

“IPL Win – Loss and Score Predictor”

Mr. Sanket Singh
(TCET-IT, Student)
Mumbai, India.
Email id:
sanketsingh98.5@gmail.com

Mr. Vivek Vikram Singh
(TCET-IT, Student)
Mumbai, India.
Email id:
singhvivek@gmail.com

Mr. Ankit Sharma
(TCET-IT, Student)
Mumbai, India.
Email id:
ankitsharma2326@gmail.com

Mr. Vijaykumar Yele
(TCET-IT, Asst.Professor)
Mumbai, India.
Email id:
vijaypyele@gmail.com

Abstract- — The goal of this project is to use variety of algorithms on a comprehensive dataset of cricket match results to predict the team agnostic match outcomes. The target competition was The Indian Premier League (IPL) which is a twenty over format cricket league. We have extracted the results of every IPL match from the years 2008 to 2018. This data allowed us to calculate arbitrary batsman and bowler statistics before every game. The models were optimized with individual player features and then both team and player features. The models were trained based on the data of 11 IPL seasons from 2008 to 2018. The optimal model was a simple prediction method which can predict the outcome of a match based on the opponents and the players in particular team. Although it is difficult to take into account all features that influence the results of the matches, an attempt to find the most significant features is made and various classifiers are tested to solve the problem.

Overview:

Introduction:

As a sport cricket is played globally across 106 member states of the International Cricket Council (ICC), with an estimated 1.5 billion fans worldwide (ICC, 2012-2013). However, much of the global finance and interest is focused upon the 10 full ICC member nations and more specifically upon ‘the big three’ of England, Australia and India. Cricket is the India’s most popular sport. It is played, watched, and enjoyed by billions of people country wide. And in India particularly, the sport has more in common with religion than entertainment. The IPL is the most-attended cricket league in the world and ranks sixth among all sports leagues. In 2010, the IPL became the first sporting event in the world to be broadcast live on YouTube. The brand value of IPL in 2017 was US\$5.3 billion, according to Duff & Phelps. According to BCCI, the 2015 IPL season contributed

₹11.5 billion (US\$182 million) to the GDP of the Indian Economy. So, one can easily understand that Cricket is a multi-billion-dollar market. Although Cricket is a game of uncertainties there is a strong incentive for models that can predict the outcomes of matches and beat the odds provided by bookers. The aim of this study is to investigate to what degree it is possible to predict the outcome of cricket matches.

Background

The main objective of this project is to develop an IPL win-loss and score predictor that will help the Management to find the best combination of player. The goal of this project is to use variety of algorithms on a comprehensive dataset of cricket match results to predict the team agnostic match outcomes. The target competition is The Indian Premier League (IPL) which is a twenty over format cricket league. We will be extracting the results of every IPL match from the years 2008 to 2018. This data will allow us to calculate arbitrary batsman and bowler statistics before every game. The models will be optimized with individual player features and then both team and player features. The models will be trained based on the data of 11 IPL seasons from 2008 to 2018. The optimal model was a simple prediction method which can predict the outcome of a match based on the opponents and the players in particular team. Although it is difficult to take into account all features that influence the results of the matches, an attempt to find the most significant features is made and various classifiers are tested to solve the problem.

Importance of the project

Project will able to predict the probability of a particular team winning a certain T-20 match of IPL based on the input by the user. The project will help to decide how much a given team has the chance of winning at a particular venue considering the given conditions. It will help to form a better team which is most suited to face the opponent team and increase the overall chances of

winning. Using neural networks and machine learning our project aims to deliver an application to predict the success of a team in an IPL match.

Objectives and Scope of the project:

The aim of the project is to predict the probability of a particular team winning a certain T-20 match of IPL based on the input by the user. The input will take consideration of the following factors: teams playing the match, toss winning team, toss decision and the venue where the match is being played.

The input will go through several algorithms in order to predict the outcome of the match i.e. team winning the match. It will also try to predict the probable score range for the winning team using neural networks based on the attributes passed to it. The project will help to decide how much a given team has the chance of winning at a particular venue considering the given conditions. It will help to form a better team which is most suited to face the opponent team and increase the overall chances of winning.

