1. **Project Group 10:** Does Athleticism and Draft Priority Translate to Success in the NFL?
2. Vitush Agarwal (v8agarwal@ucsd.edu), Praveen Manimaran (pmanimaran@ucsd.edu), William Guy (wguy@ucsd.edu)
3. **Abstract Details**

- **Background:** The NFL combine and draft are extremely influential for teams to use to determine the likely success of prospective players. Teams invest significant resources into assessing athletes using metrics such as Relative Athlete Score (RAS) which ranks players on a scale from 1 (lowest) to 10 (highest) by height, weight, wingspan, and drill exercises all into one aggregate number that's then contextualized against other athletes of the same position. Predicting player performance in professional games is a challenging and imprecise science, however through applying data science and ML models to the data, we aim to build a predictive model that correlates these inputs with career outcomes, enabling a deeper understanding of the factors that drive success at the professional level. Fantasy points is a widely used metric in fantasy football leagues, providing a comprehensive and standardized measure of a player's contributions across various aspects of the game (e.g., rushing, receiving, and scoring). By applying data science and ML models to these inputs, we aim to build a predictive model that correlates combine and draft metrics with career outcomes, enabling a deeper understanding of the factors that drive success at the professional level.
- **Problem Definition:** Our goal is to predict a NFL player's average fantasy points scored per season using their combine and draft data. By doing so, we aim to correlate athletic and draft evaluations with professional performance. Our expected inputs are raw NFL Combine data *(e.g., 40-yard dash, vertical jump, Relative Athletic Score)* and draft data *(e.g., draft round, overall pick).* Our expected output is a numerical prediction of average fantasy points per season, which serves as a proxy for career success. These predictions can provide NFL teams with valuable insights into how different metrics impact player performance, allowing for more informed draft decisions and resource allocation.
- **Motivation:** Drafting is extremely high risk, with over a third of first round picks not making it to a second contract on the team that drafted them, accurately predicting player performance can reduce risks and maximize value. Machine learning models are ideal for identifying trends in high dimension data such as combine records and draft capital. We will use a publicly available Kaggle NFL dataset that has NFL combine and draft data combined with historical fantasy football stats. This dataset spans 10+ years, providing sufficient data for training, validation, and a test group. We will split the dataset into training (70%), validation (15%), and test groups (15%) to ensure robust evaluation.
- **Literature Review:** Existing studies such as Applying Machine Learning to Predict the NFL Draft and Predicting a Player's NFL Draft Round have explored the correlation between metrics and performance for specific positions, utilizing linear regression and basic statistical analysis to predict the NFL draft outcomes. Instead of predicting NFL draft outcomes, we want to predict a player's career success(using fantasy points as a metric). We will account for the complex nature of player performances by incorporating machine learning methods such as Random Forest, Regression, Gradient Boosting and a feedforward neural network.
- **Our Approach:** We plan to explore ML models such as Random Forest (robust to overfitting and interpretable), Gradient Boosting (handles non-linear relationships well), and Support Vector Machines (effective in high-dimensional spaces). We hope to also build a feedforward neural

network as it is well-suited for modelling complex non-linear interactions between combine metrics and draft data, which are likely to influence player success. Unlike traditional models, a neural network can learn intricate feature relationships without extensive manual feature engineering. By comparing its performance to simpler baseline models, we aim to demonstrate the advantages of leveraging neural networks for high-dimensional and non-linear data.

4. **Extra Details**: Ensemble methods will help capture complex patterns, while feature selection techniques will identify key predictors of success. For evaluation, we will compare predictive performance using metrics like Mean Absolute Error (MAE) and R-squared.

5. **Success Criteria:** The project's success will be evaluated based on a comparison of our predictive accuracy regarding career success measured by average fantasy points per season, the identification of key factors that point to influencing NFL success, and practical insights into how athletic and draft evaluations inform team decision-making. Achieving an MAE of less than 25 fantasy points per season and an R-squared value above 0.7 would be considered successful. We would also consider our project to be successful if we can identify key metrics that consistently influence performance (e.g., speed for wide receivers or agility for running backs) as it can provide actionable insights for NFL teams.