

 Marwadi University	Marwadi University Faculty of Technology Department of Information and Communication Technology	
Subject: Machine Learning (01CT1509)	Topic: IPL match outcome prediction	
Project Summary	Date:	Enrollment no.: - 92110133003, 92000133016

Summary: -

First of all, we add some Libraries as we needed to run our project then we uploaded dataset file to google drive and then we connect it drive with colab so now we have access of our data. Now we read "matches.csv" file. "matches.csv" have 18 features and 756 raw data. Now we are checking null values from the dataset and then change team name with some updated team name. We print bar chart of IPL matches won so from the bar chart we can easily visualize which team won most the match and then we plot bar chart of that stadium where most of the team won there so like this, we can easily visualize how many teams won match at this stadium. We make the bar chart of toss decision where we can know about after won the toss what is team going to choose fielding or bating. We print data of where IPL match played in the city and print toss decision, print result of the match and print dl_applied in the match like this type of data we are printing. Now we are going to check which classification we can use. We are printing names of "man of the match" of top 5 players and plot a bar chart of it how many times that player being a "man of the match". We are printing number of toss win by each team and extract the records where a team won batting first and make a histogram of it. Now we are going to find out the number of wins each team after batting first and make a bar plot for top 3 teams with most wins after batting first and make a pie chart also. Now we also extract those records where a team has won after batting second and we make a histogram for frequency of wins by number of wickets and we finding out the frequency of number of wins by each time after batting second and make a bar plot for top-3 teams with most wins after batting second and also make a pie chart also of second inning. We also print number of matches played each season. We also take a dataset from drive name of "deliveries.csv" file it has 179079 raw data and 21 features. We also print only match_id from that dataset. Now from all of this we shorted data with the match_id 1 and 2 and compare it so after sorted we have 21 features and 248 raw data now, we count runs of batsman and printed also and also printed how they out like caught, bowled and run out so we print that in a number. We take sklearn libraries for to get accuracies of train model so we can take best classification from that. We plot bar chart of total number of matches played in each season now we are making short forms from the test data which are used in full form in training data. Also, there have been changes in the names of the teams and different teams have participated in different years which we need to remove. We print winner team name. We change text to numerical using LabelEncoder and train model. We check algorithms and choose only that which gives most accurate rate so we got new two algorithms one is Random Forest and Second is Decision Tree but we choose Random Forest because in that we don't have to make and divide data into many models in Random Forest we can easily do in one model and In Decision tree we have to make many models so it takes to much time to get the output so that's why we take Random Forest Classification. We also take the dataset of IPL

 Marwadi University	Marwadi University Faculty of Technology Department of Information and Communication Technology	
Subject: Machine Learning (01CT1509)	Topic: IPL match outcome prediction	
Project Summary	Date:	Enrollment no.: - 92110133003, 92000133016

2020 because now by using Random Forest Classification we are going to predict the matches of IPL match of 2020. In "Testset Matches IPL 2020.csv" file there is listed of team names, venue, city, etc. Now we are going to sorting and comparing previous data and live data so we can get our predicted team like this team is going to win this match. From all of this we learn how we can train our models and tested it and then find the accuracy and from the train model we list out which team are going to win this match in IPL season 2020.

Drive Link: -

<https://drive.google.com/drive/folders/1DfKO-PSj5ZbOeHXisDBO4D7-iZFmXHSB?usp=sharing>