



Cricket Match Analytics and Prediction

- Brinal Jason Machado
- Manideep G
- Manish S B
- Saurav Khandelwal
- Sandeep Kumar

Introduction

1,000

- ▶ Cricket generates a plethora of data related to different aspects of the game.
- ▶ T20 being a popular format, opens opportunities for various statistical analysis.
- ▶ We attempted to describe some of the statistics for IPL and predict batsman performance in a fixture.

Goal

Generate and plot Batsman, Bowler, and Match analytics using Pandas and plotting libraries

Can result be predicted based on historical data?

Which model is suitable for prediction?

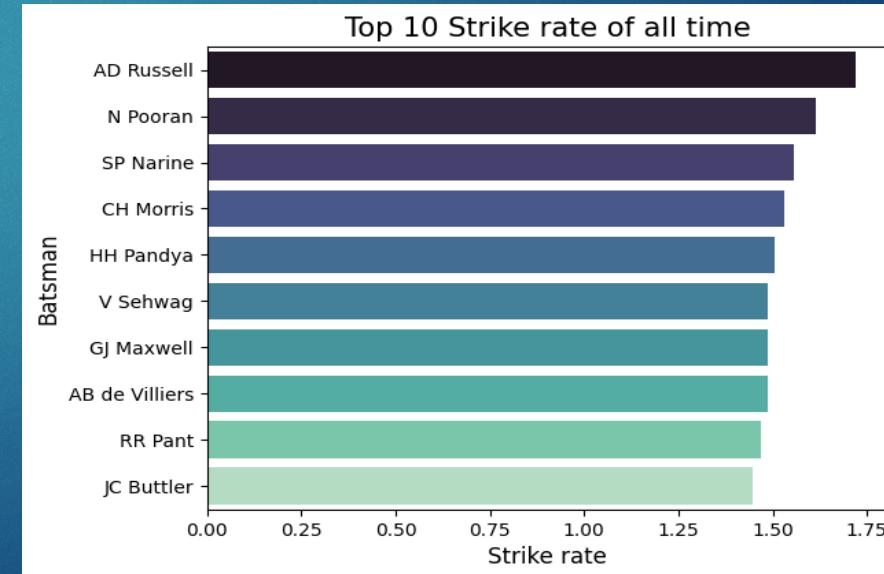
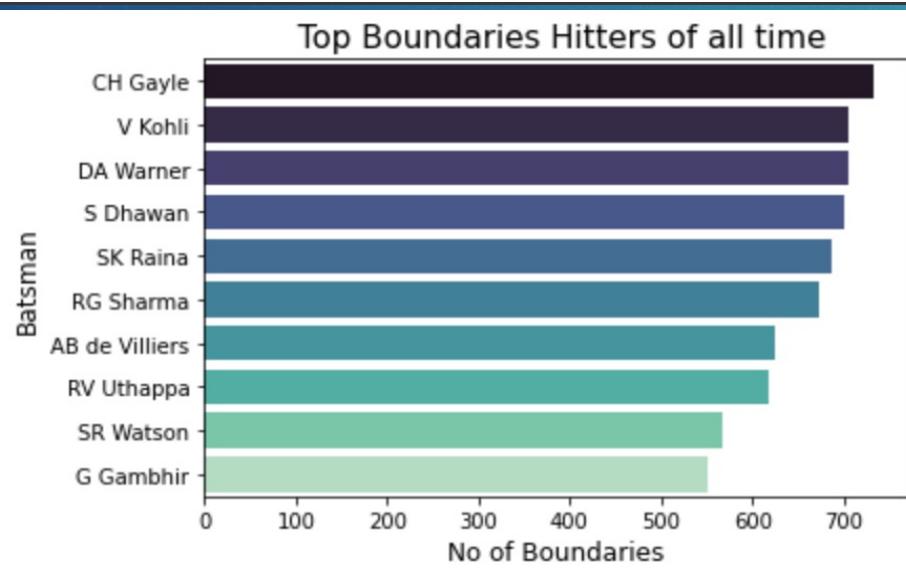
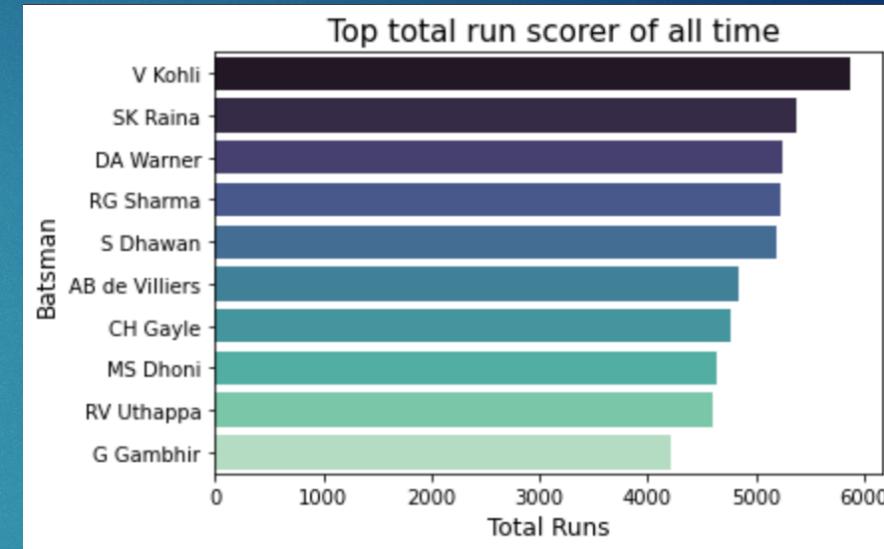
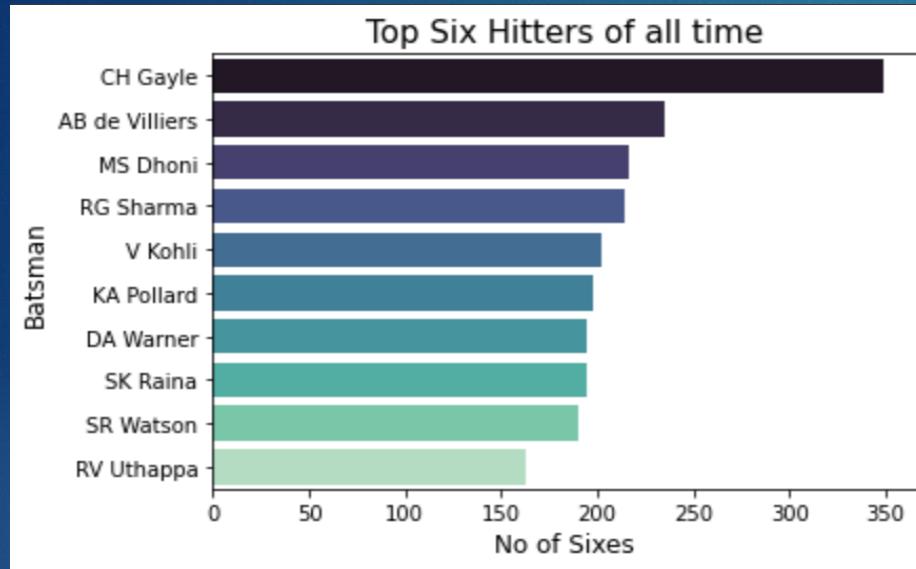
How accurate is the model?

