CriCat match win predictor application For international cricket matches

ABHYUDAY RAI*, ADITYA RAJ*, ADITYA SINGH*
School of Computer Science Engineering and Technology, Bennett University
E21CSEU0130, E21CSEU0143, E21CSEU0152

Abstract

Whenever we talk about data and statistics in real life, we occasionally see their practical use. It is not that if they are used less by us, means they are used less by others too. Probability and Statistics play a very heavy role in the sectors of science, technology, and also many other things. Finally, our goal would be to use statistical analysis and probability tools in real life for solving some practical purpose, which here is predicting the result of a cricket match.

Introduction

Using Data Analysis we are creating a cricket match win predictor, which would use tons of data, major statistical tools, and concepts to predict which team would have the highest probability of winning the game. Herein, the application would be taking data like team name, player 11, toss result, place of the match, etc. to predict the most probable result for the match.

Now all these mentioned factors don't have the same values in everybody's opinion, some people believe the past records to be the most important criterion for a team to win, whereas some believe it's their playing 11. So our goal is to focus on this factor and make it user based.

Proposed Method

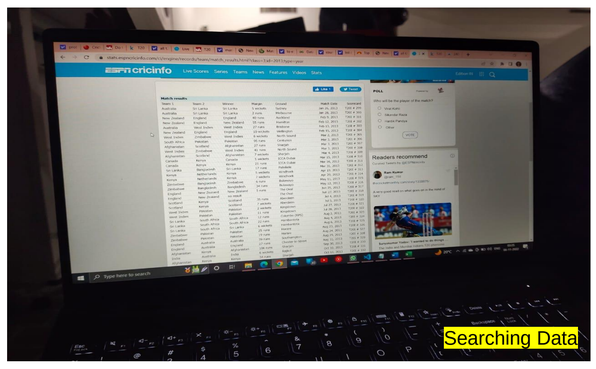
HOW WE GET THERE

1. COLLECTION OF DATA
- Collecting players' stats of each major team playing international matches
- Stats would be collected separately for diff formats
- Player vs Player data would be collected, toss to match history, stadium, and country analysis to be done.

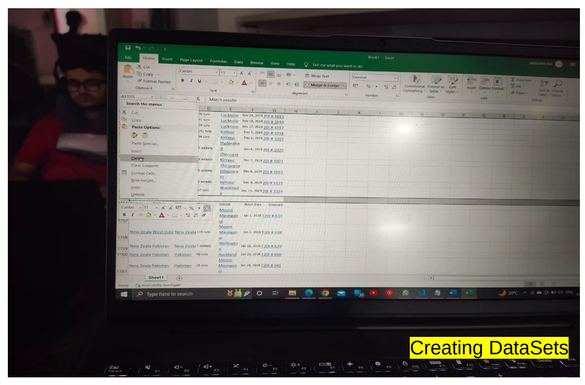
2. APPLYING OPERATIONS
- Filtering out relevant data and creating useful datasets to be able to perform functions efficiently.
- Apply required statistics on clean data to get the required results.
- Would be applied only on playing 2 teams.

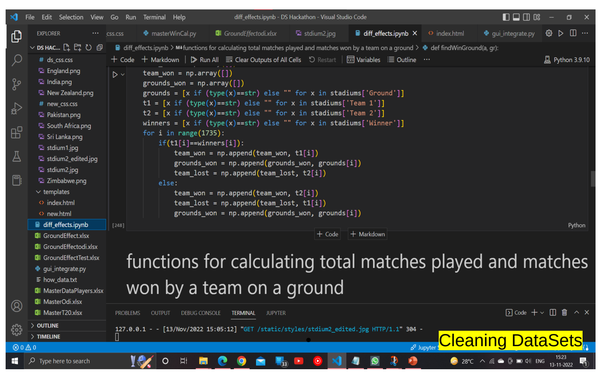
3. MAKING THE PROGRAM AND GUI
- Once the things at the back are running fine, would focus on GUI creation for a better showcase of results.
- GUI will interact with the user to take relevant inputs and provide the output.

