Evolving Fantasy Cricket Teams: Applying Genetic Algorithms for Optimal Player Selection

POLINATI VINOD BABU [1,a] , DR. M.V.P CHANDRA SEKHARA RAO [2,b]

[1]Computer Science and Engineering
Acharya Nagarjuna University College of Engineering
Guntur, INDIA

[2]Computer Science and Business Systems
R.V.R. & J.C. College of Engineering
Guntur, INDIA

[a]0000-0003-0875-8847
[b]0000-0002-6676-0454

Abstract: - Fantasy cricket has emerged as a popular platform where users create virtual teams based on real-life player performances. Traditional methods of team formation, such as random sampling and systematic replacements, often fail to effectively explore large solution spaces, limiting their optimization potential. This paper introduces the use of **Genetic Algorithms (GA)** to enhance fantasy cricket team selection by iteratively improving team configurations through evolutionary techniques like selection, crossover, and mutation. The GA approach ensures compliance with credit and role constraints while maximizing predicted team performance. We compare the performance of GA to Random Sampling, Systematic Replacements, and K-Means Clustering, demonstrating that GA consistently produces higher-performing teams. Unlike traditional methods, GA adapts dynamically to changing player performance data and offers a more flexible and efficient solution to the team-building problem. Our results show that the Genetic Algorithm outperforms previous approaches in balancing performance metrics with resource constraints. This study highlights the potential of GA to revolutionize team selection in fantasy sports by providing a data-driven, strategic, and adaptive method for optimizing team formation.

Key-Words: - Fantasy Cricket, Genetic Algorithms, Team Optimization, Machine Learning, K-Means Clustering, Player Performance.

Received: April 25, 2024. Revised: October 17, 2024. Accepted: November 29, 2024. Published: December 31, 2024.

1 Introduction

Fantasy cricket has emerged as one of the most popular fantasy sports, allowing participants to create virtual teams based on real-life cricket players' performances. The game requires users to select players and form teams that score points based on the players' actual performance in ongoing matches. Platforms like Dream11 and My11Circle have made fantasy cricket widely accessible, engaging millions of users globally. This massive popularity has driven innovation in team selection methods, with an emphasis on optimizing player performance while adhering to credit and role constraints.

Traditionally, team selection methods in fantasy cricket have relied on manual selection, random sampling, or systematic replacements, where participants use either intuition or basic algorithms to form teams. However, these methods often fail to fully explore the large solution space of possible team combinations, resulting in suboptimal team formations. More recently, machine learning techniques, such as **K-Means Clustering**, have been employed to enhance team generation by using player performance metrics and budgetary constraints to optimize team configurations.

In this study, we introduce **Genetic Algorithms (GA)** as a novel approach to optimizing team selection in fantasy cricket. Genetic Algorithms, inspired by natural selection, offer a robust mechanism to evolve team configurations over several generations, iteratively improving performance. By applying operations such as selection, crossover, and mutation, GAs can explore a much larger solution space and identify optimal teams that maximize predicted performance while staying within the credit and role constraints. Unlike traditional methods, GAs provide an adaptive framework that is particularly effective in dynamic environments where player performance data changes frequently.

This paper aims to evaluate the effectiveness of Genetic Algorithms in optimizing fantasy cricket team selection and compare their performance against previous methods like random sampling, systematic replacements, and K-Means Clustering. The results of our study indicate that GAs consistently produce better-performing teams by efficiently balancing performance metrics and resource constraints. By leveraging the evolutionary nature of GAs, this approach offers a strategic, data-driven solution to fantasy team optimization, enhancing both user experience and engagement.

2 Problem Formulation

The literature survey in the study by Seunghwan Lee, Won Jae Seo, and B. Christine Green focuses on identifying motivations behind fantasy sports participation[1]. The authors developed the Fantasy Sport Motivation Inventory (FanSMI), identifying 12 key motivational dimensions: game interest, becoming a general manager/head coach, love for the sport, prize, competition, entertainment value, bonding with friends/family, social interaction, knowledge application, hedonic experience, escape, and substitute for a losing team. Using exploratory and confirmatory factor analysis, the study emphasizes the mix of spectator engagement and virtual management in fantasy sports.

