

SPORTS PREDICTION

DATA WIZARDS

Siddhi Patil

Anushka Kulkarni

Tulsi Kabade

Charul Wankhede

OUR OBJECTIVE

1. Predicting match outcomes (win or loss) to assist in strategic decision-making for teams, betting, and fantasy leagues.
2. Estimating match scores to understand team performance and aid in pre-match analysis.
3. Predicting player performances (runs, wickets, etc.) to identify key players and their impact on the match.
4. Forecasting specific match events (wickets, boundaries) to provide detailed insights into match dynamics.
5. Offering inning-wise predictions to understand match progression and momentum shifts.
6. Providing real-time predictions during live matches to engage fans and enhance their viewing experience.



PROPOSED APPROACH

1. Data Loading and Preprocessing:

- Loaded and preprocessed IPL datasets, handling missing values and encoding categorical variables.

2. Exploratory Data Analysis (EDA):

- Explored IPL trends like home team performance and toss decisions, visualizing insights.

3. Feature Engineering and Model Building:

- Engineered features (e.g., strike rates) and built RandomForest/XGBoost models for match winner prediction.

4. Model Evaluation and Improvement:

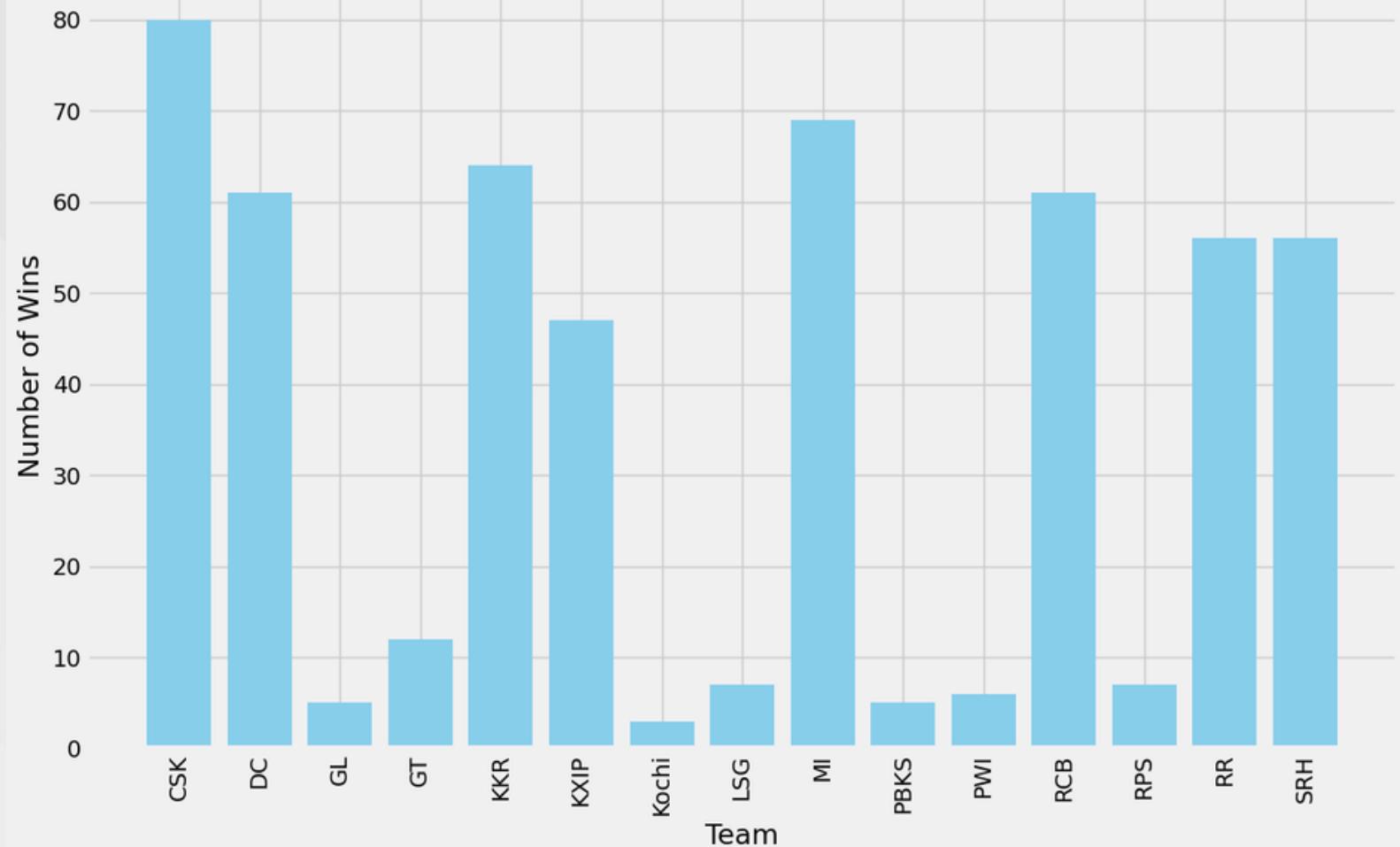
- Evaluated models' accuracy, utilized StandardScaler for feature scaling, and explored RandomForest/XGBoost effectiveness.

