IPL Score Prediction using Machine Learning Algorithm

Abstract

The IPL season is going on and we are all eager to know who will win the match beforehand and in the media, with the power of Machines and Deep Learning, you can do these types of amazing stuff. How you can train a model from scratch and embed it in the web app using simple and powerful libraries like sklearn, pandas, and flask. Also, some web development is involved.

Domain

Machine learning (ML) -

It is a type of artificial intelligence (AI) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so.

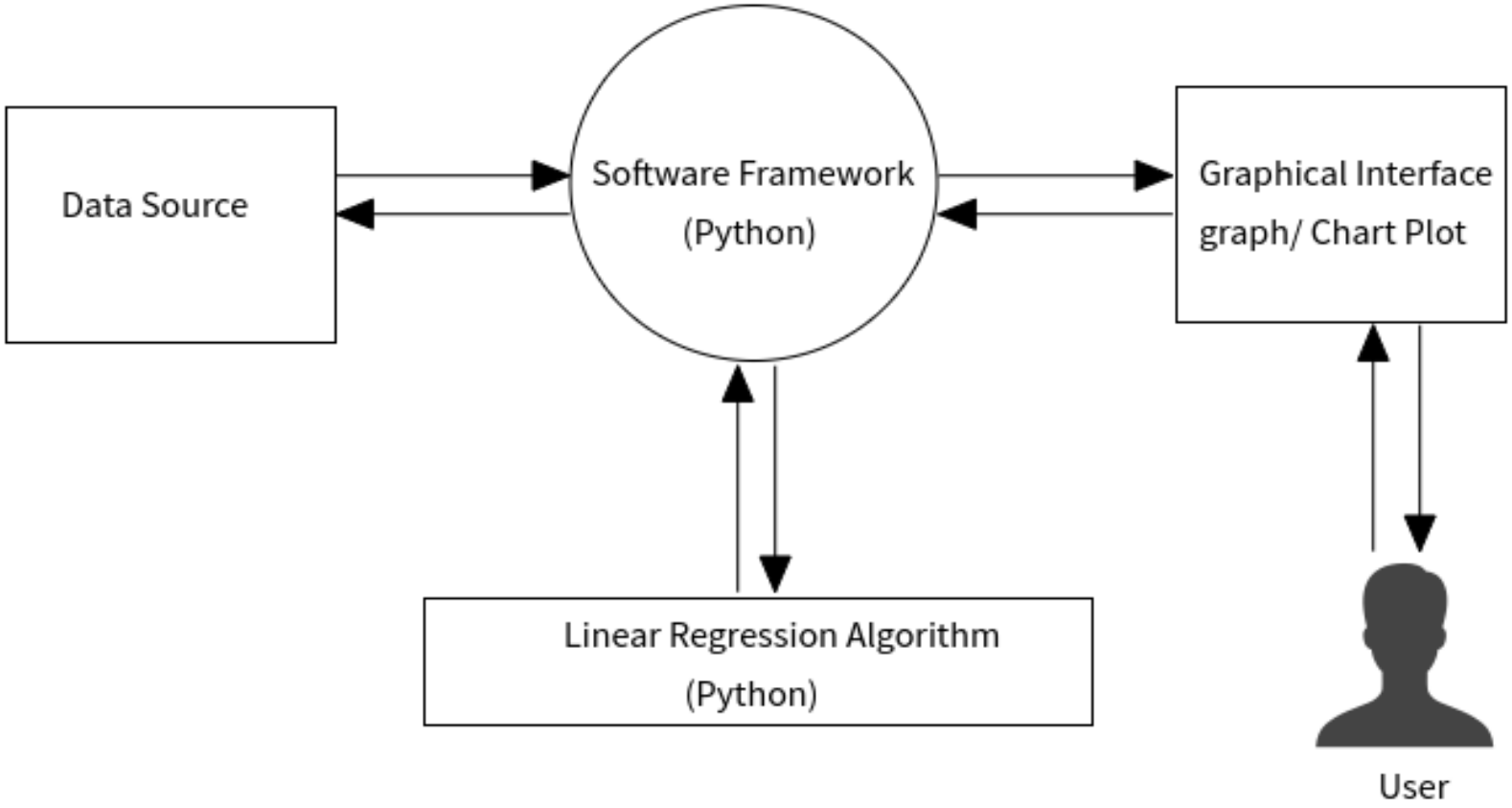
Problem Statement

- Using the IPL dataset, Predict the score of your favourite team.
- Let us read the data file and take an initial look at the data. You can access the data.

Proposed System

- Given IPL datasets of past 9 years, the main objective of this paper is to predict the outcome of an IPL match between two teams based on the analysis of previously stored data using Machine Learning algorithms.
- The information will be analyzed and pre-processed. After pre-processing the data will be used to train different models in order to give the outcomes.
- We will analyze the various datasets and use key variables such as strike rate, bowler economy, etc. and feed it as input to an algorithm will help us get the probable outcome of the match.

System Architecture



Algorithm

Linear regression:

- To predict the continuous values, Linear regression is used.
- Certain known parameters are given to the machine learning algorithms, it predicts the continuous values as output.
- It cannot used for the classification problems.
- The proposed model predicts the score using the Linear Regression

Random forest:

- Random forest classifier creates multiple decision trees and find out the individual output.
- It combines all the results together and give the results with more accuracy.
- It can be used as both classification and regression.

Support Vector Machine(SVM):

- A SVM classifier a vector space based machine learning method where the goal is to find a decision boundary between two classes that in maximally far from any point in the training data.
- It is evident from the results that SVM gives us a higher accuracy of 66.23% than other algorithms for this data distribution.

List of Modules

- 1- **Data Preprocessing** and some **Exploratory Data Analysis** to understand the data
- 2- **Data preparation**:1 **Feature Engineering** and Scaling
- 3- **Feature Selection** using **RFE** and Model Building
- 4- **Regression Assumptions** Validation and **Outlier Removal**
- 5- **Rebuilding the Model** Post Outlier Removal: Feature Selection & RFE
- 6- **Removing Multicollinearity**, Model Re-evaluation and Assumptions Validation

SCREENSHOT

First Innings Score Predictor for Indian Premier League (IPL)

A Machine Learning Web App, Built with Flask.

--- Select a Batting team ---

--- Select a Bowling team ---

--- Select a Venue ---

Overs (>= 5.0) eg. 7.2

Runs eg. 64

Wickets eg. 4

Runs scored in previous 5 Overs eg. 42

Wickets taken in previous 5 Overs eg. 3

Predict Score

Conclusion

There is a lot of room for improvement our model.it

Is all up to individual imagination. deploying model is a good way to showcase to the world what one is capable of long with fixing real-life.

As can be seen following step in procedural manner can simplify our problem solving and is generally preferred in the industry.

With this,we have come to an end with this daunting article/guide.