

Predictive Modeling and Team Performance Analysis in IPL



Supervisor: Mr. Vivek Kshirsagar

Name: Rohit Moholkar

Student Id: 10613711

Presentation Includes Following Points:

- Introduction.
- Literature Review.
- Methodology.
- Data Analysis.
- Feature Engineering.
- Machine Learning.
- Graph of Win/Lose Probabilities.
- Results.
- Challenges.
- Future Scope.
- Conclusion.



INTRODUCTION

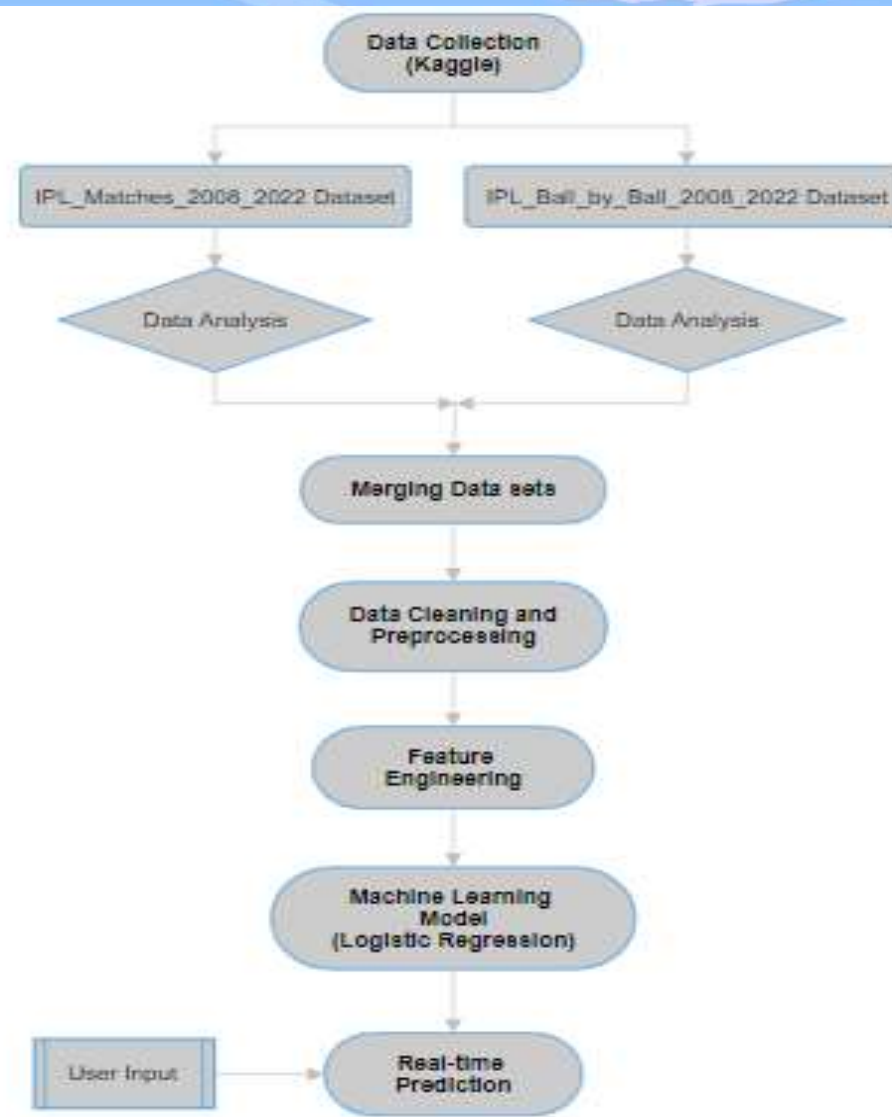
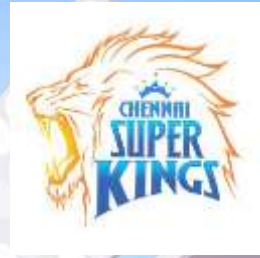
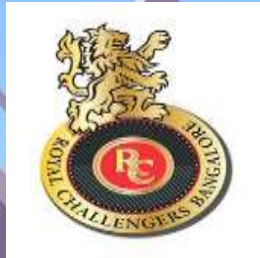
- Cricket's Global Impact.
- T20 Cricket Format.
- IPL's Evolution and Innovation.
- Overview of IPL Franchises and Teams.
- Understanding about Data.
- Data Analytics in Professional Cricket.
- Empowering Stakeholders Through Understanding.
- Machine Learning's Impact on Cricket Strategy.
- Evolutionary Impact on Cricket Culture.
- Enhancing the Cricketing Experience.