5. Prediction Functionality:

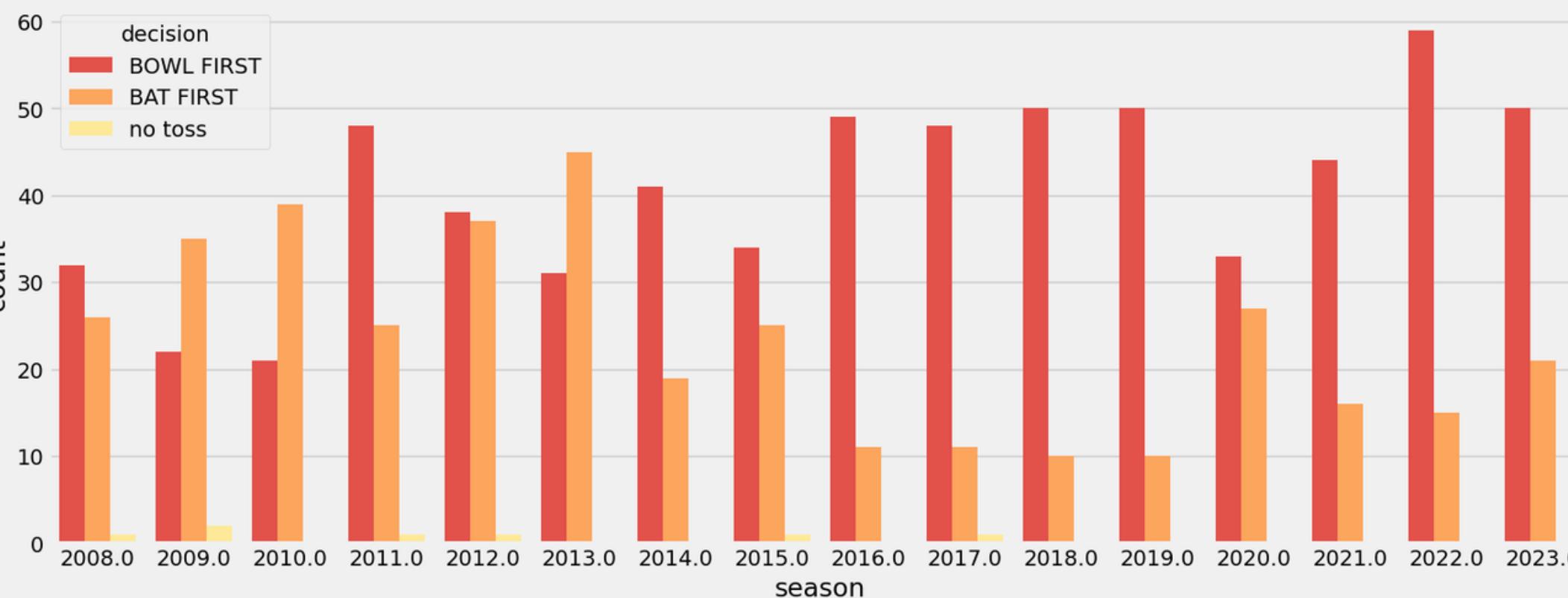
- Developed a function to predict match winners using input features, integrating trained models for forecasts.