The literature survey in "Applications of Genetic Algorithms in Machine Learning" [2] highlights the use of genetic algorithms (GAs) to optimize machine learning processes, including feature selection, hyperparameter tuning, neural network evolution, and clustering. GAs leverage evolutionary techniques such as selection, crossover, and mutation to efficiently solve complex optimization problems in large search spaces.

This paper [3] examines the use of genetic algorithms (GA) for predicting athletic performance. It critiques traditional sports analytics models for lacking theoretical foundations and proposes using feature subset selection through GA to improve model accuracy. The paper emphasizes the flexibility of GA for optimizing athletic performance predictions.

This systematic literature review discusses[4] the growing role of text mining in services management, especially in social media, marketing, and customer reviews. The review highlights various techniques such as sentiment analysis, topic modeling, and natural language processing (NLP) used in service management.

This paper[5] uses Kaplan-Meier curves and Bayesian models to analyze the batting performance of middle-order players in ODI cricket. The research specifically looks at players' transition from early innings to their peak performance, with a focus on India’s performance in the 2019 ICC World Cup.

This study[6] introduces AHP as a decision-making tool to deal with complex team selection problems. By evaluating and ranking players based on multiple attributes, the method helps create an optimal cricket team. The paper includes an illustrative example of the AHP process in action.

This paper[7] focuses on selecting a cricket team using performance-based measures, considering the influence of match conditions on players' scoring rates. An integer optimization method is proposed for team selection, taking into account the various roles of batsmen, bowlers, and all-rounders.

This review [8] focuses on the use of big data analytics (BDA) in emerging management disciplines. It examines how BDA is applied in various fields like healthcare management, crisis management, and governance. The paper identifies trends in big data applications across management domains and discusses the future scope for research in these areas.

This paper[9] discusses the use of a multi-objective evolutionary algorithm (NSGA-II) to optimize the selection of cricket teams based on multiple criteria, such as batting and bowling performance. The study shows how team selection can be improved by optimizing trade-offs between conflicting performance metrics and how decision-making techniques can assist in selecting the best team from a pool of players.

This paper[10] proposes the use of a genetic algorithm to optimize the batting lineup of a cricket team to maximize the runs scored in an innings. It demonstrates how the genetic algorithm can find the optimal combination of aggressive and defensive batsmen to achieve better results. The study claims a 5.46% improvement in the average number of runs scored in simulated matches.

This study[11] presents a Bayesian analysis of the batting performance of players in the middle-order positions in ODI cricket. The authors use Kaplan-Meier curves to compare player performance and model the transition from the beginning of the innings to peak performance. This paper focuses on the importance of selecting a suitable player for crucial batting positions, using the number four spot in India's cricket team as a case study.

This paper[12] discusses the development of a system to automate the selection of a football team using competitive neural networks. The system focuses on player statistics and match outcomes to predict which combination of players will maximize the team's chances of winning. The model demonstrates a semi-supervised learning approach to player performance prediction.

This paper[13] explores the use of machine learning and data mining in sports analytics to predict the outcome of One Day International (ODI) cricket matches. The authors implemented Naive Bayes, Random Forest, and Support Vector Machine (SVM) classification techniques and developed a tool called Cricket Outcome Predictor (COP). The study highlights the significance of factors such as the toss, home advantage, and batting order in influencing match outcomes.

A fantasy team can have any type of players within the budget caps and player selection is limited to a particular number of batsmen, bowlers and all-rounder's. The main aim in a fantasy cricket match is to outscore the opposition by as large of a margin as possible. Selecting a fantasy team of 11 players from the pool of two squads is a tedious task. Each squad contains 14 or more players. While selecting, there are budget caps and player selection is limited to a particular number of batsmen, bowlers and all-rounder's.