LITERATURE REVIEW

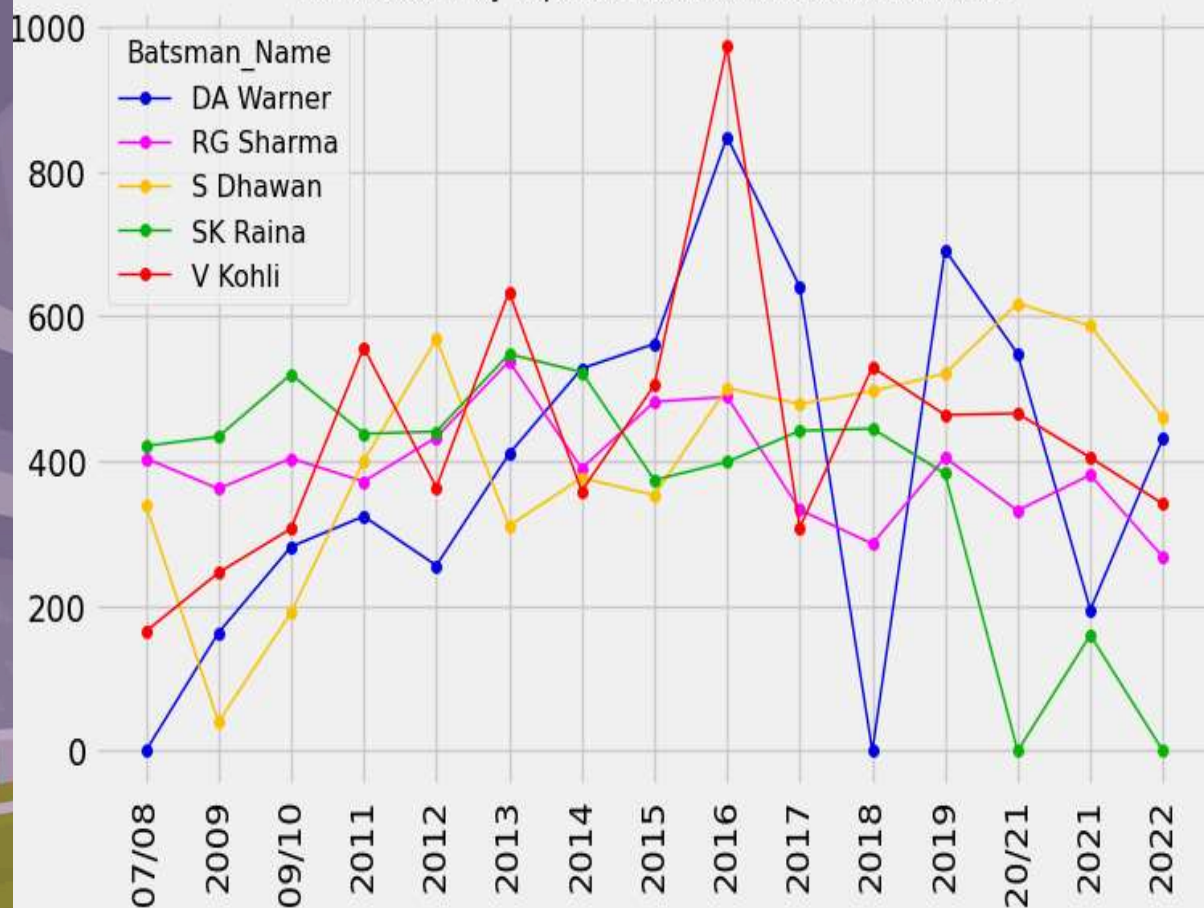
- (Nimmagadda, n.d.) Cricket score and winning prediction using data mining.
- (Petersen, 2008) IPL 2008 season analysis.
- (Vistro, 2019) Using machine learning and data analysis predicting cricket winners.
- (Jaipuria, 2022) Toss and Venue impact on the match result.
- (Sinha, 2020) Enhancing performance of the team with support vector machine in IPL.
- (Thorat, 2021) Review paper on cricket score prediction.
- (Patil, n.d.) Cricket team prediction using machine learning techniques.
- (Kapadia, 2019) A study on match outcome using machine learning.