Match Analytics

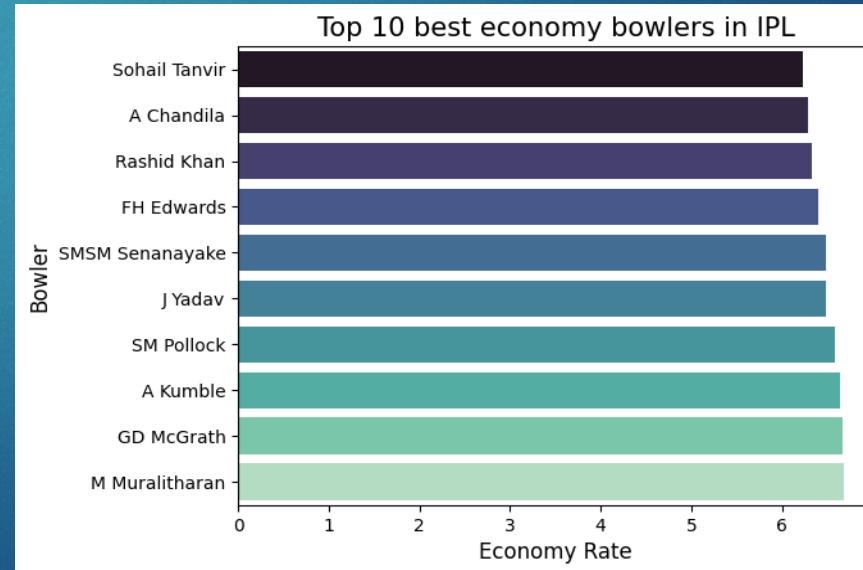
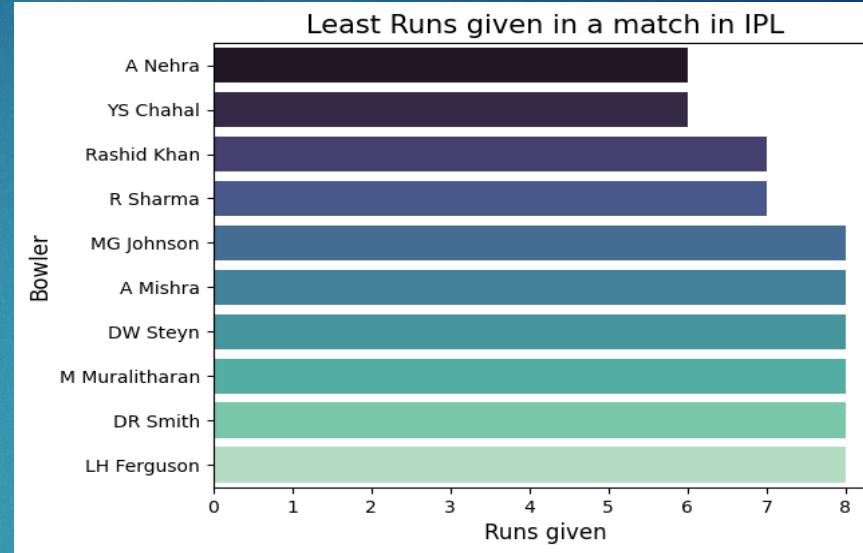
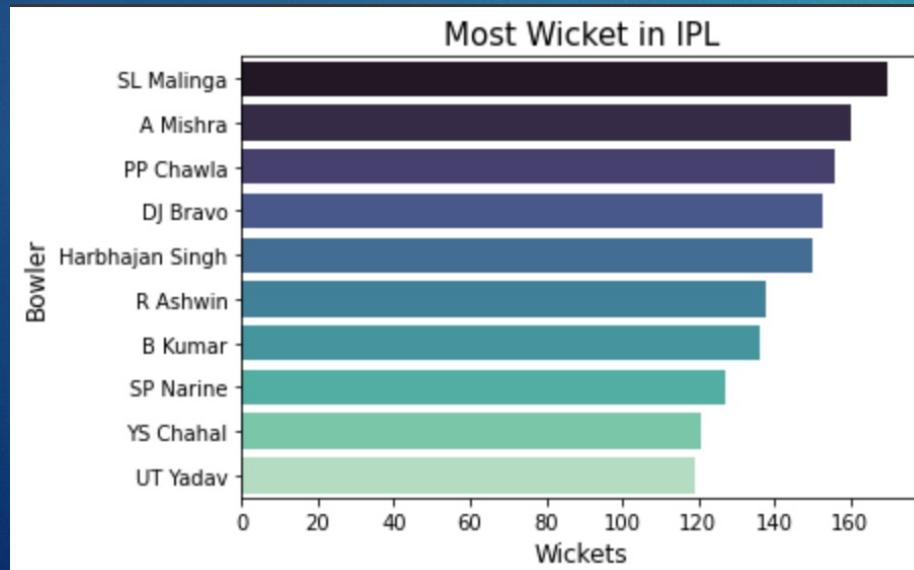
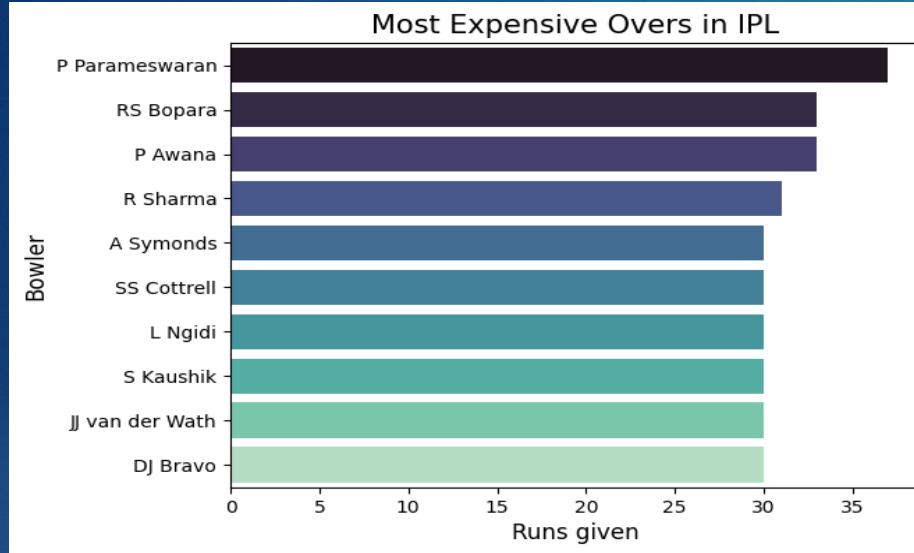
- ▶ Generate various statistics for batsman, bowler and match.
- ▶ Batsman statistics consists of highest run scorer, best strike rate, etc.
- ▶ Bowling statistics consists of best economy, highest wicket takers etc.
- ▶ Match statistics consists of toss winner, match winner given toss result etc.



