IPL SCORE PREDICTOR

KPR Institute of Engineering and Technology

Kishore Balaji B | Suryaprakash S | KPRIET

Motivation/ Introduction

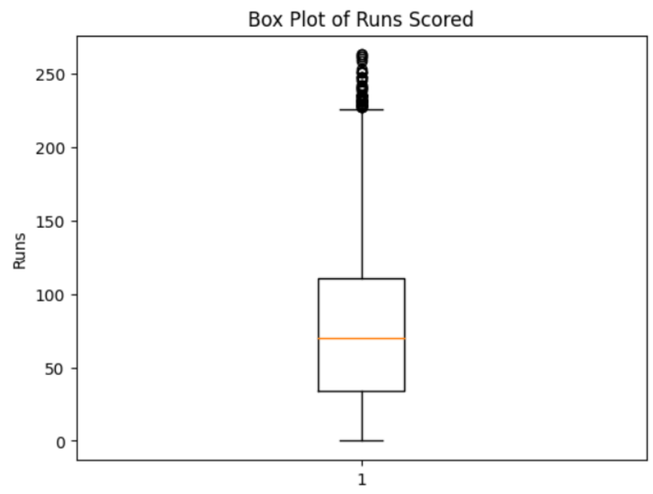
The Indian Premier League (IPL) is one of the most popular cricket leagues in the world. In this project, we will use machine learning to build a model that predicts the total runs scored in an IPL match. We will use a dataset of historical IPL matches to train our model. The model will be able to learn the relationships between different factors, such as the batting team and bowling team, current team score, current wickets, overs bowled, run scored in last 5 overs and wickets gone in last 5 overs, and use this knowledge to predict the total runs scored.

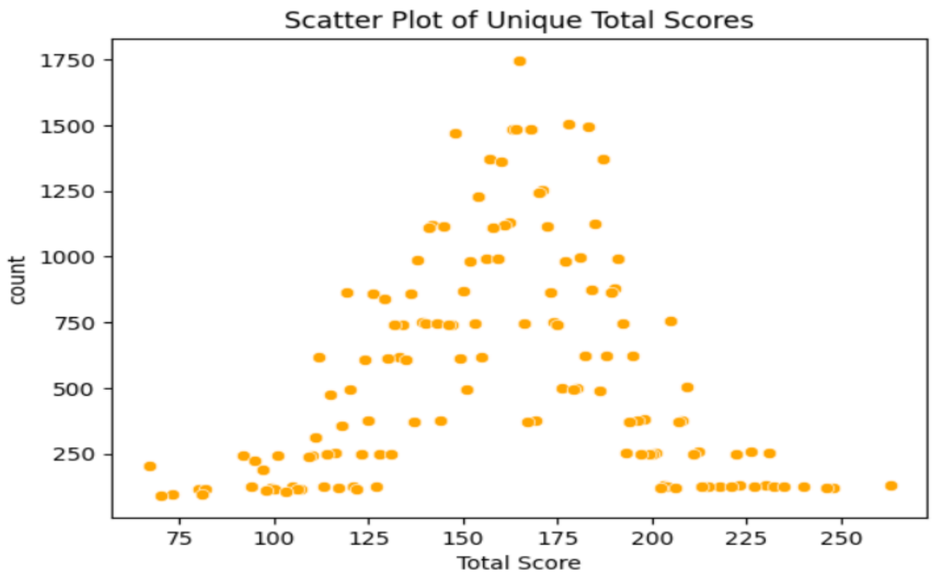
Scope of the Project

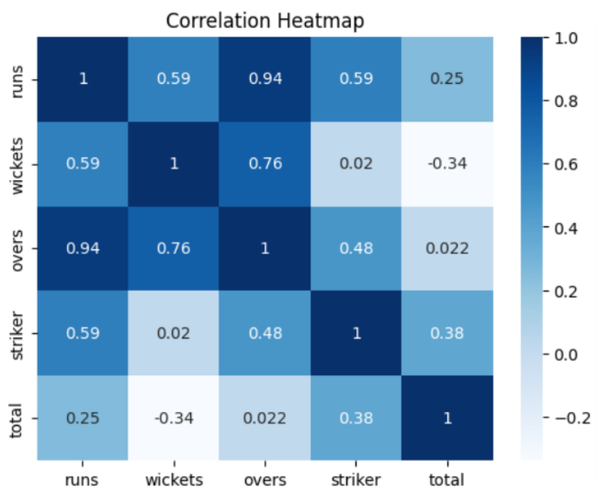
• We have achieved 93.49 % accuracy on the IPL 2008-2017 DataSet.csv dataset on which earlier researches have proposed a maximum of average accuracy of 93.77 % on Random Forest Regression, 86.83% on Decision Tree Regressor and 66.63% on Linear Regression.
• We believe that our model solves the IPL teams to assess their strengths and weaknesses and plan their strategies accordingly.
• Cricket fans to predict the outcome of matches and make informed betting decisions.
• Broadcasters to provide more engaging and informative coverage of IPL matches.
• Data scientists to learn more about the factors that influence the outcome of cricket matches.