Tools used: Python, HTML, CSS, Flask lib., Pandas lib., SciPy lib., NumPy lib., MS Excel;



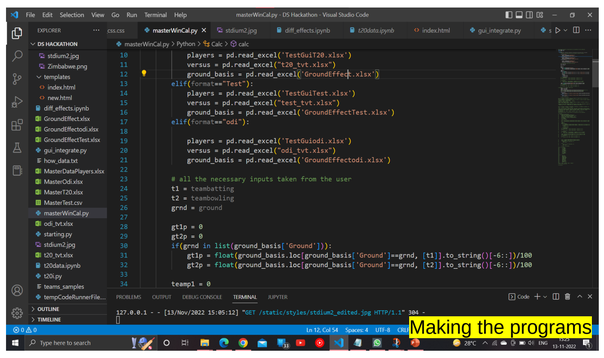
The Process



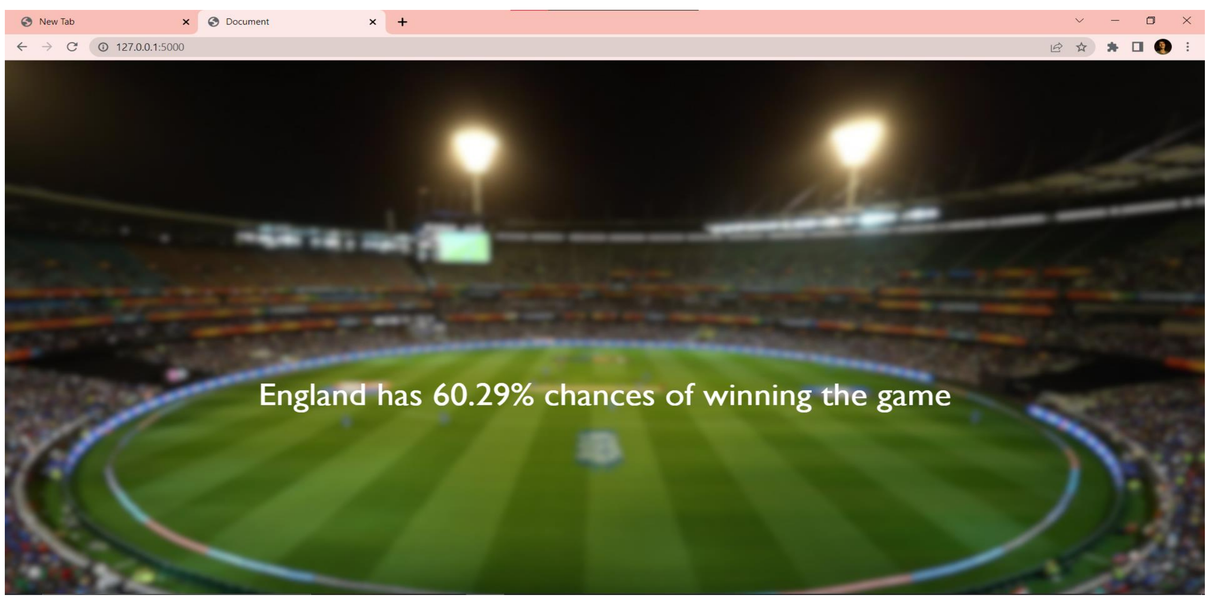


	against	Afghanistan	Australia	Bangladesh	England	India	New Zealand	Pakistan	South Africa	Sri Lanka	Zimbabwe
0	Afghanistan	0.000	100.000	33.333	100.000	100.000	100.000	100.000	100.000	75.000	7.692
1	Australia	0.000	0.000	44.444	55.000	62.500	38.462	52.174	31.250	46.154	0.000
2	Bangladesh	66.667	55.556	0.000	100.000	90.909	83.333	92.308	100.000	69.231	36.842
3	England	0.000	45.000	0.000	0.000	55.000	47.826	39.130	47.619	23.077	0.000
4	India	0.000	37.500	9.091	45.000	0.000	35.294	30.000	42.105	34.783	20.000
5	New Zealand	0.000	61.538	16.667	52.174	64.706	0.000	54.167	77.778	33.333	0.000
6	Pakistan	0.000	47.826	7.692	60.870	70.000	45.833	0.000	48.000	52.941	13.333
7	South Africa	0.000	68.750	0.000	52.381	57.895	22.222	52.000	0.000	35.000	16.667
8	Sri Lanka	25.000	53.846	30.769	76.923	65.217	66.667	47.059	65.000	0.000	0.000
9	Zimbabwe	92.308	100.000	63.158	0.000	80.000	100.000	86.667			

Creating MasterDataSets



Working and Results



Conclusion

As mentioned in the abstract and the introduction, the application works perfectly fine and give precise results if the factors are given proper weightage. Tough there are many flaws too, on which we would have to work and those are mainly related to frontend and backend integration.
Also, for the future plans, we plan the following add-ons in the application:
- Adding more dependencies like Weather factors, captaincy Records etc.
- Making it more user friendly, efficient and easy to use.
- Making it available for diff cricket leagues like IPL.