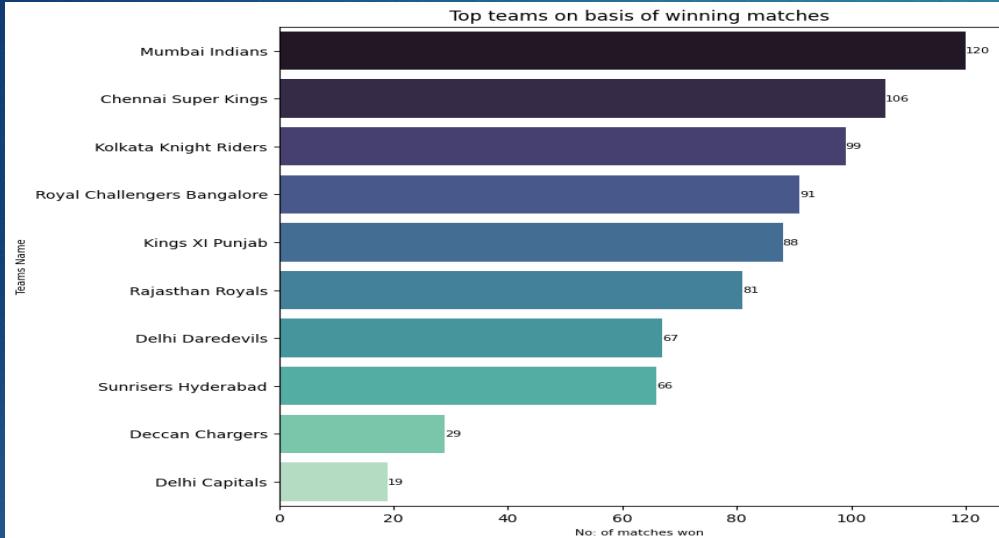
Batsman Analytics



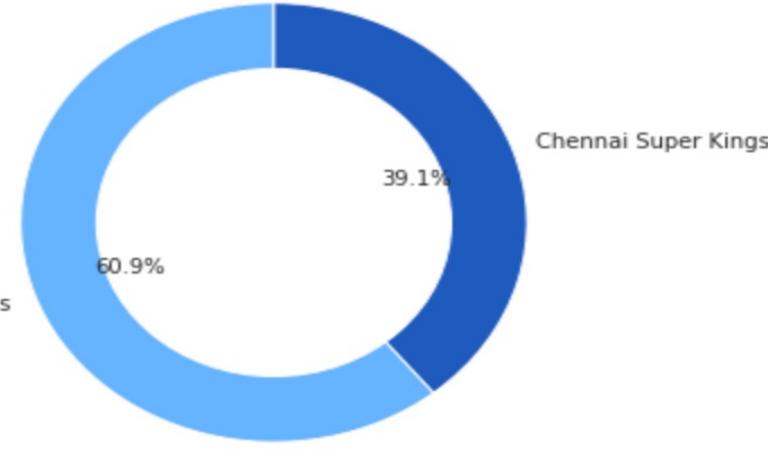
Bowler Analytics



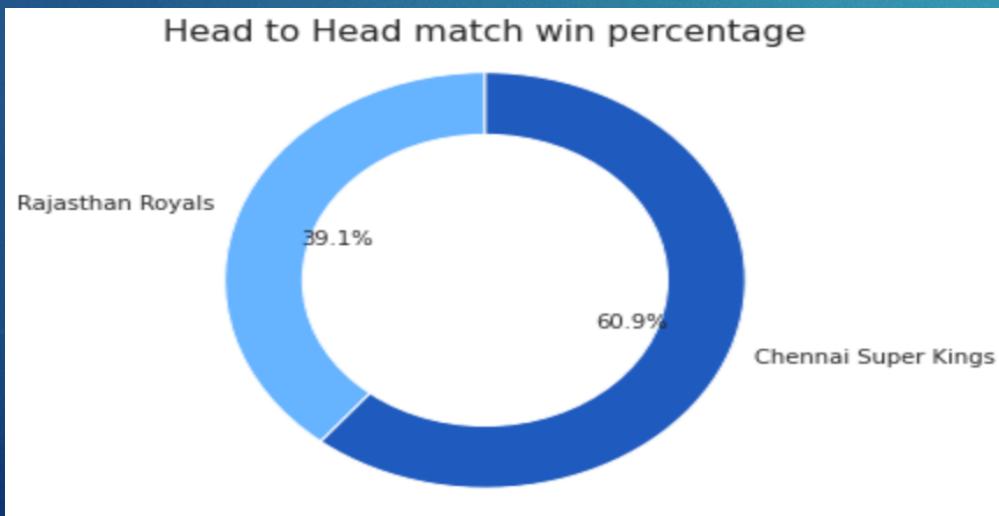
Match Analytics



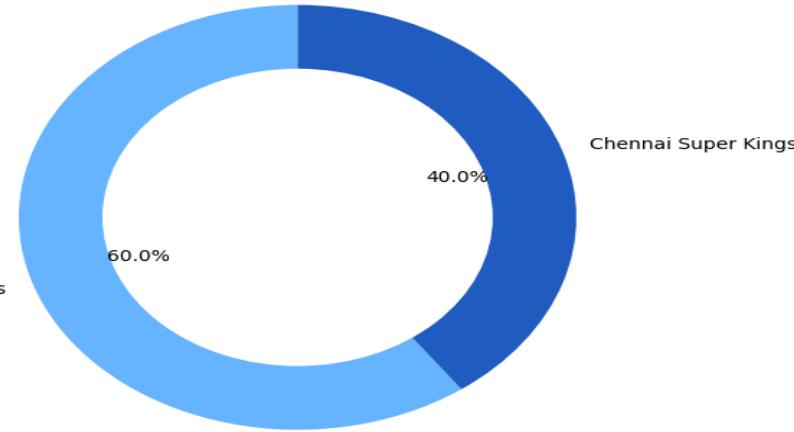
Head to Head toss win percentage



Head to Head match win percentage



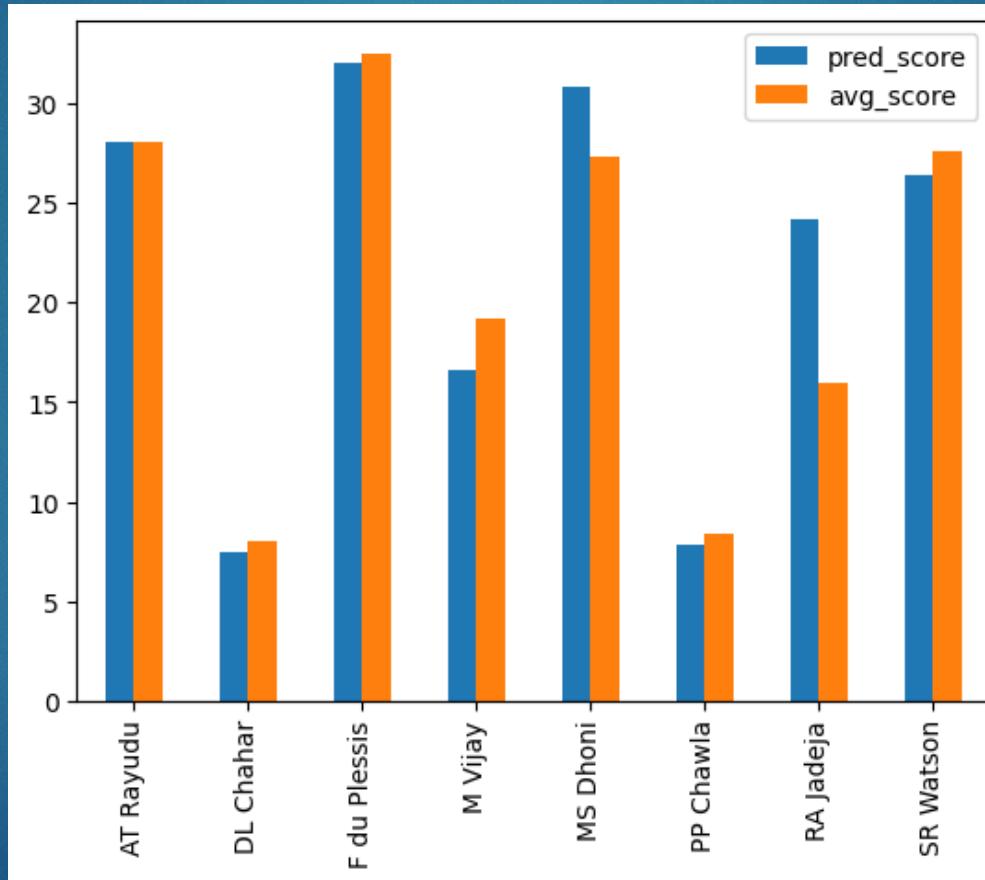
Head to Head match win percentage when toss is won



Prediction Model

- Apply different regression models on the generated match analytics
 - Model is trained on historical dataset
 - Can we predict team performance from the above models ?
 - Data Split: Train (First 80%) Test (Rest 20%)
 - Source: kaggle.com
- Model features:
 - ❖ Balls faced
 - ❖ Innings
 - ❖ Fours hit
 - ❖ Sixes hit
 - ❖ Batsmen Avg
 - ❖ Total runs
 - ❖ Strike rate
 - ❖ Runs per inning