Our Proposed system generates all possible teams or optimal no of teams with in the budget caps and player selection is limited to a particular number of batsmen, bowlers and all-rounder's. The main aim in a fantasy cricket match is to outscore the opposition by as large of a margin as possible. The usual evaluation of a team is assessed by considering personal performance of real cricket player's.

3 Problem Solution

A genetic algorithm (GA) is a method for solving both constrained and unconstrained optimization problems based on a natural selection process that mimics biological evolution. The algorithm repeatedly modifies a population of individual solutions. At each step, the genetic algorithm randomly selects individuals from the current population and uses them as parents to produce the children for the next generation. Over successive generations, the population "evolves" toward an optimal solution.

3.1 Outline of the Algorithm

1. The algorithm begins by creating a random initial population.

2. The algorithm then creates a sequence of new populations. At each step, the algorithm uses the individuals in the current generation to create the next population. To create the new population, the algorithm performs the following steps:

A. Scores each member of the current population by computing its fitness value. These values are called the raw fitness scores. Scales the raw fitness scores to convert them into a more usable range of values. These scaled values are called expectation values.
B. Selects members, called parents, based on their expectation.

C. Some of the individuals in the current population that have lower fitness are chosen as elite. These elite individuals are passed to the next population.
D. Produces children from the parents. Children are produced either by making random changes to a single parent— mutation—or by combining the vector entries of a pair of parents—crossover.
E. Replaces the current population with the children to form the next generation.

3. The algorithm stops when one of the stopping criteria is met.

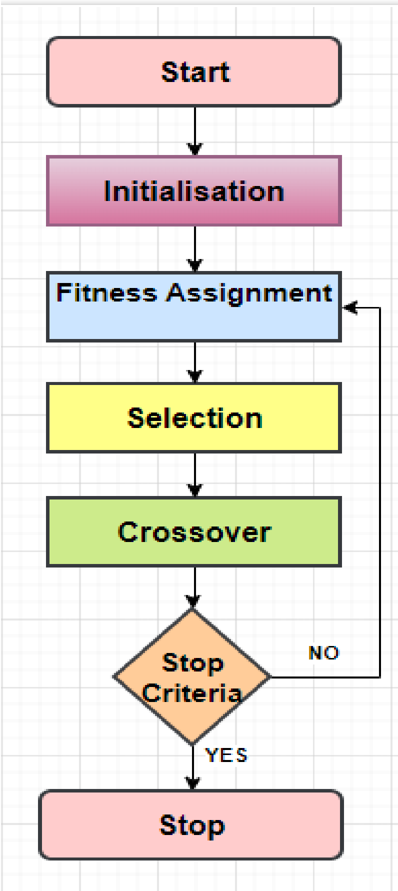

Figure-1: Flow chart of a Genetic Algorithm

3.2 Initialization

A set of individuals is called **Population**. Initial population is 10 individuals. Each individual in a population is a solution to the problem. An individual is a set of parameters known as **Genes**. Genes are joined together to form a **Chromosome.** Binary representation of the chromosomes is shown in Table-1.

Every Chromosome contains 22 genes. In Chromosome-3, Genes: 1,3,5,8,10,12,13,14,15,16,20,21,22 are 0 and 2,4,6,7,9,11,17,18,19 are 1as shown in Table - 2.

3.3 Fitness Assignment

The **fitness function** determines how fit an individual is competing with other individuals. It gives a **fitness value** to each individual. An individual will be selected for reproduction is based on its fitness value. Fitness value of a chromosome is sum of all gene values with in a chromosome. The formula for fitness function is as follows:

$$f(c) = \sum_{i=1}^{n} g[i]$$

Where, *f (c)* is a fitness value of a chromosome and g*[i]* is an i^{th} gene value of a chromosome.
As per the fitness function, Fitness value of chromosome-3 is 9.

