

# Analysis of T20 Cricket Player Performance

## Analysis of T20 Cricket Player Performance

### Final Project Report

DATA 602 – Spring 2024

Intro to Data Analysis and Machine Learning

University of Maryland Baltimore County

### Group 3

Hyma Roshini Gompaa (Student ID: CS35058)

Adarsh Rao Akula (Student ID: Jq00578)

Sujeeth sithagari (Student ID: EG21094)

Saivarun Kotha(Student ID: Fw61697)

Shri Aiswarya Gorle (Student ID: RU45776)

## Analysis of T20 Cricket Player Performance

### 1. Introduction

T20 cricket has rapidly gained popularity worldwide, attracting large audiences with its intense and exciting matches. This project aims to deeply analyze the performances of T20 cricketers, uncovering patterns, extracting insights, and exploring predictive modeling to support team strategies and management. In the fast-paced and ever-evolving landscape of T20 cricket, where matches are decided within a few hours and every run or wicket can sway the outcome, leveraging data analytics becomes indispensable for teams aiming to stay ahead of the curve. By harnessing the power of advanced statistical techniques and machine learning algorithms, this project endeavors to provide teams with actionable insights into player performances, facilitating informed decision-making in team selection, strategic planning, and performance optimization. Through meticulous analysis of batting, bowling, and fielding statistics, coupled with the application of predictive modeling, teams can gain a deeper understanding of player strengths, weaknesses, and potential, enabling them to adapt their strategies dynamically to match situations and opponents. Ultimately, the objective is to empower teams with the tools and knowledge needed to excel in the exhilarating world of T20 cricket, where split-second decisions and strategic acumen often make the difference between victory and defeat.

### 2. Objective

**In-depth Statistical Analysis:** Conduct a thorough examination of batting, bowling, and fielding statistics spanning from the inception of T20 cricket in 2007 to the present, leveraging advanced statistical techniques to uncover hidden trends and performance patterns.

**Data Preprocessing and Cleaning:** Preprocess the raw data to handle missing values, standardize column names, and remove redundant or irrelevant columns to ensure data integrity and streamline analysis.

**Application of Machine Learning:** Apply a range of supervised and unsupervised learning algorithms to categorize players, predict performance outcomes, and identify talent for strategic team selection.

**Feature Engineering:** Implement feature engineering methods such as weighted averages to provide a more nuanced understanding of player performance, considering factors like consistency and longevity.

**Predictive Modeling:** Develop predictive models to classify player performance levels and forecast future outcomes, enabling teams to make data-driven decisions in player selection and strategy formulation.

**Actionable Insights:** Provide actionable insights for team management and strategic planning, focusing on identifying key players, tactical opportunities, and potential areas for improvement.

### 3. Data Source

The project utilizes comprehensive datasets sourced from reputable sources like Kaggle, encompassing batting, bowling, and fielding statistics of T20 cricket players. These datasets offer a rich repository of individual player metrics, team performance records, match outcomes, and contextual data, providing a holistic view of the intricacies within T20 cricket.

### 4. Data Preprocessing

**Handling Missing Values:** Employ advanced imputation techniques to address missing values, ensuring robustness and accuracy in subsequent analyses.

**Standardizing Column Names:** Standardize column names across datasets to enhance clarity and consistency, facilitating seamless data integration and interpretation.

**Feature Selection:** Utilize domain knowledge and statistical techniques to identify and retain relevant features while discarding redundant or noisy ones, streamlining the data preparation process.

### 5. Supervised Learning

**Decision Tree:** We used this algorithm in predicting whether a bowler is likely to take more than the median number of wickets based on features like 'Econ' (Economy Rate), 'Ave' (Average Runs per Wicket), and 'SR' (Strike Rate).

**Logistics regression:** For the task of predicting whether a batsman's performance based on their batting average and strike rate is high, logistic regression offers a clear, interpretable model that quantifies the likelihood of being a high performer.

### 6. Unsupervised Learning

**K-Means Clustering:** Employ K-means clustering to group players based on similarities in their performance profiles, aiding in talent scouting, team selection, and strategic planning by

Analysis of T20 Cricket Player  
uncovering hidden patterns and player categories.

## **7. Feature Engineering Methods**

**Weighted Average Method:** Implement weighted averages to compute a comprehensive measure of batting and bowling performance, accounting for both the quality and quantity of player contributions over time, thus providing deeper insights into player consistency and effectiveness.

## **8. Advanced Analytical Techniques**

**Batting Analysis Using Linear Regression:** Conduct linear regression analysis to explore the relationships between various batting metrics such as average, strike rate, and overall player performance, shedding light on key factors influencing batting prowess and aiding in predictive modeling.

**Bowling Analysis Using Decision Tree Regression:** Employ decision tree regression to analyze bowling performance, identifying critical performance indicators such as bowling average, economy rate, and wickets taken, and predicting future bowling outcomes.

## **9. Cricket Bowling and Batting Performance Clustering**

**Bowling Performance Clustering:** Conduct clustering analysis to categorize bowlers based on their performance metrics, enabling teams to identify distinct bowling styles, strengths, and weaknesses, and tailor bowling strategies accordingly.