Other scopes of this project are:

Using neural networks and machine learning our project aims to deliver an application to predict the success of a team in an IPL match.

To predict the success of the team we are taking the statistics of every innings of batsman as well as the bowler through every IPL season from 2008-2018. Also, the team statistics at a particular venue are considered.

The prediction will be based on the outcomes to all the IPL matches in the past 11 years.

Since neural networks work on experience, we will be providing more than 600 different matches to our network, each with complete ball by ball data. This will help to gain accurate results. The number of matches won with a particular team combination, the average score of the team as well as the progress of the team through the season will be the deciding parameters.

Summary:

The project is to develop a “IPL win-loss and score predictor” this will help the Management to find the best combination of player. This project will use variety of algorithms on a comprehensive dataset of cricket match results to predict the team agnostic match outcomes. The target competition is The Indian Premier League (IPL) which is a twenty over format cricket league. We will be extracting the results of every IPL match from the years 2008 to 2018. This data will allow us to calculate arbitrary batsman and bowler statistics before every game. The input will go through several algorithms in order to predict the outcome of the match i.e. Team

winning the match. It will also try to predict the probable score range for the winning team using neural networks based on the attributes passed to it.

The project will help to decide how much a given team has the chance of winning at a particular venue considering the given conditions. It will help to form a better team which is most suited to face the opponent team and increase the overall chances of winning.

Using neural networks and machine learning our project aims to deliver an application to predict the success of a team in an IPL match. To predict the success of the team we are taking the statistics of every innings of batsman as well as the bowler through every IPL season from 2008-2018. Also, the team statistics at a particular venue are considered.

The prediction will be based on the outcomes to all the IPL matches in the past 11 years. Since neural networks work on experience, we will be providing them various different matches to our network, each with complete ball by ball data. This will help to gain accurate results.

Design and Implementation:

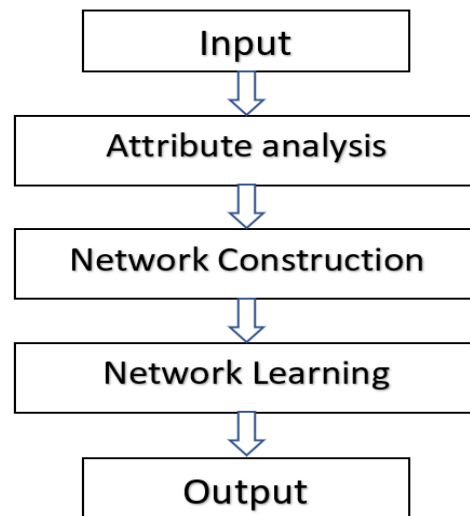


Figure No 1: Proposed Methodology

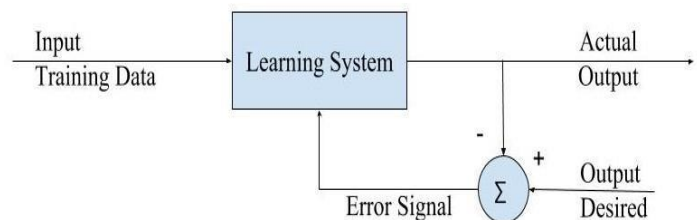


Figure No 2: Block diagram of Supervised Learning

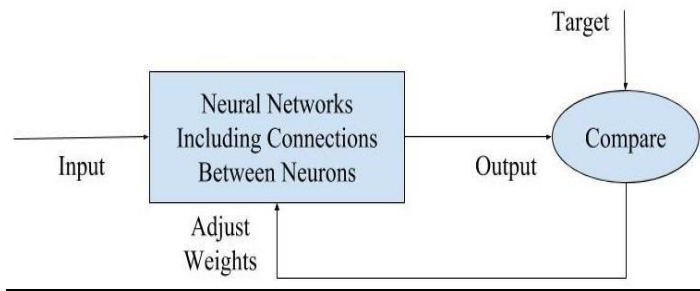


Figure No 3: Block diagram of Unsupervised Learning

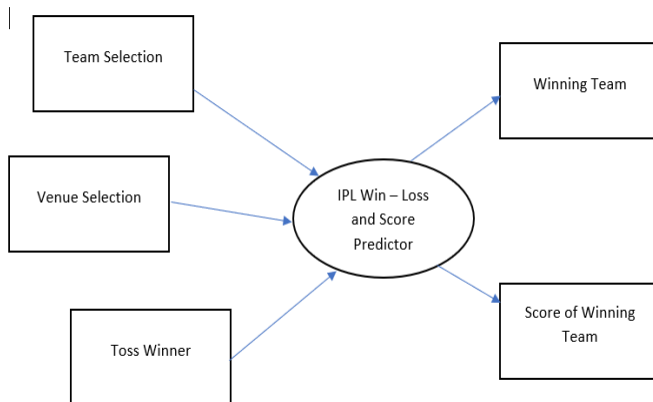


Figure No 4: DFD

Algorithms

It is a classification technique based on Bayes Theorem with an assumption of independence among predictors. In simple terms, a Naïve Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

Likelihood
Class Prior Probability

Posterior Probability
Predictor Prior Probability

$$P(c|X) = P(x_1|c) \times P(x_2|c) \times \dots \times P(x_n|c) \times P(c)$$

Figure No. 5: Naïve Bayes Algorithm

Random Forests grows many classification trees. To classify a new object from an input vector, put the input vector down each of the trees in the forest. Each tree gives a classification, and we say the tree "votes" for that class.

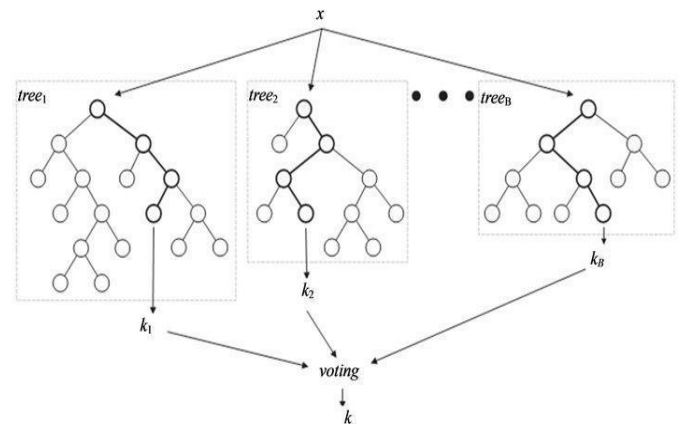


Figure No. 6: Random Forest Tree

A Support Vector Machine (SVM) is a supervised machine learning algorithm that can be employed for both classification and regression purposes. SVMs are more commonly used in classification problems and as such, this is what we will focus on in this post. SVMs are based on the idea of finding a hyperplane that best divides a dataset into two classes, as shown in the image below.

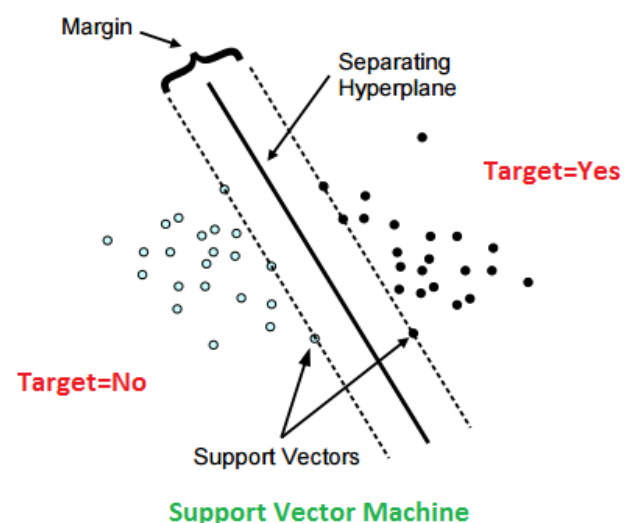


Figure No. 7: Support Vector Machine

Results and Discussion

Actual Results:

