

# Cricket Score Prediction using Machine Learning Algorithms

**Sudhanshu Akarshe**

*Student*

*Department of Computer Engineering*

*D Y Patil Institute of Technology, Pimpri, Savitribai Phule Pune University, Pune, India*

**Rohit Khade**

*Student*

*Department of Computer Engineering*

*D Y Patil Institute of Technology, Pimpri, Savitribai Phule  
Pune University, Pune, India*

**Nikhil Bankar**

*Student*

*Department of Computer Engineering*

*D Y Patil Institute of Technology, Pimpri, Savitribai Phule  
Pune University, Pune, India*

**Prashant Khedkar**

*Student*

*Department of Computer Engineering*

*D Y Patil Institute of Technology, Pimpri, Savitribai Phule  
Pune University, Pune, India*

**Prof. Prashant Ahire**

*Professor*

*Department of Computer Engineering*

*D Y Patil Institute of Technology, Pimpri, Savitribai Phule  
Pune University, Pune, India*

## Abstract

Cricket is 11 player team sport played on ground. Cricket has huge fan base in India. With such great spectator support and many people try to predict the outcome of matches based on their individual cricket sense. The game has certain rules and scoring system. Factors viz, match venue and individual player performance have great impact on outcome of the match. Such various parameters are highly interdependent on each other which makes it difficult to make precise prediction of the match. In this project, we are going to build a prediction system that takes in data of matches played in past and makes a prediction of future match events such as final score and results in a victory or loss. Our system will predict match outcome by analysing pre-stored match data using various machine learning algorithms. We intend to use more features such as pitch condition, weather condition, outcome of toss, individual player performance with respect to match venue. Our system finally present quantitative results displayed by best suited algorithm having highest accuracy. Also, demonstrating the performance of our algorithms in predicting the number of runs scored which is one of the most important parameter of match outcome.

**Keywords-** Prediction System, Machine Learning Algorithms, Score Prediction

## I. INTRODUCTION

Cricket started in the 16th century in England. Cricket is a sport with multiple formats, different playing standards and varying duration. Twenty20 is one of the three current forms of Cricket which is recognized by the International Cricket Council (ICC). In that format, two teams have a one innings each with a maximum of 20 overs. Because of the short time duration and the excitement, it generates, Twenty20 cricket has become such a huge success. There are many annual tournaments conducted at both domestic and international level. There is huge commercial interest in player performance prediction. This has motivated many analyses of individual and team performance, as well as prediction of future games, across all formats of the game. Currently, strategists rely on a combination of player experience, team constitution and "cricketing sense" for making instantaneous strategic decisions. We choose to focus our testing and evaluation on the most popular format, namely ODI.

Principles of Machine Learning are used for developing the prediction system. There are two types of machine learning namely supervised machine learning and unsupervised machine learning. In supervised machine learning we must train the machine by providing huge data sets and the outcomes. In this project, we use various classification algorithms such as Naïve Bayes, Support Vector Machine, K-Means also Logistic Regression for regression analysis and Decision Tree algorithms for making effective decisions feature selection.

## II. PROPOSED SYSTEM

The main aim is to predict the match outcome, performance of each player based on the historical data. In order to achieve a reliable accuracy, we need to analyse a large amount of data. Therefore, the initial step of the implementation was to collect data for all possible matches. Dataset is collected from various websites such as ESPN, Kaggle, etc.

We collected data which provides ball by ball details for all the matches. Various analytical rules are used in order to filter collected data with respect to the selected feature.

Features can be match venue, playing 11, weather condition, performance of individual player.

Further, cleaned data is split into training (80%) and testing data (20%). Training data is fed to all machine learning model and accuracy of each model is noted. Model having highest accuracy is selected for further prediction.

This model is used to make prediction over the input data provided, input data may include – home team, opposition team, current weather condition and analysis done on historical data.