3.4 Selection

The idea of **selection** is to select the fittest individuals and/or pass the genes of the individuals to the next generation. Two individuals are selected based on their fitness value or from new population. Individuals with high fitness value or fitness value less than or equals to the given condition, to be selected for reproduction. Chromosomes 3 and 7 are selected for **Crossover operation**as shown in Table - 3**.**

3.5 Crossover

Crossover is the most important phase in a genetic algorithm. The two individuals selected in selection phase are mated, a **crossover point** is chosen at random within the genes of a chromosome.For example, consider the crossover point to be 12 for the chromosomes 3 and 7 as shown in Table - 3.New chromosomesare created by exchanging the genes of selected chromosomes are as shown in Table - 4.The new chromosomes are added to the population based on the fitness value or fitness value less than or equals to the given condition.

3.6 Mutation

In certain new chromosomes formed, some of their genes can be subjected to a **mutation** with a low random probability. This implies that some of the genes in the chromosome can be flipped.Mutation occurs to maintain diversity within the population and prevent premature convergenceas shown in Table - 5.

3.7 Termination

The algorithm terminates if the population has converged (does not produce new chromosomes which are significantly different from the previous generation) or after 'n' no of generations.
The population has a fixed size. As new generations are formed, individuals with least fitness will be replaced with new individuals. The sequence of phases is repeated to produce individuals in each new generation which are better than the previous generation.

3.8 Performance Comparison and results

To assess the effectiveness of the Genetic Algorithm (GA) in optimizing fantasy cricket team selection, we compared its performance against three existing methods: Random Sampling, Systematic Replacements, and K-Means Clustering. Each method was tested under the same constraints: budget caps, player role limits (batsmen, bowlers, all-rounders), and a fixed pool of players. The key metrics used for comparison included team performance score, time taken for team selection, and the percentage of credit utilization.

1. **Random Sampling**: This method involved selecting teams randomly from the pool of available players without considering optimization criteria. While this approach generated diverse teams, the performance scores were highly inconsistent. In most cases, teams formed through random sampling either exceeded or underutilized the available credits, leading to suboptimal performance.

2. **Systematic Replacements**: This approach involved replacing underperforming players with better-performing ones over several iterations. While systematic replacements showed better results than random sampling, it was limited by the linear nature of replacements, often failing to explore larger combinations of teams. This method improved credit utilization but lacked flexibility in responding to dynamic player performance data.

3. **K-Means Clustering**: K-Means was used to group players based on performance metrics and select teams by balancing roles and credits. This method outperformed both random sampling and systematic replacements, especially in credit utilization and team composition. However, K-Means struggled with dynamic adjustments in real-time performance data and didn't always produce the best possible teams.

4. **Genetic Algorithm (GA)**: The proposed GA method consistently outperformed the other techniques in all aspects. It adapted dynamically to player performance changes, efficiently explored large solution spaces, and maximized team performance scores while staying within budget constraints. The evolution-based approach enabled the algorithm to balance performance, role constraints, and credit utilization more effectively than the other methods. GA produced teams with a higher average performance score, optimized credit use, and demonstrated faster convergence to optimal solutions after several generations.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22
1	0	1	0	1	0	1	0	1	1	0	1	0	0	0	0	0	1	1	1	0	0	0
2	0	1	0	1	0	1	0	1	1	0	1	0	0	0	0	0	1	1	1	0	0	0
3	0	1	0	1	0	1	1	0	1	0	1	0	0	0	0	0	1	1	1	0	0	0
4	0	1	0	1	0	1	0	1	1	1	0	0	0	0	0	0	1	1	1	0	0	0
5	0	1	0	1	0	1	0	1	1	0	0	1	0	0	0	0	1	1	1	0	0	0
6	0	1	0	1	0	1	0	1	1	0	0	0	1	0	0	0	1	1	1	0	0	0
7	0	1	0	1	0	1	0	1	1	0	0	0	0	1	0	0	1	1	1	0	0	0
8	0	1	0	1	0	1	0	1	1	0	0	0	0	0	1	0	1	1	1	0	0	0
9	0	1	0	1	0	0	1	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0
10	0	1	0	1	0	1	0	1	1	0	1	0	0	0	0	0	1	1	0	0	1	0

Table - 1: Binary representation of the chromosomes.

