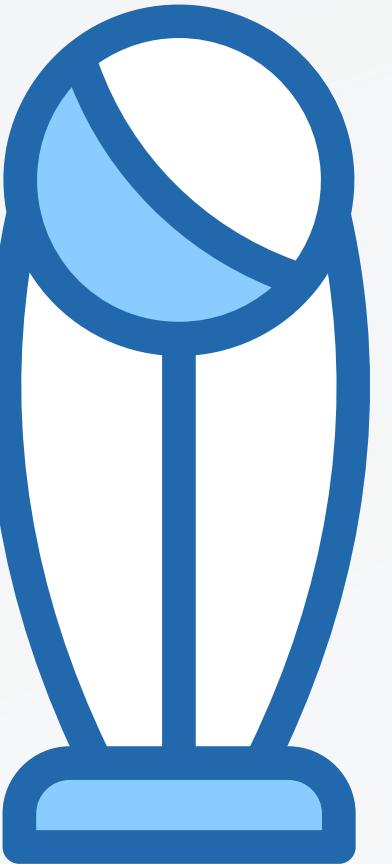
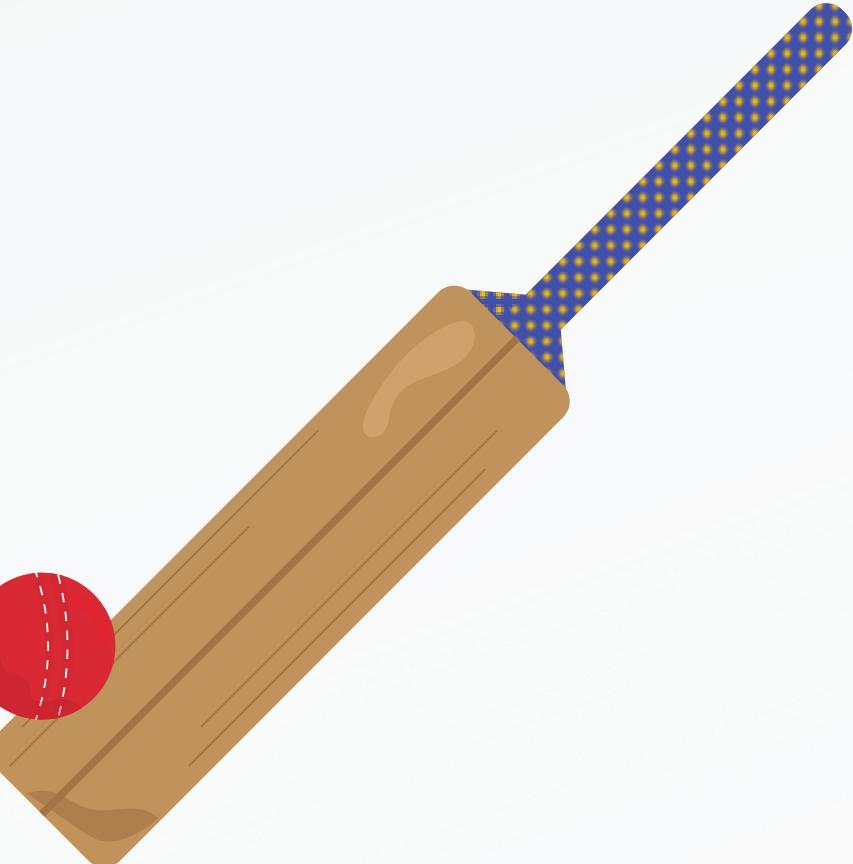