Implementation Details

The implementation plan for an IPL match win-loss prediction using neural networks is as follows:

1. A database of all the IPL matches containing various match statistics is stored in MS Excel.
2. The user will be given a form that will be generated using a suitable programming language. This form will contain various attributes that will be used as testing data.
3. All the coding will be done in Python.
4. The Excel files are imported to Python, where the initial Excel database is taken as Input and later as target.
5. Training is performed on 70% of the data and testing on the remaining 30% of the data.
6. Initially, random weights are assigned all the entries in the database.
7. The weighted sum of all input nodes is then calculated and the activation for the node is calculated.
8. Necessary adjustments will be done according to error difference obtained between target output and achieved output.



Figure No. 8: Team Selection

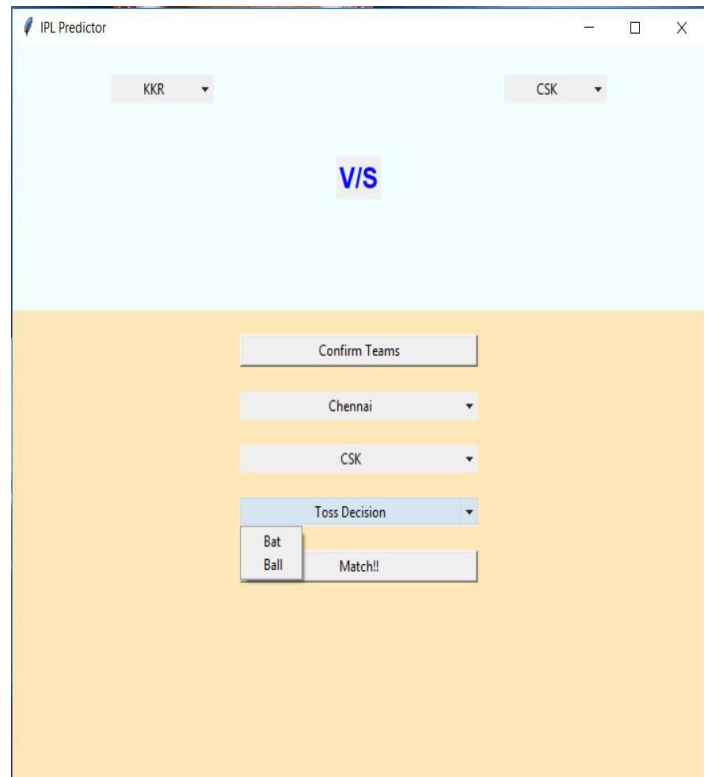


Figure No. 9: Toss Decision



Figure No. 10: Match Winner and Winning Team Score Prediction

Outputs:

We have used accuracy as a metric to evaluate correctness for all the algorithms.

Win-Loss prediction:

We obtained following accuracy for –

- 1) Support Vector Machines: 69.08%
- 2) Naïve Bayes Classifier: 64.70%
- 3) Random Forest Classifier: 51.55%

The algorithm with the maximum accuracy was used to predict the winning team for an IPL match.

Score Prediction:

Accuracy of neural network: 62.02%

Using the above model with given accuracy, the score for the winning team was predicted.

Outcomes:

The project intends to correctly analyse the input data on the basis of 14-16 parameters to correctly judge the winning team and its score. It uses complex algorithms like Naïve Bayes, Decision Tree, and Random Forest.

Discussion:

AI systems are now a part of a variety of specializations from stock trading to the setting of real estate prices; from fraud detection to translation of vernacular languages; from designing our daily shopping lists to forecasting which movies we might like and diagnosis of medical condition.

This is merely the beginning. Not long from today, AI will be used to advance our understanding of biological phenomenon through analysis of highly dense and huge datasets, help us invent new drugs and make treatments more thorough and personal. Self-driving vehicles can revolutionize transportation and allow new paradigms in planning of architecture. Machines will automate our homes more efficiently, make businesses more productive and help predict risks to society.

