# Sports Prediction

## Problem Statement 2

Team Name: Code Blooded

# TABLE OF CONTENTS

## OBJECTIVE

Develop a model which will predict the winner of a match

Provided Data set:
https://www.kaggle.com/datasets/rajsengo/indian-premier-league-ipl-all-seasons

all_season_batting_card.csv
all_season_bowling_card.csv
all_season_details.csv
all_season_summary.csv
points_table.csv

## FUTURE SCOPE

- Real Time Analysis
- Player Training Modules
- Team training
- Team selection
- Wearable devices
- Education and Research
- Multimodal Data Fusion
- Personalized Predictions
- Adaptation to Dynamic Environments
- Global Adoption and Accessibility

# Proposed Approach

# Approach

1. Exploratory Data Analysis

2. Data Preprocessing

3. Model training And Prediction

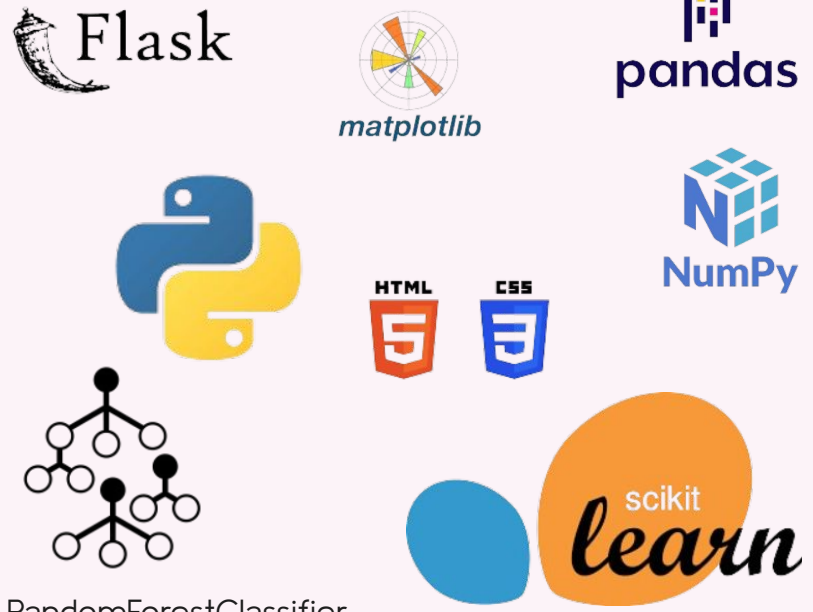
4. Model Deployment

# EDA and Tech Stack

## EDA

- Data set contained data from 2008 to 2023
- Using data from 2018 to 2023
- Data Types: Float, int, object
- Found the unique values of every column from Summary table
- Identifying dependent and independent columns

## Tech Stack



# Data Preprocessing

- Dropped the null values
- Used `StandardScaler` for runs
- Calculated the average strike rates and economy rates for both teams in every match using batting and bowling cards
- Calculated the average strike rates and economy rates of current 2024 teams on the basis of past 5 years dataset and current players
- Mapped the textual data to numerical values: team names, venue, toss won, toss decision
- Created a merged data set for the model
- Split the data in a 80:20 ratio

# Model Training and Prediction

## Architecture and Training

- Model chosen is RandomForestClassifier, accuracy: 49.35%
- RandomForestClassifier gave the highest accuracy out of Logistic Regression, SVM, and XGBoost for the given dataset
- Trained on 9 input features, mainly: home_team, away_team, toss_won, decision, venue_name, away_avg_strike_rate, home_avg_strike_rate, home_avg_economy_rate, away_avg_economy_rate
- Target variable: Winner

## Predictions

- Input from user: home_team, Away_team, toss_won, toss_decision, venue_name
- Example 1: Match on 3rd April 2024 KKR vs DC
    - Predicted: KKR, Actual: KKR
- Example 2: Match on 2nd April 2024 LSG vs RCB
    - Predicted: LSG, Actual: LSG
- Example 3: Match on 22nd March 2024 CSK vs RCB
    - Predicted: CSK, Actual: CSK

# Flow Of The Model

# Flow Of the Model