IPL SCORE PREDICTION





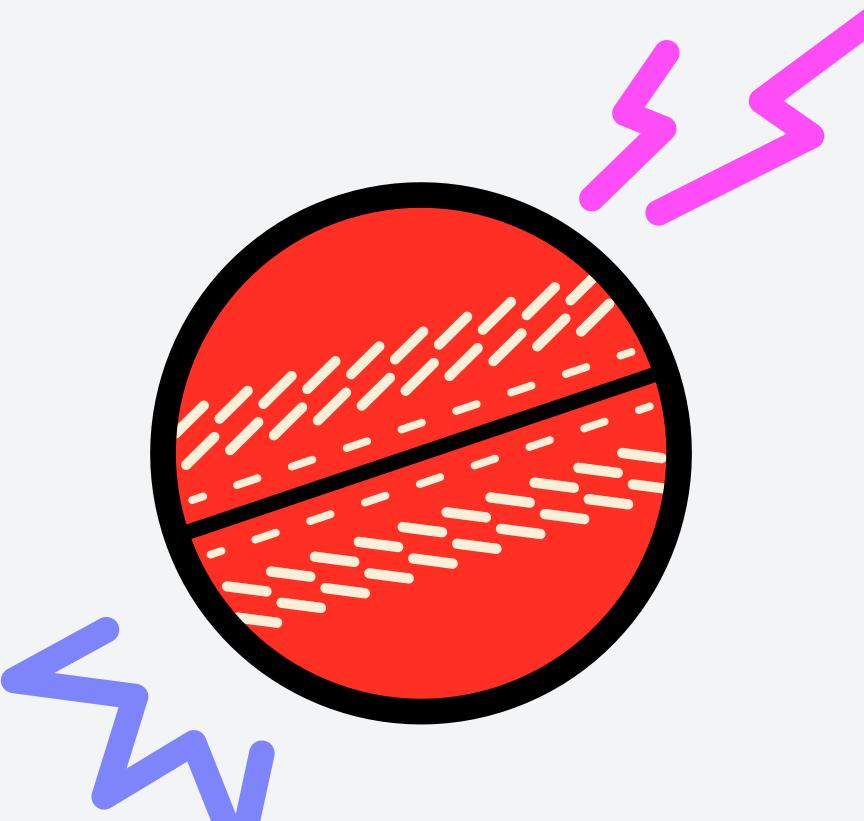
INTRODUCTION

The Indian Premier League (IPL) is a fast-paced cricket tournament brimming with excitement and strategic complexities. Predicting its outcomes presents a captivating challenge due to the dynamic nature of the game and the influence of various factors. This project delves into the world of IPL score prediction, leveraging data analysis and machine learning techniques to unveil the underlying patterns hidden within player performance, team compositions, and historical trends. By harnessing the power of data, we aim to develop a model that can forecast match scores with greater accuracy, providing valuable insights for cricket enthusiasts and potentially even aiding strategic decision-making within the IPL ecosystem.



Tools and Technologies:

- Python for data analysis and model building
- Streamlit web app enables users to predict total runs between teams using current runs and wickets.



