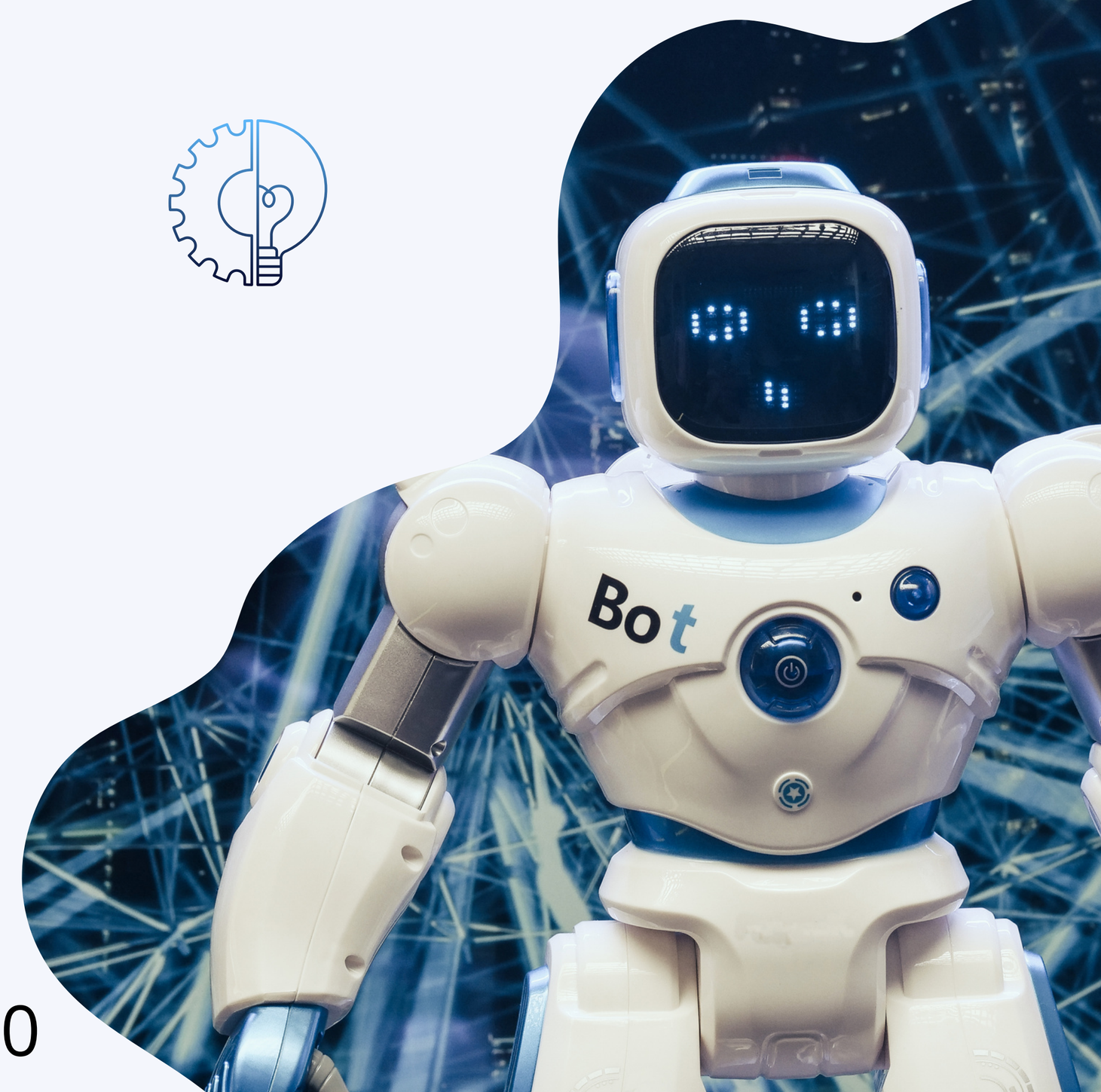
Cricket Match Winner Prediction

Sports Analytics

Submitted by-
S Surya-20BCE1071
Nilavan.I-20BCE1080



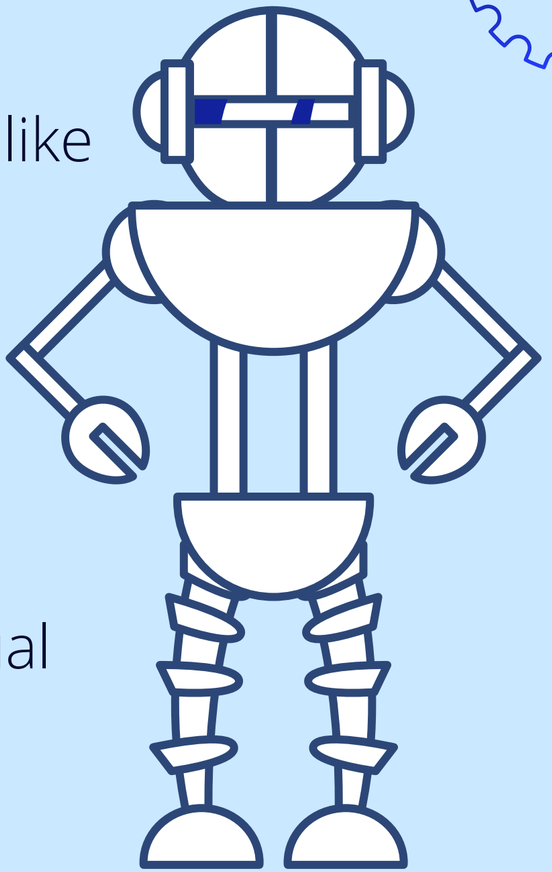
SCOPE

- In cricket, particularly the twenty20 format is most watched and loved by the people, where no one can guess who will win the match until the last ball of the last over.
- In India, The Indian Premier League (IPL) started in 2008 and now it is the most popular T20 league in the world.
- So we decided to develop a machine learning model for predicting the outcome of its matches.

OBJECTIVE

Winning in a Cricket Match depends on many key factors like

- home ground advantage
- past performances on that ground
- records at the same venue
- the overall experience of the players
- record with a particular opposition
- overall current form of the team and also the individual player



(continued.)

- Our main objective is to find the key factors that affect the match outcome and select the best machine learning model that best fits this data and gives the best results.
- Some works already have been published in this area of predicting the outcome of a cricket match. In some paper, only a few key factors are taken for prediction so the accuracy is less.
- Whereas in some paper the machine learning model is not appropriate. So it is important to take all the key factors that can affect the match outcome and as well as to select the best model for training and testing the data.
