AI-DS MINI PROJECT

TOPIC : IPL MATCH ANALYSIS USING MACHINE LEARNING

PROFESSOR : DR. RAVITA MISHRA

STUDENT : AMAN YADAV
ROLL NO : 60

INTRODUCTION

- The Indian Premier League (IPL) is one of the most popular T20 cricket leagues globally.
- Analyzing IPL match data can reveal performance patterns, predict match outcomes, and provide strategic insights.
- This project applies data science and machine learning techniques to analyze and model IPL data.

DATASET USED

- Source: Kaggle IPL datasets including match and delivery data.
- Data includes details like teams, scores, players, match winners, etc.
- Total Matches: Over 600 (from 2008 to 2020).
- Columns: Match ID, Team names, Toss winner, Toss decision, Match winner, Player of the match, etc.

LITERATURE SURVEY

Title	Authors	Year	Techniques Used	Key Findings	Drawbacks
Predicting IPL Match Results Using ML Algorithms	Pratiksha Bhagat, Rucha Mahadik	2021	LR, DT, KNN, RF	RF ~74% accuracy. Toss & venue matter.	Old dataset, limited features
Cricket Outcome Prediction Using ML	Meenal Goyal et al.	2021	KNN, SVM, XGBoost	XGBoost ~78% accuracy. Good feature importance.	No real-time/live data
IPL Outcome Prediction Using Deep Learning	Akash Rathi et al.	2023	LSTM (RNN)	Improved live predictions using innings data.	High complexity, slow training

AI MODELS USED

- • Logistic Regression
- • Decision Tree
- • Random Forest
- • Gradient Boosting
- • Support Vector Machine (SVM)
- • K-Nearest Neighbors (KNN)

MODEL EVALUATION

- Evaluation based on Confusion Matrix, Accuracy, Precision, and Recall.
- Gradient Boosting and Random Forest performed well with high accuracy.
- SVM and Logistic Regression showed balanced results.
- KNN showed variability depending on feature scaling and data split.

CONCLUSION

- Machine learning models can predict IPL match outcomes with significant accuracy.
- Ensemble models like Gradient Boosting and Random Forest are effective.
- Model performance depends on feature selection, data preprocessing, and model tuning.
- Future work: Incorporate live data for real-time prediction and betting insights.