FEATURES OF THE DATASET



mid - match id

date - when matches are played

venue - place where matches are played

bat_team - batting team

bowl_team - bowling team

batsman - batsman

bowler - bowler

runs - runs scored

wickets - wickets

overs - next 3 are based on this

run_last_5 - runs scored in last 5 overs

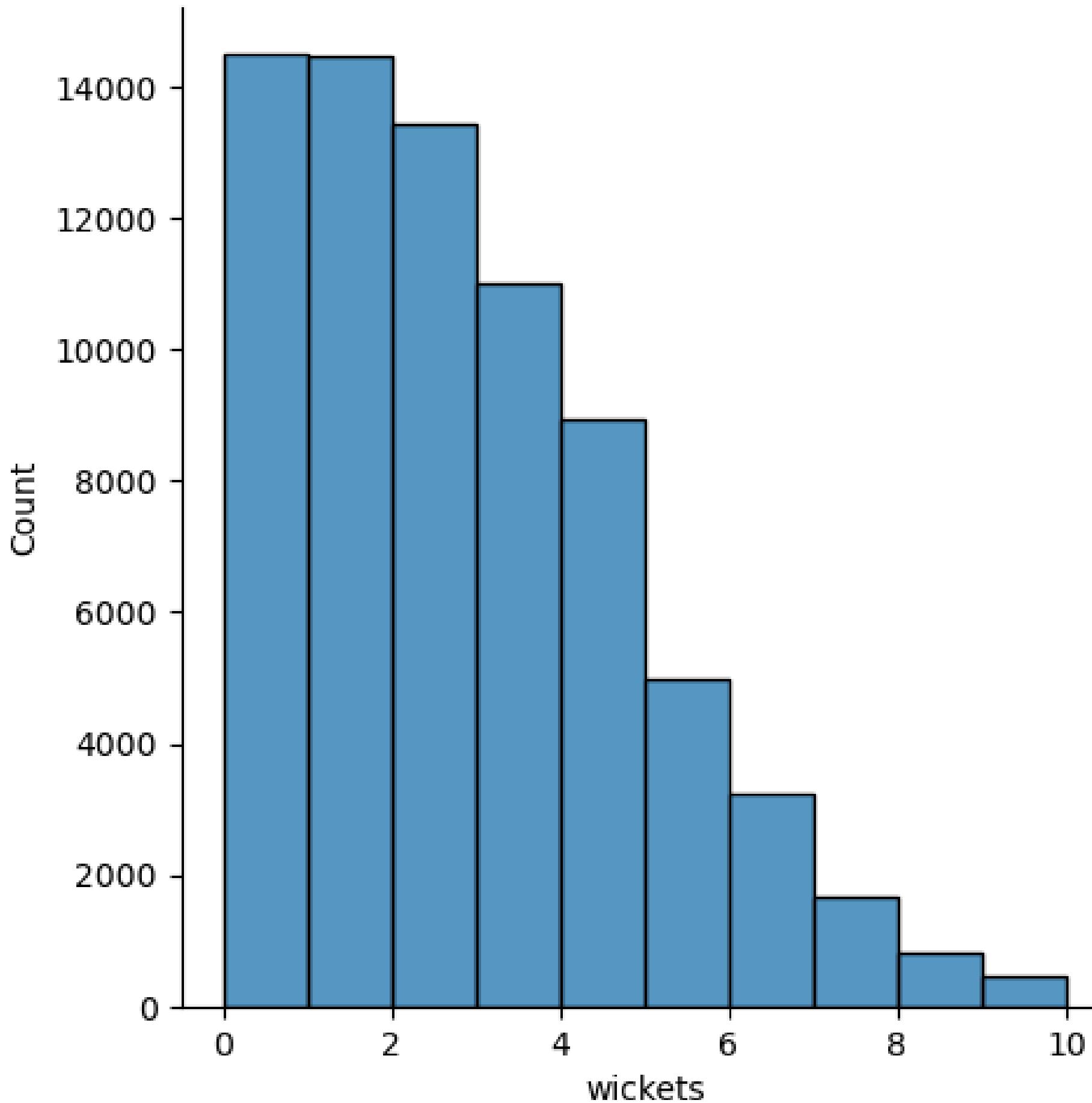
wicket_last_5 - wickets in last 5 overs

striker - batsman playing as main 1