EXPLORATORY DATA ANALYSIS

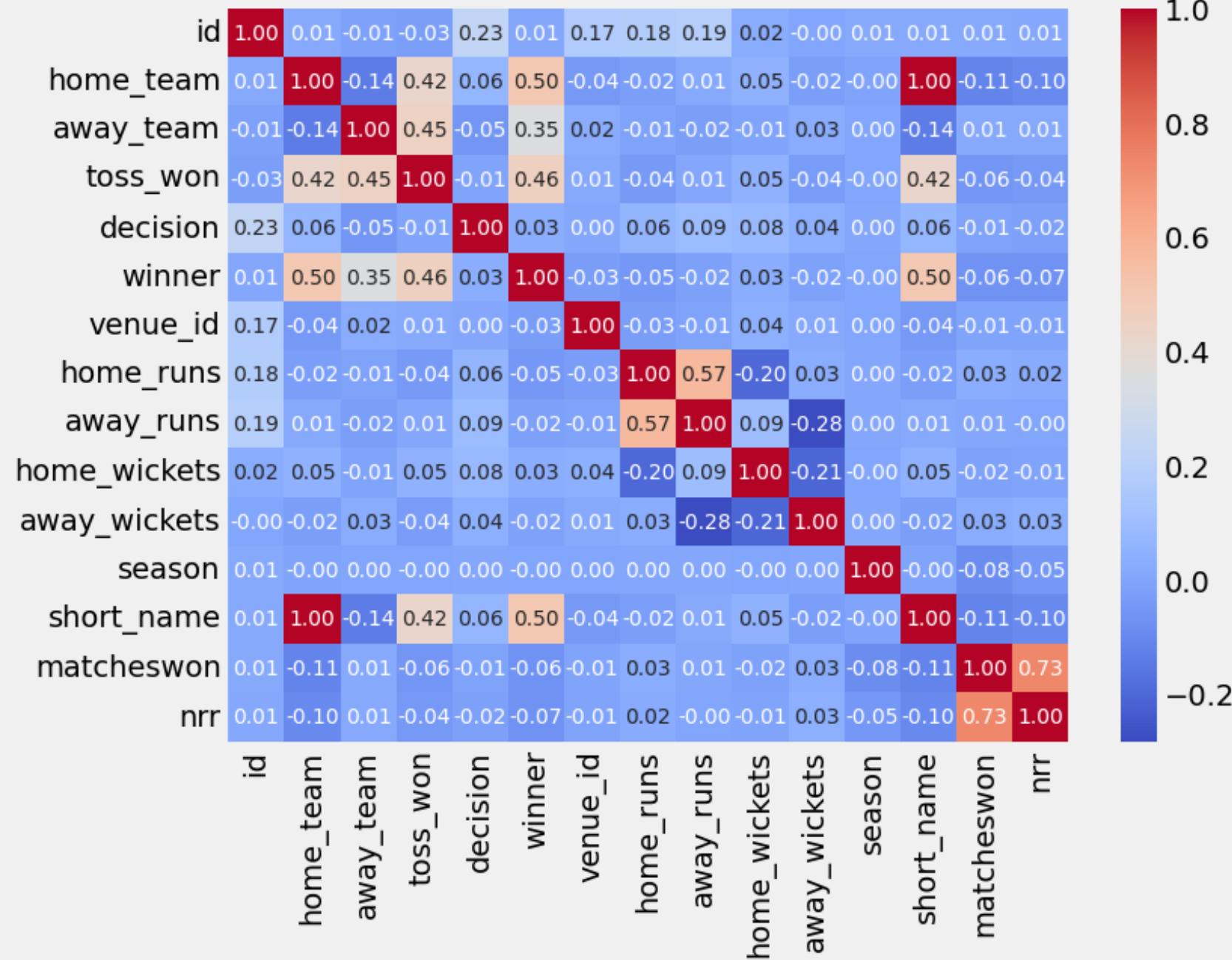
Number of Wins for Each Team at Their Home Venues