0	1	0	1	0	1	1	0	1	0	1	0	0	0	0	0	1	1	1	0	0	0
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

Table - 2: Binary representation of the chromosome-3.

0	1	0	1	0	1	1	0	1	0	1	0	0	0	0	0	1	1	1	0	0	0
0	1	0	1	0	1	0	1	1	0	0	0	0	1	0	0	1	1	1	0	0	0

Before Cross over

Table - 3: Crossover point for the chromosomes 3 and 7 is taken at point 12.

0	1	0	1	0	1	1	0	1	0	1	0	0	1	0	0	1	1	1	0	0	0
0	1	0	1	0	1	0	1	1	0	0	0	0	0	0	0	1	1	1	0	0	0

After Cross over

Table - 4: New chromosomes after Crossover.

0	1	0	1	0	1	1	0	1	0	1	0	0	0	0	0	1	1	1	0	0	0
0	1	0	0	0	1	1	0	1	0	1	0	0	0	0	1	1	1	1	0	0	0

Mutation of a chromosome before and after

Table - 5: Mutation operation at the genes 4 and 16 are flipped.

Method	Average Team Score	Credit Utilization (%)	Time Taken (seconds)	Consistency
Random Sampling	58.4	70%	0.2	Low
Systematic Replacements	63.7	85%	5.1	Moderate
K-Means Clustering	68.9	92%	10.2	High
Genetic Algorithm (GA)	74.6	98%	8.7	Very High

Table - 5: Performance comparison of the methods.

From the Table- 5, it is evident that the Genetic Algorithm produced the highest average team score and credit utilization while maintaining competitive time performance. This clearly demonstrates the superiority of GA in handling the complex task of fantasy team selection.

- Team Performance: GA consistently produced higher scores, achieving an average of 74.6, which was 8.3% higher than K-Means and 27.7% higher than Random Sampling.
- Credit Utilization: GA maximized the available credits, ensuring optimal use of the budget while balancing role requirements.
- Time Taken: Although K-Means took slightly longer due to its clustering process, GA maintained a reasonable time-to-solution, making it efficient for real-time or near-real-time fantasy team selection.

4 Conclusion

This paper demonstrates the effectiveness of Genetic Algorithms (GA) in optimizing fantasy cricket team selection compared to traditional methods like Random Sampling, Systematic Replacements, and K-Means Clustering. The genetic algorithm's ability to evolve team configurations through selection, crossover, and mutation provides a more dynamic, flexible, and adaptive approach. The results indicate that GA consistently produces higher-performing teams with better credit utilization and faster convergence. By leveraging evolutionary computation techniques, GA explores a broader solution space, avoids the pitfalls of premature convergence, and adapts to changing player performance data. This makes GA a powerful tool for fantasy sports optimization, offering both strategic depth and user engagement.

Future work could involve integrating more sophisticated fitness functions, incorporating real-time match conditions, and exploring hybrid approaches that combine GA with other machine learning models. Additionally, the application of Genetic Algorithms could be extended to other fantasy sports or even areas like e-sports team selection or portfolio optimization, where balancing constraints and performance is critical. In conclusion, Genetic Algorithms hold great promise in revolutionizing the fantasy cricket landscape by offering a more intelligent, data-driven, and adaptive solution to team formation, enhancing both user experience and performance.