METHODOLOGY

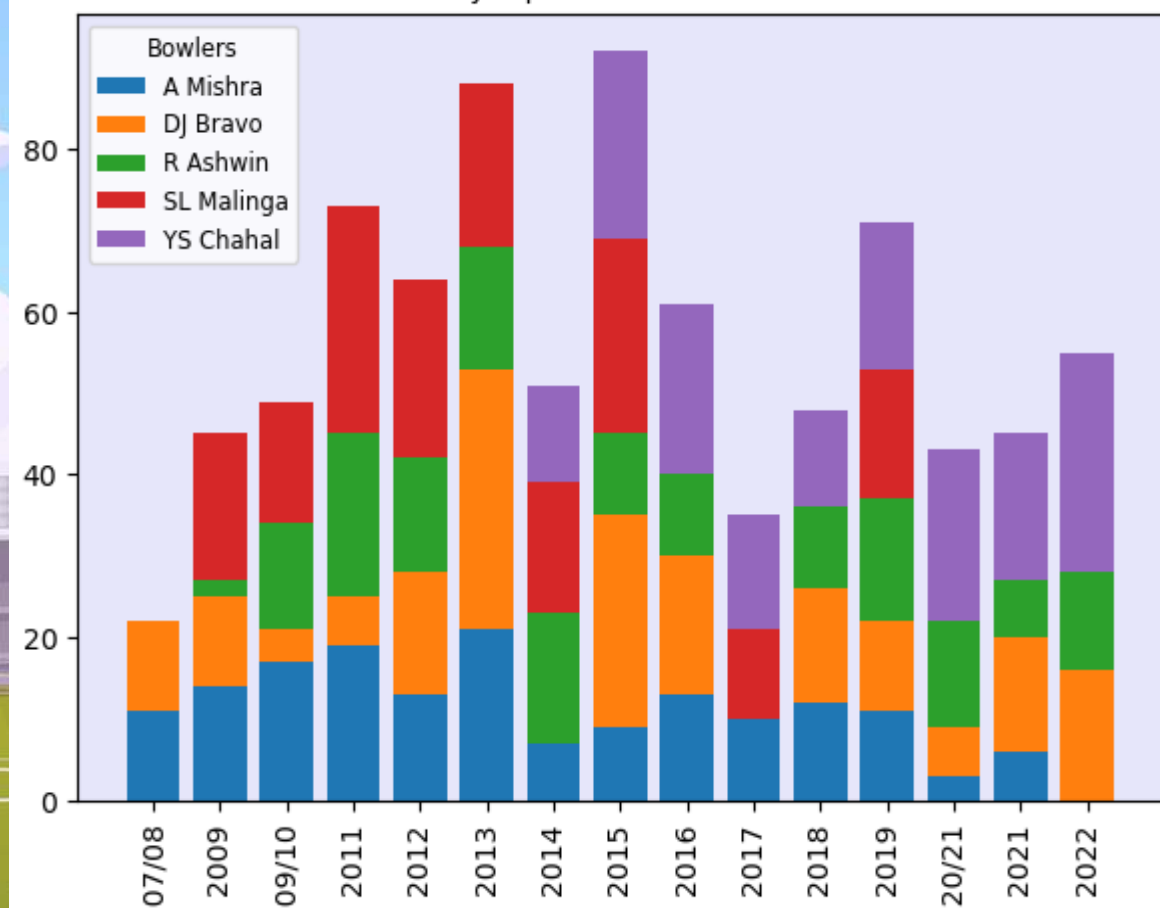


DATA ANALYSIS

Runs Scored by Top Five Batsmen in each IPL Season



Wickets Taken by Top Five Bowlers in each IPL Season



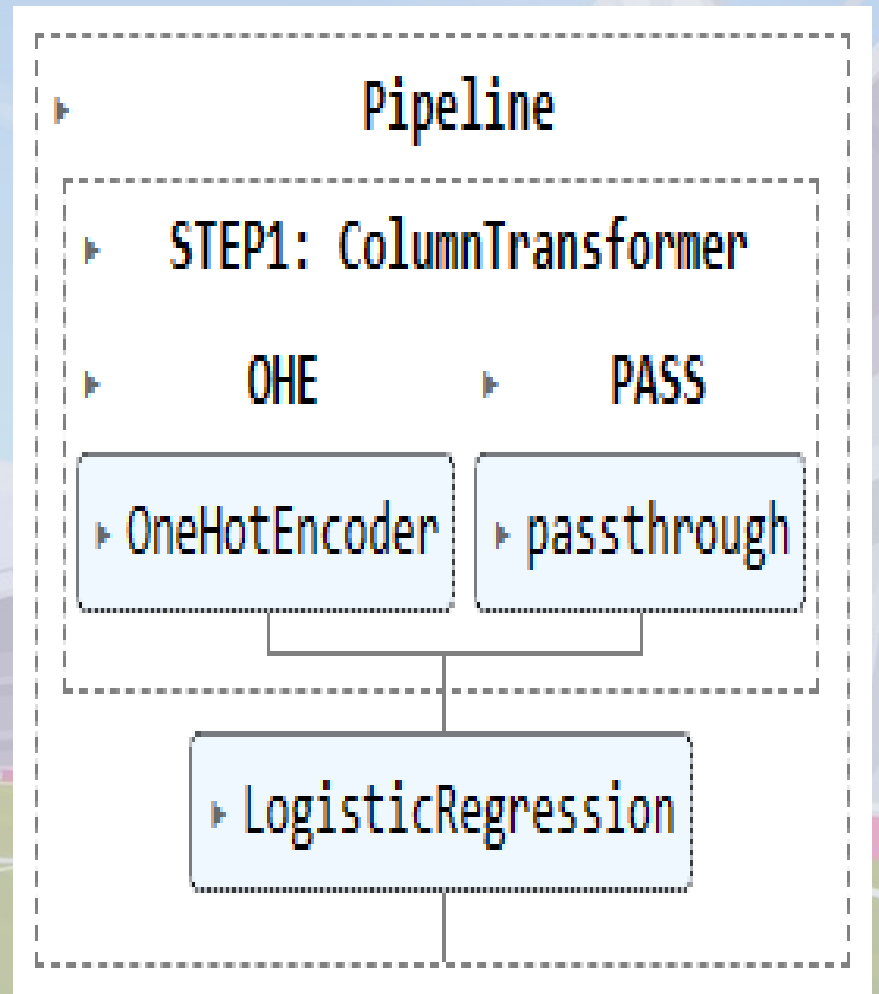
FEATURE ENGINEERING

- Addressing Missing Values.
- Merging Datasets.
- Data Cleaning.
- Ensuring Consistency in Columns.
- Creating Additional Columns.
- Fixing Inaccuracies.
- Generating Dynamic Features.
- Handling Null and Infinite Values.
- Creating Target Variable.
- Removing Unnecessary Columns.