**Batting Performance Clustering:** Similarly, cluster batsmen based on performance metrics to delineate different batting styles and capabilities, facilitating strategic decision-making in team composition and batting order optimization.

## **10. Statistical Analysis of Batting Performance:**

This section delves into the intricacies of batting performance analysis, exploring key metrics such as batting average, strike rate, and consistency. Through detailed statistical analyses, trends in player performance over time are identified, shedding light on factors influencing batting prowess. Moreover, advanced techniques such as time-series analysis unveil patterns in player form and

Analysis of T20 Cricket Player  
performance fluctuations across different phases of a tournament or season.

### **11. Bowling Performance Evaluation:**

In this segment, the focus shifts to the bowling department, where metrics like economy rate, bowling average, and wicket-taking ability take center stage. Through rigorous statistical scrutiny, standout performers and areas for improvement are identified, guiding strategic decisions around bowling rotations, matchups, and tactics. Furthermore, the analysis extends beyond individual performance to assess bowling partnerships, uncovering synergies and optimizing bowling combinations for maximum impact.

### **12. Fielding Metrics and Impact:**

Fielding is often the unsung hero of cricket, influencing match outcomes through crucial catches, run-outs, and fielding efficiency. This section employs a multifaceted approach to evaluate fielding performance, incorporating metrics such as catches per match, run-saving contributions, and fielding efficiency ratings. By quantifying the impact of fielding prowess on match results, teams can prioritize fielding drills, enhance agility, and elevate overall team performance on the field.

### **13. Team Dynamics and Cohesion:**

Beyond individual performances, successful teams thrive on strong camaraderie, communication, and cohesion. This chapter delves into team dynamics, analyzing factors such as team spirit, leadership effectiveness, and collective resilience. By leveraging qualitative insights alongside quantitative metrics, teams can foster a culture of collaboration, unity, and shared purpose, driving collective success on and off the field.

### **14. Strategic Recommendations and Decision Support:**

Drawing upon the rich tapestry of analytical insights, this section offers actionable recommendations to cricket team management. From player selection dilemmas to in-match strategy adjustments, each recommendation is grounded in empirical evidence and strategic foresight. By aligning data-driven insights with tactical acumen, teams can chart a course towards sustained success in the competitive landscape of T20 cricket.

## **15. Case Studies and Success Stories:**

This chapter showcases real-world examples of teams implementing data-driven strategies to achieve remarkable success in T20 cricket. From underdog triumphs to dominant performances on the global stage, each case study highlights the transformative power of analytics in shaping cricketing fortunes. By distilling key learnings and best practices from these success stories, teams can glean valuable insights to inform their own strategic journeys.

## **16. Fantasy 11 Analysis**

**Data Preparation:** Aggregate and preprocess data for top T20 bowlers and batsmen, transforming them into Pandas DataFrames for efficient manipulation and analysis.

**Random Data Shuffling:** Randomly shuffle the DataFrames to eliminate any bias in player selection, ensuring fairness and reproducibility in fantasy team formation.

**Performance-Based Selection:** Select top performers based on key metrics such as batting average, bowling average, and strike rate, providing valuable insights for fantasy team managers and enthusiasts.

## **17. Conclusion**

This project underscores the importance of data-driven decision-making in T20 cricket, offering valuable insights into player performance and team dynamics. By leveraging advanced analytics techniques and predictive modeling, teams can optimize player selection, devise effective strategies, and gain a competitive edge in the fast-paced world of T20 cricket.

In this final chapter, the report culminates with a compelling call to action, urging stakeholders across the cricketing ecosystem to embrace the transformative potential of data-driven decision-making. By fostering a culture of curiosity, collaboration, and continuous improvement, teams can harness the power of analytics to propel cricketing excellence to unprecedented heights. Together, let us embark on a journey of innovation, insight, and inspiration, as we redefine the boundaries of what's possible in the thrilling world of T20 cricket.

## **18. Future Work**

As the cricketing landscape continues to evolve, this section explores emerging trends and future

### Analysis of T20 Cricket Player

directions in T20 cricket performance analysis. From the rise of player tracking technologies to the integration of artificial intelligence in decision-making processes, the horizon is brimming with possibilities. By embracing innovation and staying abreast of technological advancements, teams can stay ahead of the curve and unlock new frontiers of performance excellence

**Advanced Machine Learning Algorithms:** Explore the application of advanced machine learning algorithms such as Random Forests, Gradient Boosting Machines, and Neural Networks to enhance the accuracy and robustness of predictive models.

**Integration of Additional Data Sources:** Incorporate additional data sources such as live match data, player fitness metrics, and historical performance data to enrich the analysis and provide more comprehensive insights.

**Sentiment Analysis and Social Media Data:** Integrate sentiment analysis of match commentary and social media data to gauge public perception, player popularity, and fan sentiment, providing valuable contextual insights for team management.

**Interactive Visualization Tools:** Develop interactive visualization tools and dashboards for coaches, analysts, and fans to dynamically explore player performance data, uncover trends, and derive actionable insights, fostering greater engagement and understanding of the game.