There are several techniques that can be used to correctly classify and predict when a new tuple is entered.

Future Scope:

Using neural networks and machine learning our project aims to deliver an application to predict the success of a team in an IPL match. By using neural networks and machine learning our project can be extended to develop a world cup-based application or for any other international sports tournaments.

To predict the success of the team we are taking the statistics of every innings of batsman as well as the bowler through every IPL season from 2008-2018. Also, the team statistics at a particular venue are considered. The prediction will be based on the outcomes to all the IPL matches in the past 11 years.

Since neural networks work on experience, we will be providing more than 600 different matches to our network, each with complete ball by ball data. This will help to gain accurate results. The number of matches won with a particular team combination, the average score of the team as well as the progress of the team through the season will be the deciding parameters.

Future research in this project can be predicting the bowl by bowl outcome of a match while the match is being played. This will require real time data collection as well as statistical measures for successful prediction.

Deployment:

Deployment can be a major challenge in software which involves machine learning. Huge amounts of processing is required just to get it up and running. Data generated in the process is humongous as well.

The ML model will be deployed using Google firebase or spring. The cloud platform will provide a base for future developments where flexibility, portability and reliability are required.

The GUI by itself is not very difficult to deploy. We have multiple ways of offering interfaces, such as web interfaces and command line interfaces. It will expand to a desktop GUI in the future as well.

The web interface will be deployed using Heroku or AWS, depending on the relative cost of hosting.

The command line can be used as a windows executable file. Users can simply download it and run it natively on their windows PCs.

Conclusion:

By observing recent trends and strategies, the scope of machine learning is growing day by day. The project based on machine learning will evaluate and improve the understanding by using this project as a path. By using the technologies like ml, python, neural networks.

Using our derived set of features, we will be able to predict the outcome of a match with approximately 65% accuracy, and the score with 60% accuracy. Predicting the outcome as well as the score of a cricket match has multiple factors involved, some of which we have tried to incorporate, others being difficult to correctly specify and represent in this project. Greater prediction accuracies can be achieved through more define data that better correlates to features significant for an IPL Match.

References:

- [1] Duckworth, Frank C., and Anthony J. Lewis. "A fair method for resetting the target in interrupted one-day cricket matches." *Journal of the Operational Research Society* 49.3 (1998): 220-227.
- [2] Beaudoin, David, and Tim B. Swartz. "The best batsmen and bowlers in one-day cricket." *South African Statistical Journal* 37.2 (2003): 203.
- [3] Lewis, A. J. "Towards fairer measures of player performance in one-day cricket." *Journal of the Operational Research Society* 56.7 (2005): 804-815.
- [4] Swartz, Tim B., Paramjit S. Gill, and David Beaudoin. "Optimal batting orders in one-day cricket." *Computers and operations research* 33.7 (2006): 1939-1950.
- [5] Norman, John M., and Stephen R. Clarke. "Optimal batting orders in cricket." *Journal of the Operational Research Society* 61.6 (2010): 980-986.
- [6] Kimber, Alan." A graphical display for comparing bowlers in cricket." *Teaching Statistics* 15.3 (1993): 84-86.
- [7] Barr, G. D. I., and B. S. Kantor." A criterion for comparing and selecting batsmen in limited overs cricket." *Journal of the Operational Research Society* 55.12 (2004): 1266-1274.
- [8] Van Staden, Paul Jacobus." Comparison of cricketers bowling and batting performances using graphical displays." (2009).
- [9] Sankaranarayanan, Vignesh Veppur, Junaed Sattar, and Laks VS Lakshmanan." Auto-play: A Data Mining Approach to ODI Cricket Simulation and Prediction." *SDM*. 2014.
- [10] Khan, Mehvish, and Riddhi Shah." Role of External Factors on Outcome of a One Day International Cricket (ODI) Match and Predictive Analysis."