**Project Proposal**

Project 3: UTA-VIRT-DATA-PT-04-2023-U-LOLC-MTTH Cohort

**Project Proposal:** Predicting NFL Draft Round/Pick Order

**Project Overview:**

* **Project Team Members:**
* Charles Quinn [inlikequinn1@gmail.com](mailto:inlikequinn1@gmail.com)
* Alice Liu [alice.liu053@gmail.com](mailto:alice.liu053@gmail.com)
* Brandon Loredo [bloredo1987@gmail.com](mailto:bloredo1987@gmail.com)
* Bryan Miller
* **Project Scope and Objectives:**

The objective of this project is to develop a predictive machine learning algorithm leveraging NFL Combine results data spanning from 2000 to 2015. This algorithm is designed to assess an athlete's likelihood of transitioning into a professional NFL career based on their performance metrics during the Combine. By thoroughly analyzing this historical dataset, we intend to uncover crucial patterns and trends that significantly influence a player's potential NFL success. The primary aim is to provide valuable insights for both players and scouts, aiding in decision-making processes and potentially revolutionizing talent evaluation within the NFL.

* **Key Objectives:**
* Assess NFL Combine data from 2000 to 2015
* Identify key performance metrics crucial for NFL success
* Develop a machine learning model for player evaluation
* Evaluate the model's predictive accuracy and reliability
* Provide actionable insights for players and scouts
* Enhance talent evaluation processes within the NFL
* Enable data-driven decision-making in player selection
* Link to dataset:
  + <https://www.kaggle.com/datasets/mitchellweg1/nfl-combine-results-dataset-2000-2022>
  + <https://www.kaggle.com/datasets/ulrikthygepedersen/nfl-draft-1985-2015/data>

**Rationale for the Project:**

* **Project Advantages:**
* Informed Player Choices: Equips players with data-backed insights to make strategic decisions about their NFL aspirations.
* Proactive Talent Development: Supports players in identifying areas for improvement, thereby increasing their chances of success in the NFL.
* Optimized Team Selection: Enables NFL teams to make more informed choices during player drafts, leading to more competitive and successful rosters.
* **Performance Metrics:**

To assess the impact and effectiveness of the project, the following key performance indicators will be employed:

* Performance Metrics: Success will be evaluated based on the project's ability to establish significant correlations between player metrics, combine results, and eventual NFL performance. This analysis will drive targeted recommendations for players seeking to maximize their potential.
* Draft Success Rate: The project's effectiveness will be measured by tracking the success rate of players recommended by the model in actual NFL drafts. A higher rate of drafted players achieving successful careers will indicate the model's accuracy.
* Long-Term Success Stories: Tracking the careers of players who received recommendations and eventually made it into the NFL will serve as a testament to the project's effectiveness in identifying true potential.

**Project Outline:**

* Data cleaning **(Python)**
  + merge datasets
  + perform corr heatmap
  + delete highly correlated columns
  + edit nan’s and duplicates
  + Data exploration **(Tableau)**
    - Visualize distribution of target variables
    - .describe()/.summary() for each metric
    - df.hist(bins=10)
    - 40yd dash, shuttle, vertical, bench, broad jump, etc. over the years
    - find which teams had the most draft picks
    - which colleges had the most players in the draft
  + Scale data (use StandardScaler)
  + split data into test/train (2000 - 2014 is training, 2015 is test)
* Models to Use **(Python)**:
  + Decision Tree
  + Naive Bayes
  + KNN
  + Neural net
  + Xgboost
* Analysis
  + Compare each model using MSE, MAE, Acc, Loss
  + pick the best one!