

Data-Driven Cricket Team Selection: Optimizing Squad Composition Using Binary Integer Programming

Rajesh S, Sagar R Nayak, Nishith R

B.Tech (Hons) Data Science

Vidyashilp University, India

Email: {2023rajesh.s, 2023sagar.nayak, 2023nishith.rvidyashilp.com}

Abstract—Cricket team selection is a complex decision-making process that requires evaluating multiple player abilities to form a well-balanced and competitive squad. Traditional selection methods rely on subjective judgment or simple statistical analysis, which may not always yield an optimal team composition. This paper explores an optimization approach using Binary Integer Programming (BIP) to select an optimal T20 cricket team based on statistical measures such as batting average, strike rate, economy rate, and fielding performance. The proposed model considers constraints such as team composition (batsmen, bowlers, all-rounders, and wicketkeepers), budget limitations, and foreign player restrictions. By employing Python-based solvers like PuLP, Gurobi, and CPLEX, the model ensures a data-driven, objective, and efficient selection strategy. Results indicate that BIP-based selection provides a more balanced and competitive squad compared to traditional methods, minimizing selection biases while maximizing team performance. This study contributes to the growing field of sports analytics by demonstrating how mathematical optimization can enhance team selection strategies in professional cricket.

Index Terms—Cricket, Team Selection, Optimization, Binary Integer Programming, Sports Analytics

I. INTRODUCTION

Cricket has evolved over the years, with increasing emphasis on statistical analysis for evaluating team and player performances. The emergence of T20 cricket has further intensified the need for rapid decision-making, requiring a well-balanced squad that maximizes both scoring potential and defensive capabilities.

II. PROBLEM STATEMENT

Selecting a cricket team involves multiple trade-offs between different player roles and skills. Traditional methods rely on selectors' experience and heuristics, which can introduce biases. To overcome this, we propose a mathematical optimization model using Binary Integer Programming (BIP) that ensures an objective, data-driven selection strategy while satisfying constraints such as: - The team must have exactly 11 players. - At least 5 batsmen and 5 bowlers should be selected. - At least one wicketkeeper must be included. - Additional constraints like budget and foreign player limits (if applicable).

III. METHODOLOGY

The methodology consists of three main steps:

A. Data Collection

Player performance data is gathered from SA domestic Pro20 league and includes: - Batting Metrics: Runs scored, Batting Average, Strike Rate - Bowling Metrics: Economy Rate, Bowling Average, Strike Rate - Fielding Metrics: Dismissals per match (for wicketkeepers)

B. Binary Integer Programming Model

We define decision variables where each player is assigned a binary value:

$$x_i = \begin{cases} 1, & \text{if player } i \text{ is selected} \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

The objective function maximizes the total team performance based on weighted indices for batting, bowling, and fielding:

$$\max \sum_{i=1}^n c_i x_i \quad (2)$$

where c_i represents the ability score of player i .

C. Constraints

$$\sum_{i \in \text{Batting}} x_i \geq 5 \quad (\text{At least 5 batsmen}) \quad (3)$$

$$\sum_{i \in \text{Bowling}} x_i \geq 5 \quad (\text{At least 5 bowlers}) \quad (4)$$

$$\sum_{i \in \text{Wicketkeeper}} x_i = 1 \quad (\text{Exactly 1 wicketkeeper}) \quad (5)$$

$$\sum_{i=1}^n x_i = 11 \quad (\text{Total team size is 11}) \quad (6)$$

The BIP model is implemented using Python with PuLP, Gurobi, or CPLEX to compute the optimal team selection.

The results are compared with traditional selection methods to validate effectiveness.

IV. RESULTS AND DISCUSSION

The optimization model selects a squad that maximizes team performance while satisfying role constraints. The optimized team is compared with manually selected teams, demonstrating: - A higher overall performance index - A more balanced team composition - Elimination of selection biases

V. CONCLUSION AND FUTURE WORK

This study demonstrates the effectiveness of Binary Integer Programming in cricket team selection. By integrating data-driven techniques, selection can be made more objective and optimized for performance. Future work could incorporate real-time player form adjustments and machine learning predictions to further enhance team selection.

REFERENCES

- [1] H. H. Lemmer, "The combined bowling rate as a measure of bowling performance in cricket," *South African Journal of Sports Science*, 2002.
- [2] D. Beaudoin and T. B. Swartz, "The best batsmen and bowlers in one-day cricket," *South African Statistical Journal*, 2003.
- [3] H. Gerber and G. D. Sharp, "Selecting a limited-overs cricket squad using an integer programming model," *South African Journal of Research in Sport*, 2006.
- [4] W. Brettenny, "Integer optimization for the selection of a fantasy league cricket team," *Unpublished master's dissertation*, 2010.