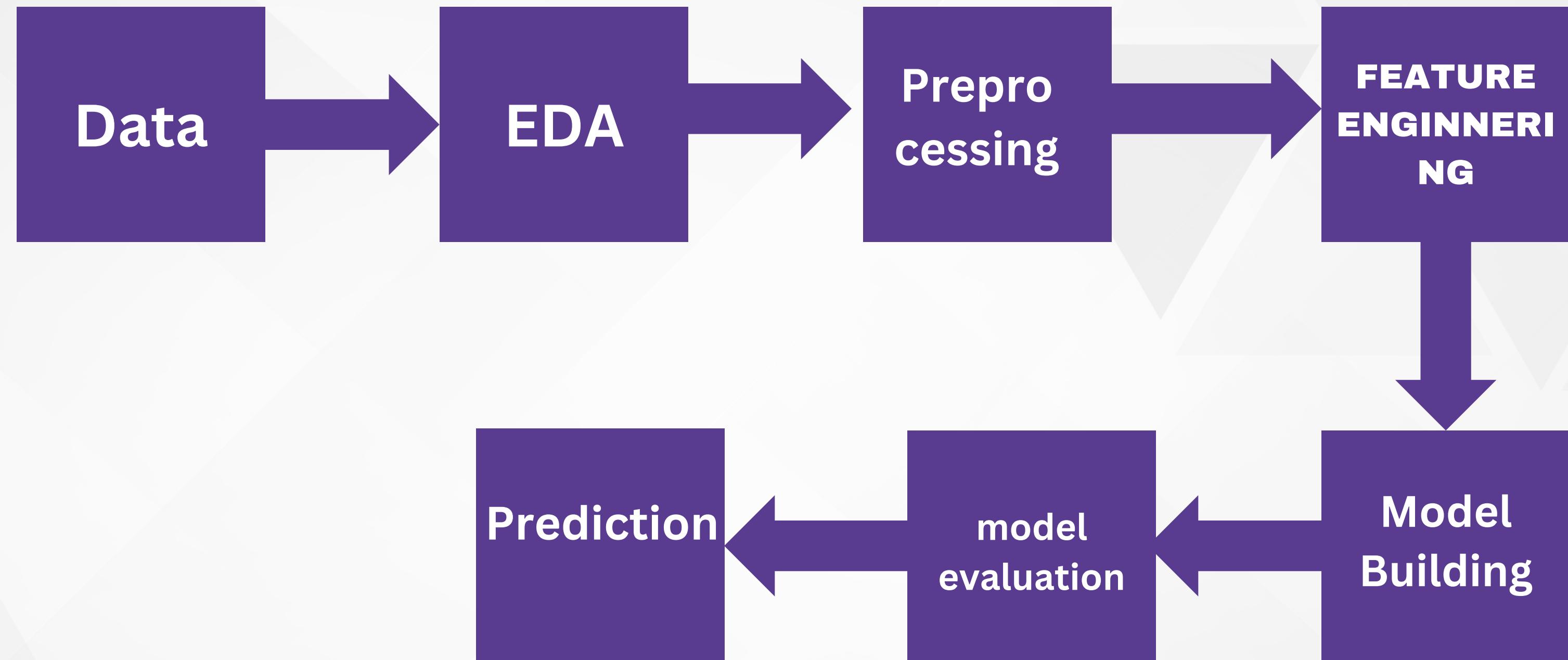
Toss Decision in each Season of IPL



Correlation Matrix



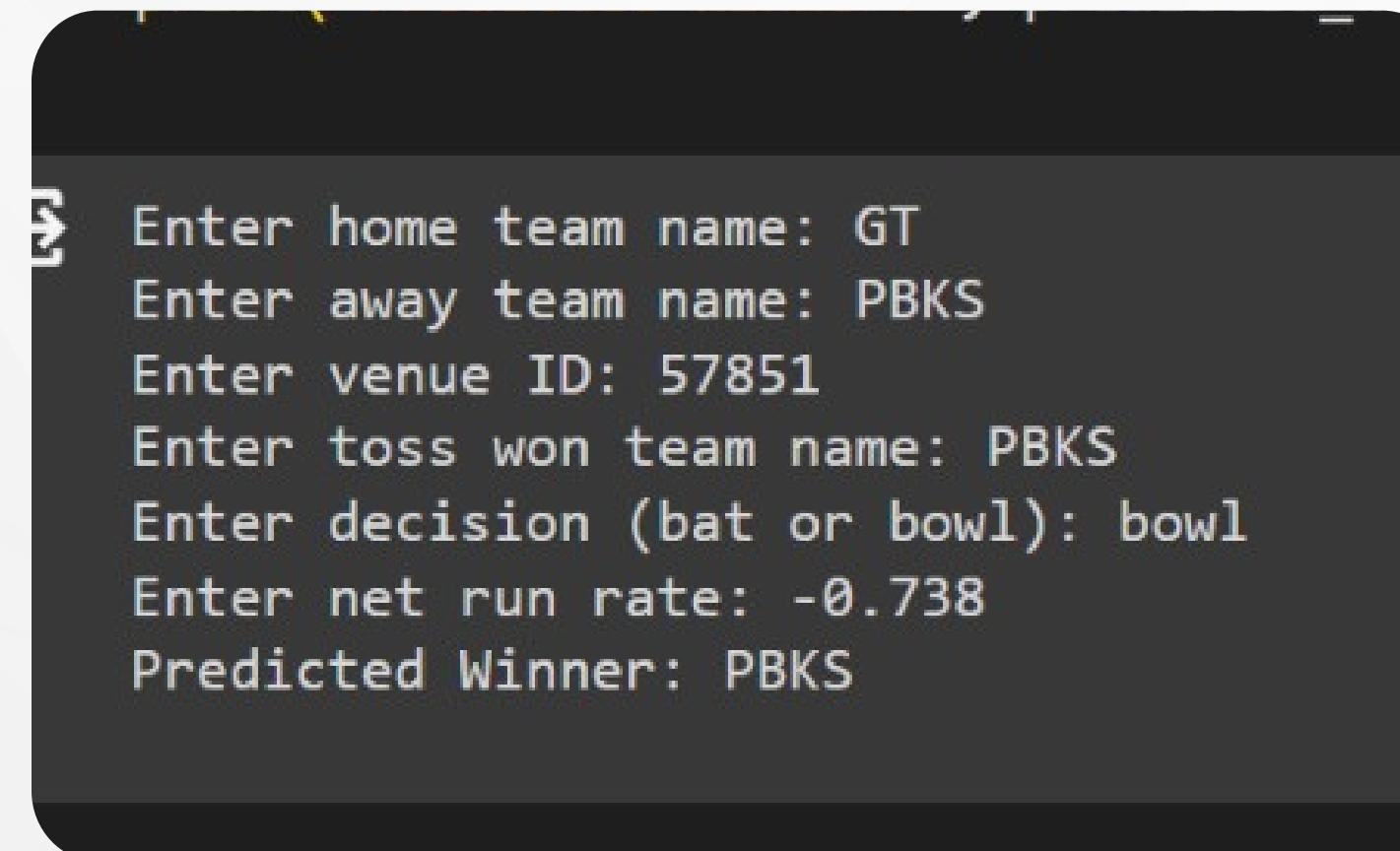
FLOW OF THE SOLUTION



ALGO USED

XGBOOST

Accuracy:
0.8299474605954466



Enter home team name: KKR
Enter away team name: DC
Enter venue ID: 57980
Enter toss won team name: KKR
Enter decision (bat or bowl): bowl
Enter net run rate: 1.528
Predicted Winner: KKR

ALGO USED

RANDOM FOREST

Accuracy:
0.8185639229422067

TECH STACK

1. Programming Languages:

- Python: Widely used for data manipulation, machine learning, and model development.

2. Machine Learning Libraries:

- scikit-learn: For building and training machine learning models.
- XGBoost, LightGBM, CatBoost: Gradient boosting libraries for more advanced model building.
- TensorFlow or PyTorch: Deep learning frameworks for implementing neural networks.

IDE Used- GOOGLE COLAB

FUTURE SCOPE

1. Incorporating more data sources like player fitness, weather conditions, and team strategies.
2. Advancing feature engineering techniques for better data extraction.
3. Implementing ensemble methods and deep learning approaches for improved accuracy.
4. Introducing explainable AI for transparency in predictions.
5. Developing real-time prediction systems and interactive visualizations.
6. Personalizing predictions based on user preferences.
7. Deploying prediction models as services for wider accessibility.
8. Collaborating with IPL teams, analysts, and betting platforms.
9. Continuously refining models with new data for better performance.





THANK YOU