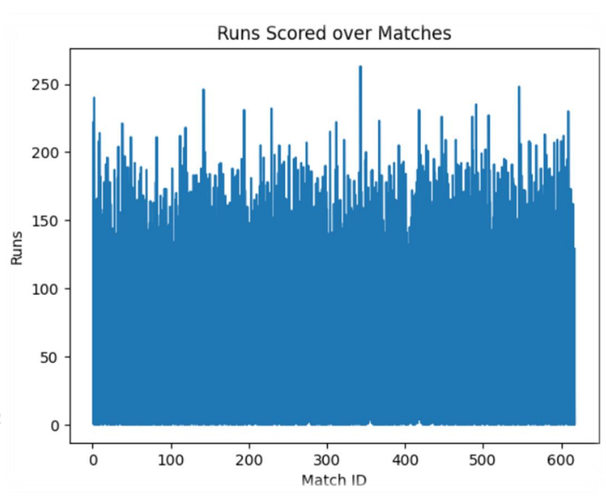
Visualization

S.NO	MODEL USED	TRAIN ACCURACY	TEST ACCURACY
1	RANDOM FOREST REGRESSION	99.01	93.49
2	DECISION TREE REGRESSOR	99.98	86.52
3	LINEAR REGRESSION	65.98	65.49



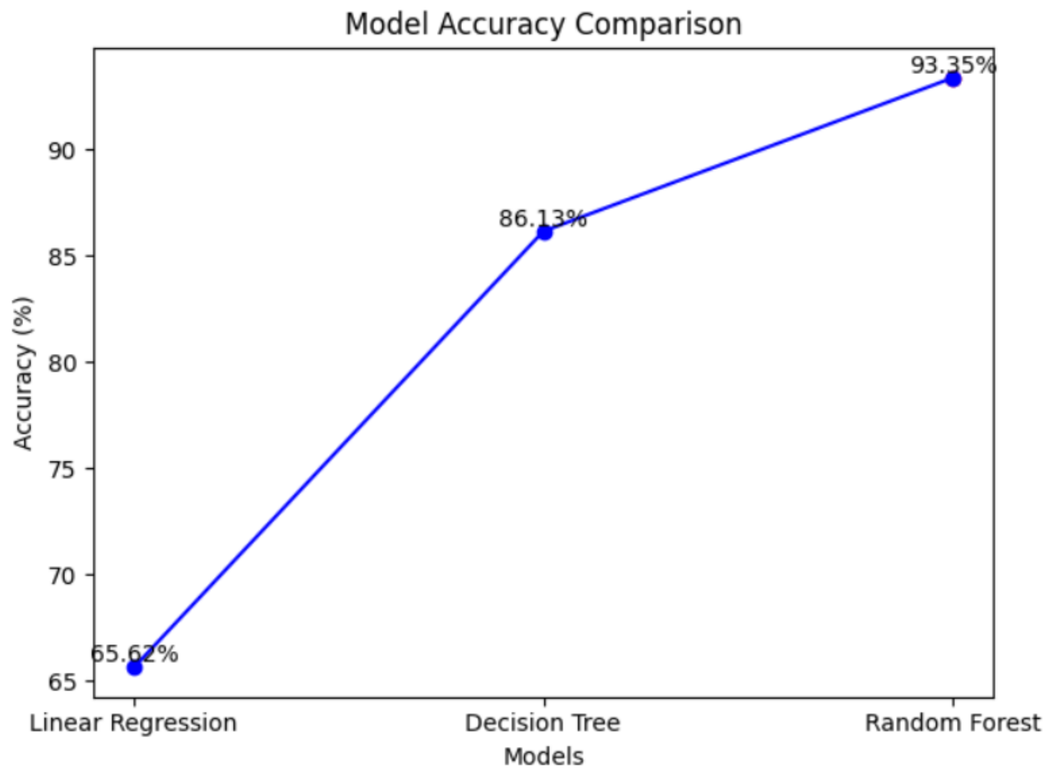


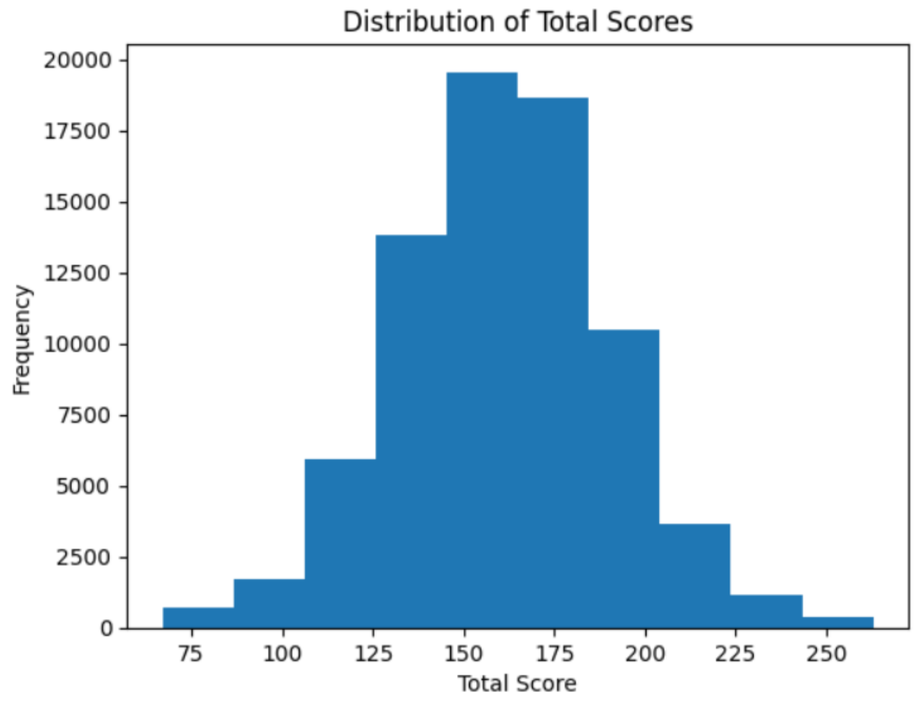




Results

- The collected data is processed and transformed into meaningful features. This may involve data preprocessing, such as handling missing values, scaling, and encoding categorical variables. Additional features may be derived or selected based on their significance in predicting IPL scores.
- The preprocessed data is split into training and testing sets. Different regression models, such as the Random Forest Regressor, Decision Tree Regressor, and Linear Regression, are trained using the training data. The models learn patterns and relationships between the features and target variable (IPL scores) during this phase.
- The trained models are evaluated using the testing data. Evaluation metrics such as accuracy is calculated to assess the performance of each model. The accuracy scores provide an estimate of how well the models predict IPL scores.
- Once the models are trained and evaluated, they can be used to make predictions on new or unseen data. The IPL Score Predictor takes the input data, applies the trained model (in this case, the Random Forest Regressor), and generates predictions of IPL scores based on the provided features.





Conclusion/ Summary

The IPL Score Predictor project yielded positive results with accurate predictions of IPL scores. The Random Forest Regressor model stood out with an impressive accuracy score of 93.35%, showcasing its superior performance compared to the Decision Tree Regressor (86.13%) and Linear Regression (65.62%) models. These findings highlight the effectiveness of the Random Forest Regressor in accurately forecasting IPL scores. The project's success suggests the potential application of machine learning techniques in predicting sports outcomes. Further evaluation and analysis could enhance the project's robustness and facilitate the development of more accurate IPL score prediction models.

Contact Details

22am031@kpriet.ac.in
22am062@kpriet.ac.in

Acknowledgments/ References

• https://github.com/topics/ipl-prediction
• https://www.analyticsvidhya.com/blog/2021/10/building-an-ipl-score-predictor-end-to-end-ml-project/
• https://youtu.be/gdI-0owdYvk