non-striker - batsman playing as runner up - not main 0

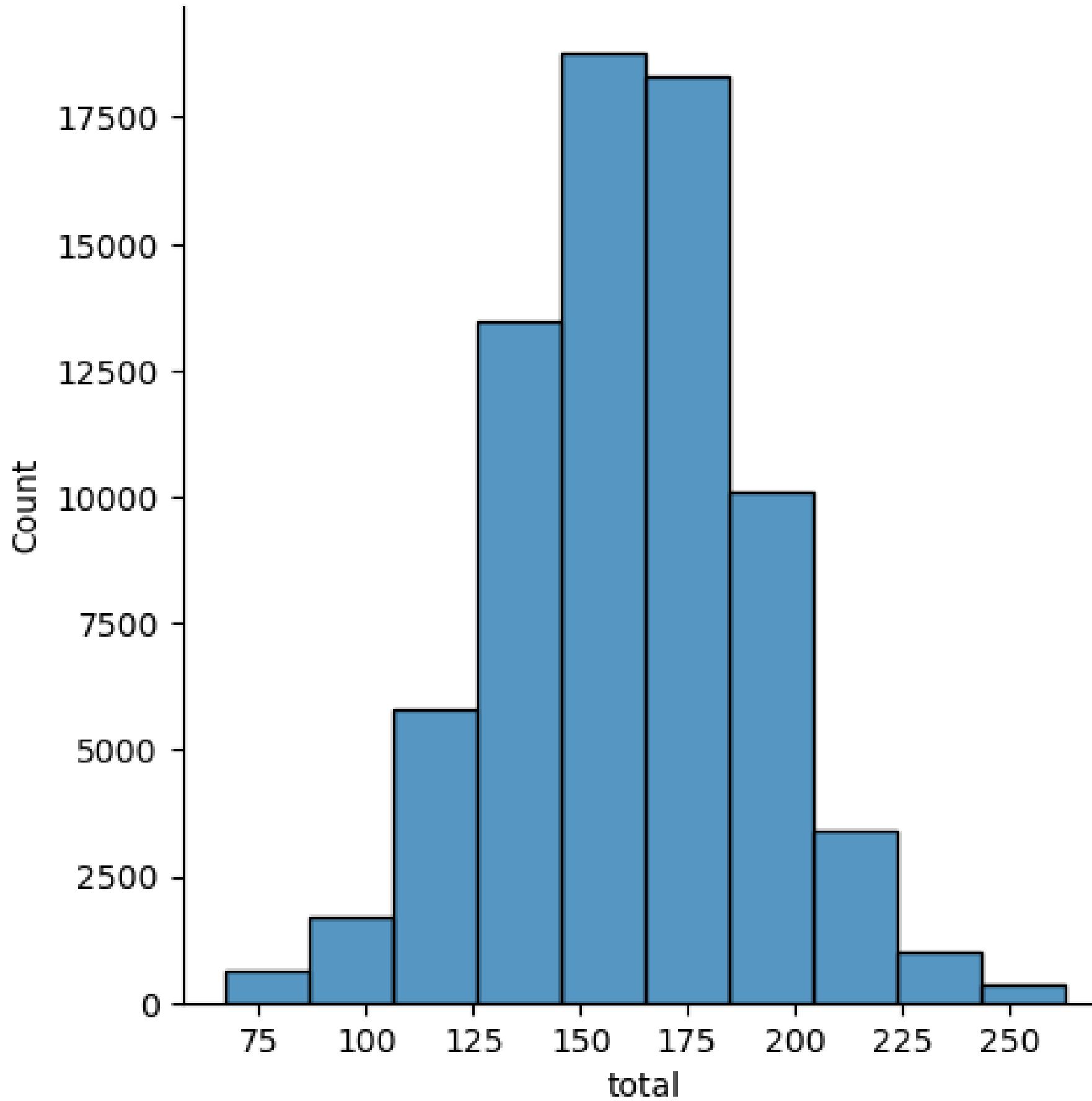
total - total score (target variable)

Wickets Distribution



From the graph we can observe the count of runs per wickets. Maximum of the wickets were fallen in the initial stages of the match.

Runs Distribution



From the above graph we can observe that the graph is similar to normal distribution.
Maximum teams scored 140 to 180 runs.