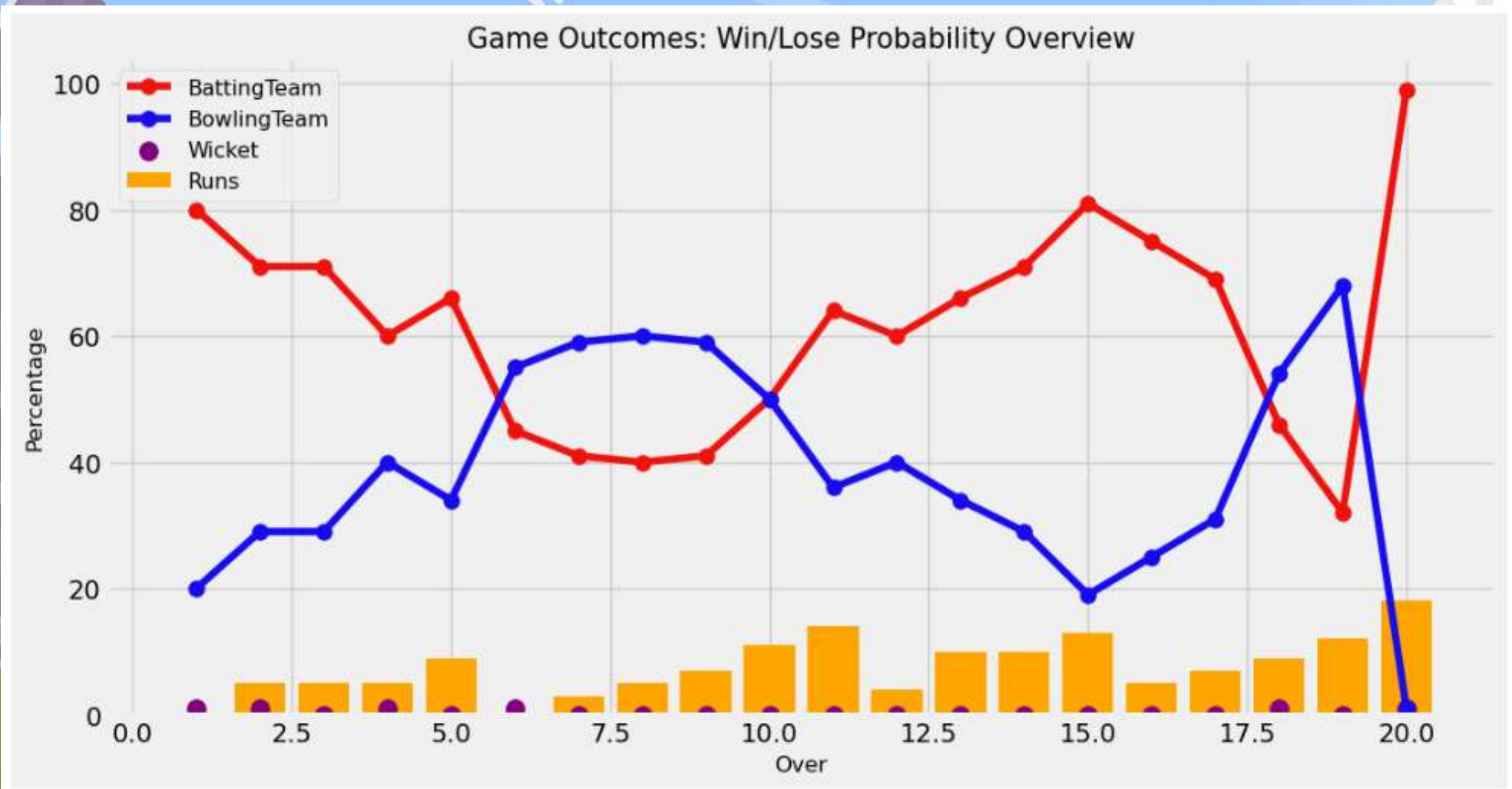


MACHINE LEARNING

- Selecting Features and Target Variable.
- Creating Column Transformer.
- One-Hot Encoding (OHE) for Categorical Column.
- Pipeline Construction.
- STEP1 - Data Transformation.
- STEP2 - Logistic Regression Model.
- Fitting the Pipeline.
- Feature Importance.
- Accuracy Score Calculation.
- Cross-Validation for Accuracy Check.
- Winning Probabilities DataFrame.