### III. SYSTEM ARCHITECTURE

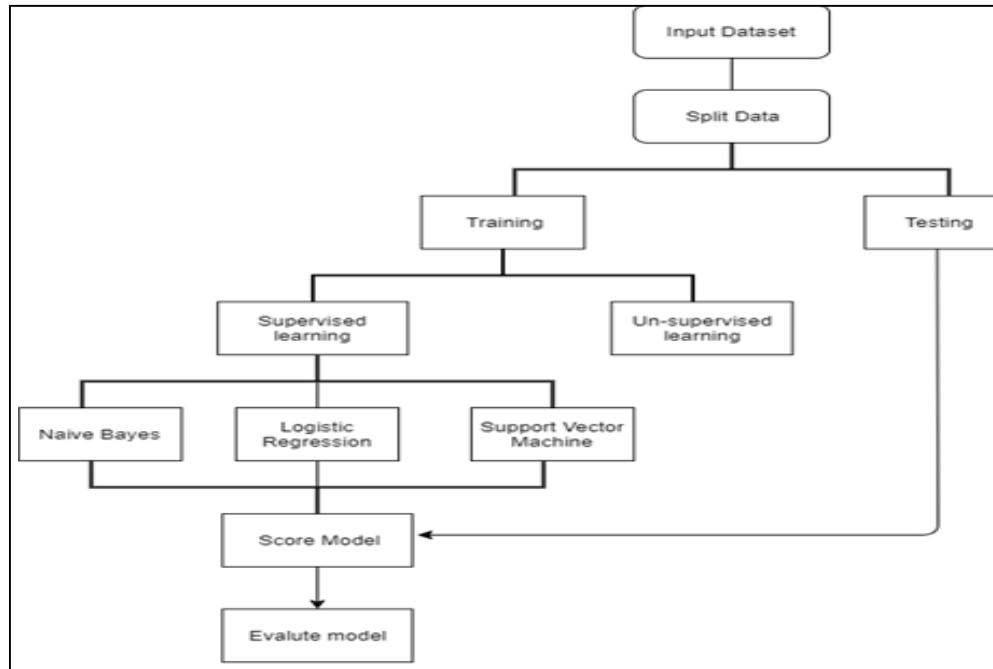


Fig. 1: Architecture of Cricket Prediction System

### IV. WORK FLOW DIAGRAM

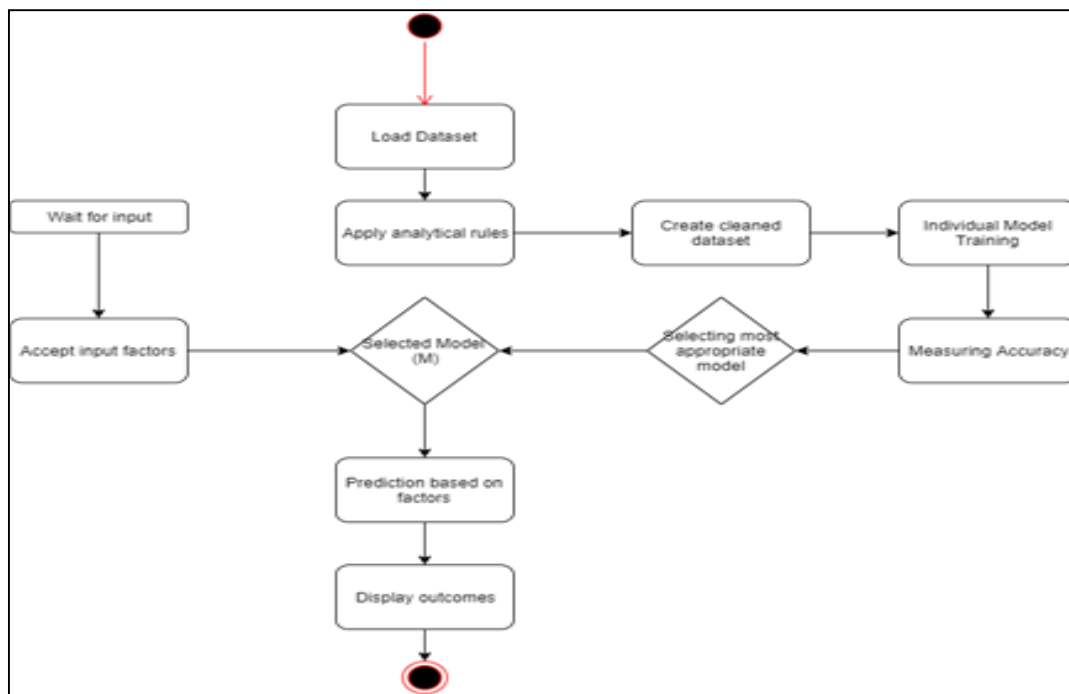


Fig. 2: Work Flow Diagram

## V. TOOLS

### A. TensorFlow

TensorFlow is software framework for machine learning, dataflow and differentiable programming for wide range of tasks. TensorFlow is open source software which consist of libraries, flexible ecosystem of tools and other resources which is helpful for developers to build and deploy Machine Learning applications easily and effectively. Researchers and engineers working at Google developed this software under Google Brain project for their internal use. On 9th November, 2015 it was released as open source project under Apache license 2.0. Flexible architecture of TensorFlow allows easy development of computation across different platforms ( CPUs, GPUs, TPUs ).

### B. Python

Python is open source, high level, interpreted programming language which is dynamically typed and garbage collected. Python can used in software development, statistics, and system scripting and web development. Python has rich source of libraries, packages, modules and repositories which makes it powerful in handling big data and machine learning operations.

## VI. RESULTS

### A. Performance Measures

To evaluate performance we use confusion matrix of predicted score. Values of True Positive(TP), True Negative(TN), False Positive(FP), False Negative(FN) are filled in confusion matrix.

#### 1) Confusion Matrix

		Predicted	
		Negative	Positive
Actual	Negative	True Negative	False Positive
	Positive	False Negative	True Positive

- Accuracy Score: Comparison of the expected outcomes with the predicted outcomes of the system for a given input dataset is done. Higher the rate of actual true labels in testing data matching with corresponding set of predicted labels, higher will be the accuracy

For measuring the success of the prediction, we use precision-recall index.

- Precision Score: This is defined as the number of True Positives (Tp) divided by the sum of True Positives and False Positives (Fp)

$$P = \frac{Tp}{Tp + Fp} \quad (2)$$