- This will increase the prediction accuracy drastically.

DATASET

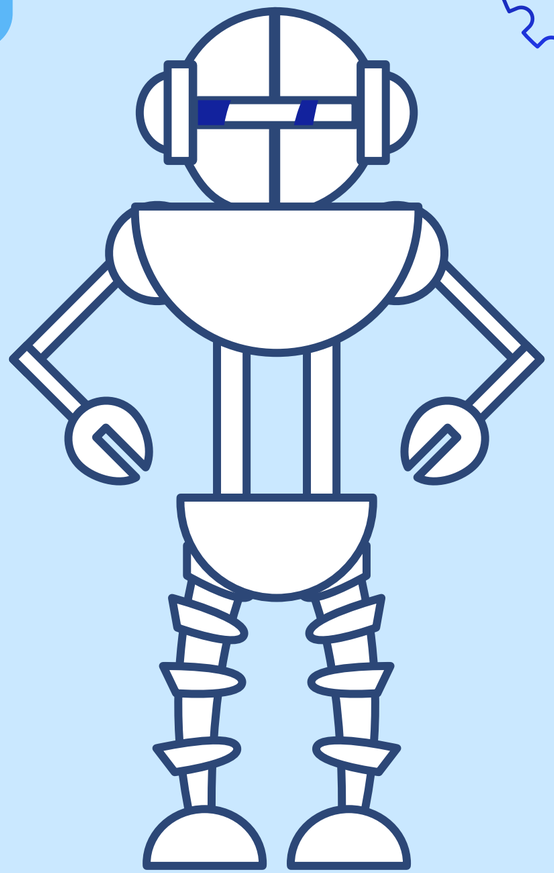
	id	season	city	team1	team2	toss_winner	toss_decision	result	dl_applied	winner	win_by_runs	win_by_wickets	player_of_match	venue
id														
1	1	2017	City1	Team1	Team5	Team5	field	normal	0	Team1	35	0	Player 467	Stadium1
2	2	2017	City2	Team2	Team4	Team4	field	normal	0	Team4	0	7	Player 405	Stadium2
3	3	2017	City3	Team3	Team6	Team6	field	normal	0	Team6	0	10	Player 85	Stadium3
4	4	2017	City4	Team4	Team8	Team8	field	normal	0	Team8	0	6	Player 145	Stadium4
5	5	2017	City5	Team5	Team7	Team5	bat	normal	0	Team5	15	0	Player 209	Stadium5

- This dataset contains the results of past IPL matches.
- It contains13 attributes with first 5 rows shown in the above figure.

METHODOLOGY

- Data Pre-Processing
- Model Building using

1. NAVIE_BAYES ALGORITHM
2. DECISION TREE REGRESSOR
3. SUPPORT VECTOR MACHINE
4. RANDOM FOREST ALGORITHM



CONCLUSION

- In this project, we selected 13 key features and 4 machine learning models and select the best possible prediction accuracy.
- We can extend this model to predict the outcome of twenty 20 matches, one-day international matches, and test matches also.
- This model can also be used for predicting the outcome of other sports also like football, hockey, tennis, baseball, rugby, etc.

Thank You