ALGORITHMS USED



Linear Regression : Train Score : 65.25%
Test Score : 65.19%

SVM: Train Score : 56.70%
Test Score : 56.24%

K-Nearest Neighbor Regression : Train Score : 86.60%
Test Score : 77.13%

XGBoost Regressor: Train Score : 88.66%
Test Score : 85.04%

Random Forest Regressor: Train Score : 99.08%
Test Score : 93.46%

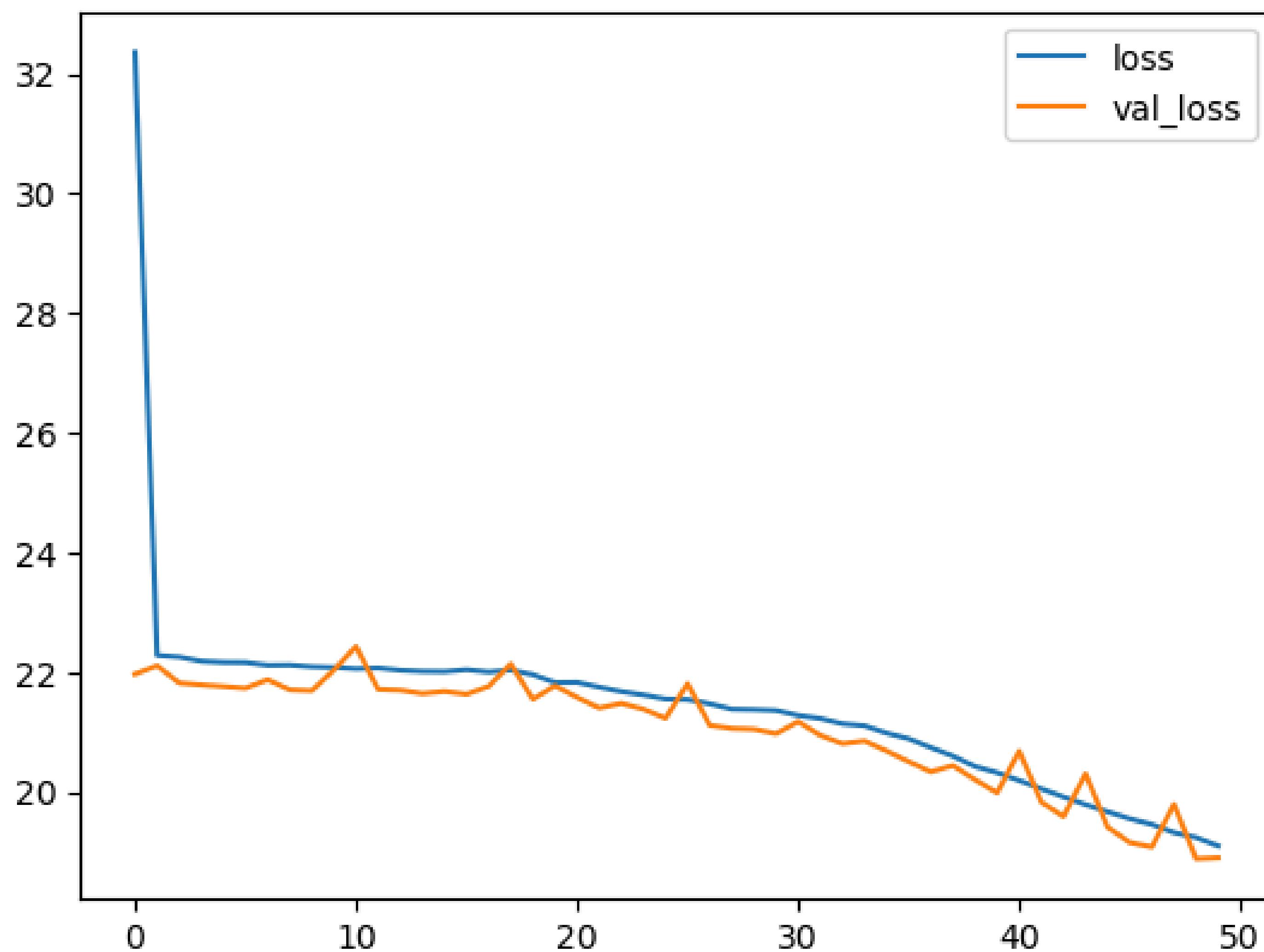
Decision Tree Regressor: Train Score : 99.98%
Test Score : 86.62%

Random Forest performed the best, closely followed by Decision Tree. So we will be choosing Random Forest for the final model.



DEEP LEARNING

Since the dawn of the IPL in 2008, it has attracted viewers all around the globe. A high level of uncertainty and last-minute nail-biters have urged fans to watch the matches. Within a short period, the IPL has become the highest revenue-generating league in cricket. In a cricket match, we often see the scoreline showing the probability of the team winning based on the current match situation. This prediction is usually done with the help of data analytics. Before, when there were no advancements in machine learning, predictions were usually based on intuition or some basic algorithms.



After the training we have stored the training and validation loss values to our neuron network during the training process and plotted a graph using the stored values. From the graph we can observe x-axis represents no of epochs, y-axis represents the loss value. Here the loss gets decreasing in both training and validation loss which means the model is learning from the training data and generalizing well to unseen data in validation set predicted using the trained neural network on the testing data.

THANK YOU