Result - Batsman Run Prediction

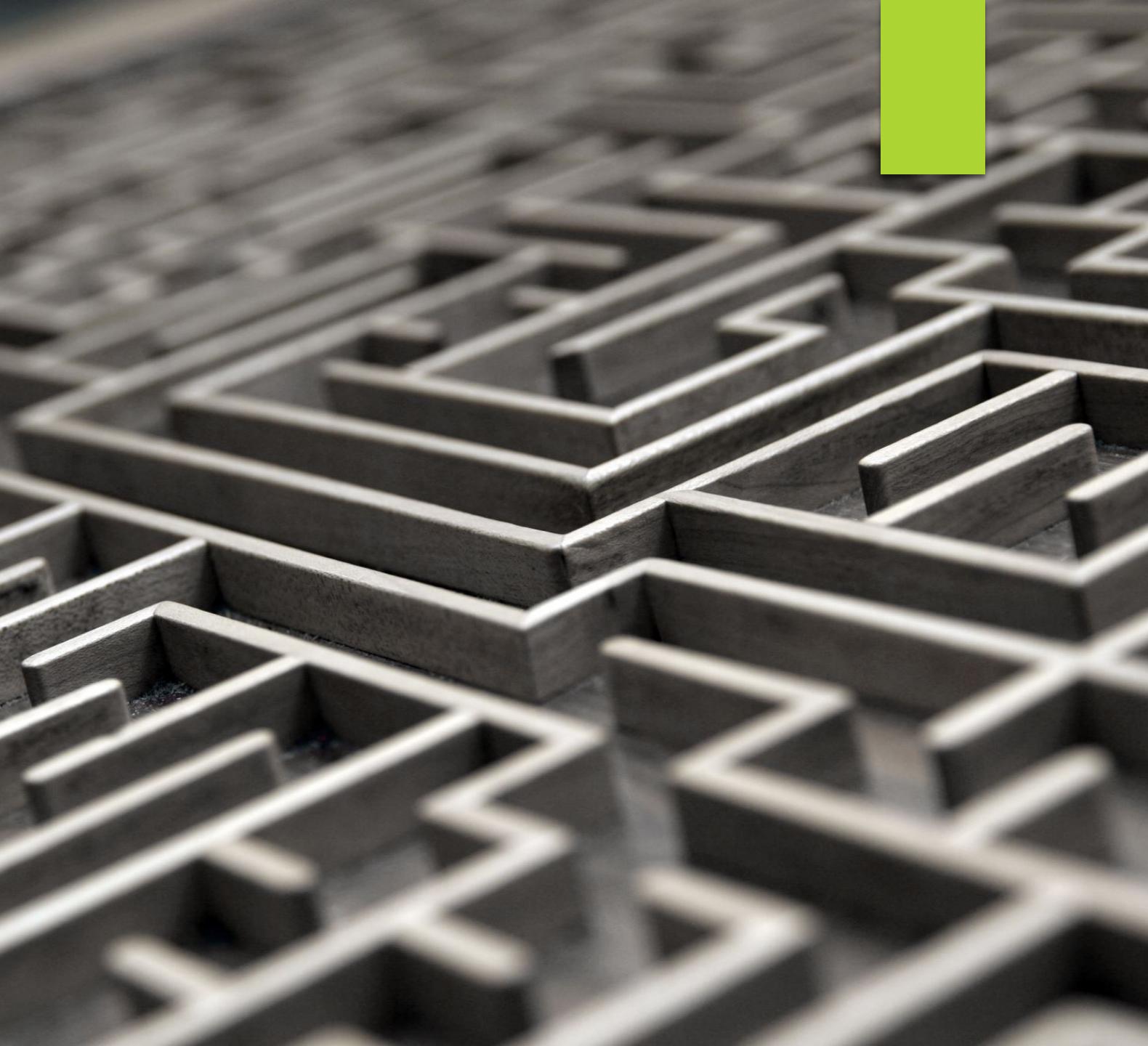


Model	RMSE	R2
Linear Regression	4.14	0.85
Random Forest	3.40	0.90
Gradient Boost	2.82	0.93

Performance metrics and Evaluation

Challenges

- ▶ Skewed data sanitization
- ▶ Selecting players who are part of both train and test data set
- ▶ Selection of right feature sets for batsman and bowler for more accurate prediction was tricky
- ▶ Choosing the most accurate model



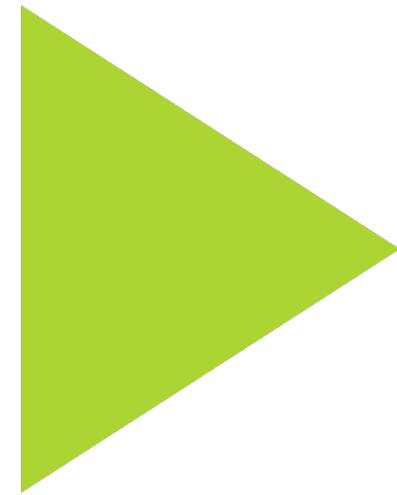


GitHub Link



- ▶ <https://github.com/s0098/IPL-Cricket-Match-Analysis>

Demo



Thank You