Best value: 1 and Worst value: 0.

- Recall Score: This is defined as the number of True Positives (Tp) divided by the sum of True Positives and False Negatives (Fn).

$$P = \frac{Tp}{Tp + Fn} \quad (3)$$

- Accuracy: Total accuracy can be easily evaluated using confusion matrix. Ratio of sum of True Positives(Tp) and True Negative(Tn) to that of sum of True Positives(Tp) +True Negative(Tn)+False Positives(Fp) and False Negative(Fn)

$$A = \frac{Tp + Tn}{Tp + Tn + Fn + Fp}$$

## VII. FUTURE SCOPE

Traditionally, single algorithm is used in such prediction system and respective performance is measured. Instead, we intend to use multiple machine learning algorithms and relatively measure their performance.

## ACKNOWLEDGEMENT

In this way, with the help of prediction system we can predict the outcome of matches on prior basis.

## REFERENCES

- [1] "M. Ahmad, A. Doud, L. Wang, H. Hong, October 2016, Prediction of rising stars in cricket".
- [2] "P. Somaskandhan, G. Wijesinghe, L. Bashitha, 2017, Identifying optimal set of attributes that impose high impact on end results of cricket match using machine learning."
- [3] "M. Rehman, O. Shamim, S. Ismail, October 2018, an analysis of Bangladesh One day international cricket: Machine learning approach".
- [4] "A. Tripathi, J. Vankar, B. Vaje, V. Varekar, Cricket Score Prediction system using clustering algorithm.
- [5] M. Bailey and S. Clarke, "Predicting the Match Outcome in One Day International Cricket Matches, while the Game is in Progress", Journal of sports science & medicine, vol. 05, no. 04, pp. 480-487, 2006.
- [6] N. Pathak and H. Wadhwa, "Applications of Modern Classification Techniques to Predict the Outcome of ODI Cricket", Procedia Computer Science, vol. 87, pp. 55-60, 2016.
- [7] P. Satao, A. Tripathi, J. Vankar, B. Vaje and V. Varekar, "Cricket Score Prediction System (CSPS) Using Clustering Algorithm", International Journal of Current Engineering and Scientific Research, vol. 03, no. 04, pp. 43-46, 2016.
- [8] [https://web.archive.org/web/20130120040151/http://www.icccricket.com/match\\_zone/historical\\_ranking.php](https://web.archive.org/web/20130120040151/http://www.icccricket.com/match_zone/historical_ranking.php)