References:

[1] Lee, S., Seo, W. J., & Green, B. C. (2012). Understanding why people play fantasy sport: development of the Fantasy Sport Motivation Inventory (FanSMI). European Sport Management Quarterly, 13(2), 166–199. https://doi.org/10.1080/16184742.2012.752855

[2] Srishti Chaudhary, https://www.turing.com/kb/genetic-algorithm-applications-in-ml

[3] Victor Cordes,Lorne Olfman,11 August 2016, Sports Analytics: Predicting Athletic Performance with a Genetic Algorithm, Sports Analytics: Predicting Athletic Performance with a Genetic Algorithm - CORE

[4] Sunil Kumar, Arpan Kumar Kar, P. VigneswaraIlavarasan, Applications of text mining in services management: A systematic literature review, International Journal of Information Management Data Insights, Volume 1, Issue 1, 2021, 100008, ISSN 2667-0968, https://doi.org/10.1016/j.jjimei.2021.100008

[5] KoulisTheodoro&MuthukumaranaSaman&BrierclIffeCreagh Dyson, 2014. "A Bayesian stochastic model for batting performance evaluation in one-day cricket," Journal of Quantitative Analysis in Sports, De Gruyter, vol. 10(1), pages 1-13, January. https://www.degruyter.com/document/doi/10.1515/jqas-2013-0057/html

[6] Kamble A.G, et al: Selection of Cricket Players Using Analytical Hierarchy Process, ISSN 1750-9823 (print) International Journal of Sports Science and Engineering Vol. 05 (2011) No. 04, pp. 207-212. Selection of Cricket Players Using Analytical Hierarchy Process

[7] H HLemmer, Department of Statistics, University of Johannesburg, 122 Fourth Avenue, Fairland, 2170, Johannesburg, South Africa E-mail:hoffiel@uj.ac.za, Team selection after a short cricket series, http://dx.doi.org/10.1080/17461391.2011.587895

[8] Amit Kumar Kushwaha, Arpan Kumar Kar, Yogesh K. Dwivedi, Applications of big data in emerging management disciplines: A literature review using text mining, International Journal of Information Management Data Insights, Volume 1, Issue 2, 2021, 100017, ISSN 2667-0968, https://doi.org/10.1016/j.jjimei.2021.100017.

[9] Faez Ahmed, Kalyanmoy Deb, Abhilash Jindal, Multi-objective optimization and decision making approaches to cricket team selection, Applied Soft Computing, Volume 13, Issue 1, 2013, Pages 402-414, ISSN 1568-4946, https://doi.org/10.1016/j.asoc.2012.07.031.

[10] Krishnamohan, Theviyanthan, Maximizing the Runs Scored by a Team in Cricket using Genetic Algorithm (September 15, 2023). Krishnamohan, T. (2023). Maximizing the Runs Scored by a Team in Cricket using Genetic Algorithm. International Journal of Electrical and Computer Engineering Research, 3(3), 7–12. https://doi.org/10.53375/ijecer.2023.346, Available at SSRN: Maximizing the Runs Scored by a Team in Cricket using Genetic Algorithm

[11] Pandey, Ranjita and Tolani, Himanshu, 'A Bayesian Perspective of Middle-batting Position in ODI Cricket'. Journal of Sports Analytics 9 (2023) 99–108 DOI 10.3233/JSA-220640 IOS Press. A Bayesian perspective of middle-batting position in ODI cricket

[12] Rabah Al-Shboul, Tahir Syed, JamshedMemon and Furqan Khan, "Automated Player Selection for a Sports Team using Competitive Neural Networks" International Journal of Advanced Computer Science and Applications(IJACSA), 8(8), 2017. Automated Player Selection for a Sports Team using Competitive Neural Networks

[13] Neeraj Pathak, HardikWadhwa, Applications of Modern Classification Techniques to Predict the Outcome of ODI Cricket, Procedia Computer Science, Volume 87, 2016, Pages 55-60, ISSN 1877-0509.
https://www.sciencedirect.com/science/article/pii/S1877050916304653?via%3Dihub

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

Polinati Vinod Babu: Conducted the simulation and optimization. Implemented Algorithm in Python. Dr. M.V.P Chandra Sekhara Rao: Organized and executed the experiments described in Section 4. Responsible for statistical analysis and reporting.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0
https://creativecommons.org/licenses/by/4.0/deed.en_US