

ABSTRACT

Data Mining and Machine Learning in sports analytics, is a brand-new research field in computer science with a lot of challenge. Machine Learning and Data Science are one of the fastest-growing technological fields. This field results in amazing changes in the medical field, production, robotics etc. The main reason for the advancement in this field is the increase in the computational power and availability of large amounts of data. In Data Science, this data is analysed and made suitable for creating machine learning models and products.

In our model we will work on IPL Data Analysis and Visualization Project using Python where we will explore interesting insights from the data of IPL matches like most run by a player, most wicket taken by a player, and much more from IPL season 2008-2020. Different Machine Learning and statistical approach will be taken to find out the best possible outcome. A very popular mathematical technique named random forest Classifier, Multiple Linear Regression is used in order to make comparison of results found. Also we are going to discuss the IPL team win prediction. Based on some match stats, we're predicting who is the winner of an IPL match. Through this project, you will get familiar with the exploratory data analysis and feature engineering techniques that need to be applied to process data.