

## Group Assignment 1: Group 4

### Optimizing NFL Team Construction Using a Genetic Algorithm

**Evan Fox, Sumedh Khodke, Dineshkumar Katta, Sanjay Aravind Loganathan  
Ravichandran, Amritha Subburayan**

**Answer the following questions to illustrate the Motivation and Significance of the project. You may answer three questions in three separate paragraphs. Proper references should be cited in the text of your report.**

**a) Discuss the significance of the project through the lens of statistical evidence. Include available recent statistics from the literature. Show the rationale for more scientific studies in this subdomain (specific to your project). [Suggested Word Count: 200 words]**

The significance of our project is to create better team construction strategies. Team management will be able to calculate the optimal team construction path, either in the offseason or for individual transactions. Roster construction is critical in any sport for the team's subsequent performance, so this solution will be applicable to other sports, outside of American Football. Many factors influence the selection process, ranging from current team construction and player skill to financial constraints. All of this complicates the process of selecting a 53-man roster, as there are too many permutations for a greedy-search heuristic and many parameters that influence success. The goal of this project is to use a Genetic Algorithm to objectively select players for a football team by integrating with an ML-based fitness function that is used to evaluate a team's quality (or survival score). The results will allow for the development of various solutions for team construction that maximize financial resources, for teams in specific situations. This can be done by initializing a roster with a pre-set portion of the roster – these players will not be mutated in the genetic algorithm.

Statistical measures of success are not particularly achievable in this sub-domain, because solutions are non-deterministic - i.e., team success is not exclusively based on team quality. For instance, the paper, *A Generic Approach for Team Selection*<sup>1</sup>, utilizes a genetic algorithm with a tailored, rather than learned, fitness function. This approach showed success in constructing Cricket teams. Furthermore, the player's individual performance was not considered throughout the team selection process. We believe this approach can be expanded on with our idea of an ML-based, learned fitness function.

---

<sup>1</sup> [A Generic Approach for Team Selection in Multiplayer Games Using Genetic Algorithms](#)

**b) Discuss the major advancements to address the significant research questions in this subdomain (specific to your project). Summarize the uncertainties, challenges, and further research opportunities centered around these major advancements. Keep the discussion succinct. [Suggested Word Count: 300 words]**

The paper, *Predicting Athletic Performance with a Genetic Algorithm*<sup>2</sup>, discusses utilizing genetic algorithms combined with decision trees to determine optimal starting players from a roster. We plan to repurpose this architecture for determining the best roster construction of a team, starting by initializing a roster with a pre-set portion of the roster as immutable players, which will complement the selection of the rest of the players for our genetic algorithm. A major advancement on standard genetic algorithm optimization is to leverage a machine learning model as our fitness function/survival score. This would allow us to train our fitness function on previous team construction and successes – leading to potentially improved roster construction. One uncertainty is whether the learned fitness function is accurate and useful, as this can be challenged to gauge, given the complexity of an NFL team's success. Another uncertainty is how we account for player contracts and salaries changing over time.

Our development plan includes opportunities for further advancement. This could include expanding on existing machine learning architectures to better suit our scenario – one option is to modify XGBoost algorithm to more appropriately consider relationships between positions. Another improvement could be in the form of model input, adding context with which players have played together and had success in previous seasons – this would add in a sense of combinatorial advantage. As a visual aide to explain the architecture and system, clustering to seek well-valued players will also be done.

**c) Explain the potentials of your project to contribute to addressing any of these uncertainties, challenges, and further research opportunities. Discuss why this contribution is crucial. [Suggested Word Count: 200 words]**

This project has potential to improve team construction analytics by combining two common approaches to prediction and optimization. We expect our project to strive to pick players with similar scores and positions, that have a high return on investment. To address uncertainty in our learned fitness function, we plan to start by utilizing pre-determined player scores from the popular video game, Madden. This will allow us to leverage a proven single-metric representing player quality. We can then take steps to improve this learned fitness function by utilizing additional, more specific, inputs such as yards, tackles, and touchdowns. By considering these factors, a weighted average is calculated for each team and the roster with best weighted average is selected by the fitness function. To address uncertainty in player salary, change over time, we will begin by assuming player pay remains constant – this is a helpful assumption for developing and testing the other aspects of our project. Future improvement can come from extrapolating future player salaries based on historical trend. Our research and genetic algorithm could further be extended to other multi-player sports which require a more robust team selection approach. This contribution to the current state of sports analytics, could help to synthesize two common strategies, machine learning and evolutionary algorithms, into a more cohesive and successful team construction tool.

---

<sup>2</sup> [Predicting Athletic Performance with a Genetic Algorithm](#)