket, Turkish Journal of Computer and Mathematics Education, Vol.12 No.6, pp.5111-5124