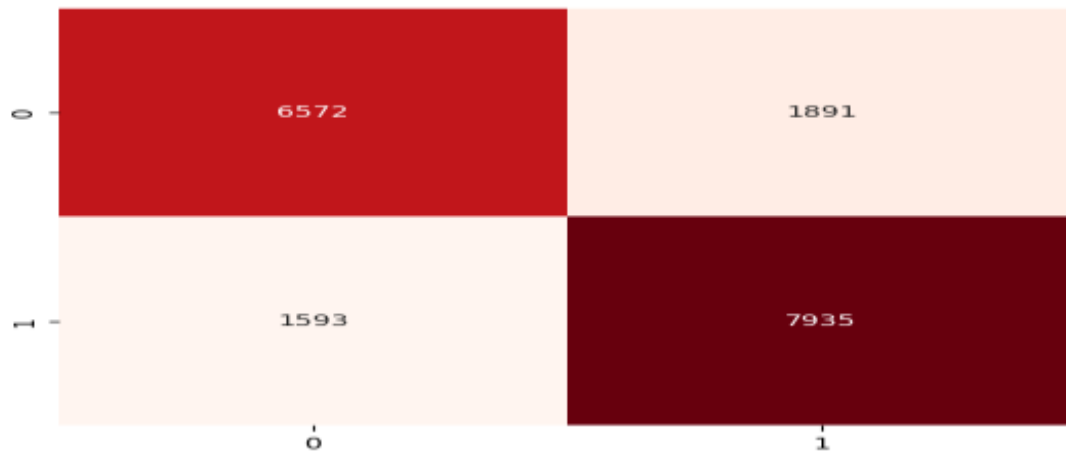


GRAPH OF WIN/LOSE PROBABILITIES



RESULTS

Confusion Matrix



	precision	recall	f1-score	support
0	0.80	0.78	0.79	8463
1	0.81	0.83	0.82	9528
accuracy			0.81	17991
macro avg	0.81	0.80	0.81	17991
weighted avg	0.81	0.81	0.81	17991

Creating function that takes raw data as input and returns dataframe ready for

```
def input_df(bat,bowl,city,target,runs,balls,wickets,crr,rrr):
```

```
    input_data = pd.DataFrame([{\n        'BattingTeam':bat,\n        'BowlingTeam':bowl,\n        'City':city,\n        'Target':target,\n        'runs_left':runs,\n        'balls_left':balls,\n        'wickets_left':wickets,\n        'CRR':crr,\n        'RRR':rrr,\n    }])
```

```
    return input_data
```

```
data = input_df('Royal Challengers Bangalore','Kolkata Knight Riders','Mumbai',180,180,120,10,0,9.00)
```

```
# Printing answer.
```

```
if loaded_model.predict(data)[0]==1:
```

```
    print('High Probability of WIN')
```

```
else:
```

```
    print('High Probability of LOSS')
```

High Probability of LOSS

CHALLENGES

- Diverse Datasets.
- Cricket Context Understanding.
- Exploratory Data Analysis. (EDA)
- Handling and Cleaning of the Raw Data.
- Robust Data Preprocessing.
- Calculation Complexity.
- Bug Fixing and Accuracy.
- Feature Engineering Challenges.
- Ensuring Data Uniformity.
- Model Evaluation.



FUTURE SCOPE

**Data
Integration in
Real Time.**

**Consider
External
Factors.**

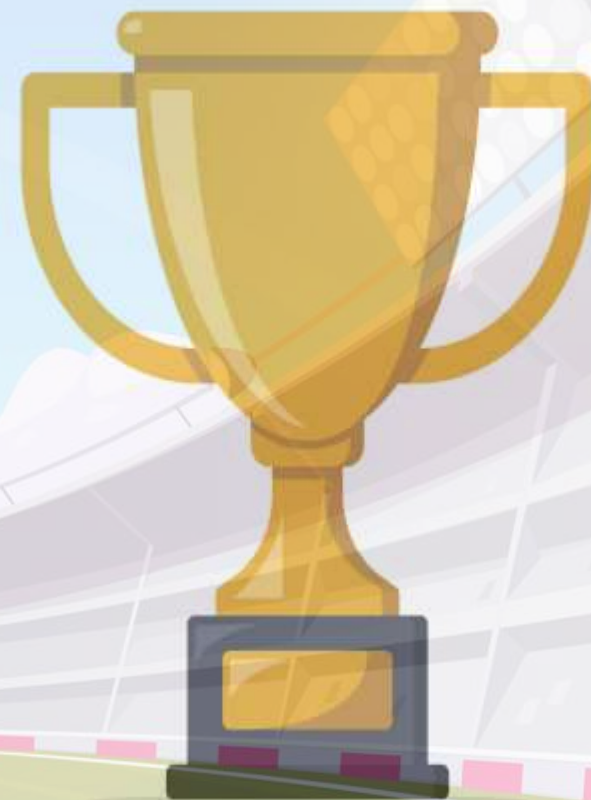
**Create
Community
Involvement.**

**Work with
Experts.**

**Develop
interactive
Application or
Website.**

CONCLUSION

- Comprehensive IPL Journey.
- Essential Analysis Stages.
- Deep Dives into Player Performances.
- Innovative Statistical Functions.
- Fusion of Statistics and Machine Learning.
- Future Possibilities and Innovations.
- Innovation in Real-Time Predictions.
- Sporting Insights Beyond Statistical Exploration.
- Data Science's Impact on Cricket Understanding.





THANK YOU