

## COE Project 3

Aidan Liu, Arabel Rachman

Professor: Joe Stubbs

### **1. Introduction and topic and/or problem statement – A short introduction and summary of the goals of the project**

For this project we are doing the data from Riot games regarding Teamfight Tactics (TFT). For a brief introduction for this game, this game is a more complex version of autochess, where players are in a lobby against 7 other players (8 counting the player). In each set of the game, there are unique champions (which can resemble pawns or bishops in normal chess) and traits each champion has. Each set contains a unique set of champions and traits. Each set is swapped out for a new set every 6 months or so as to not make the game stale. For the data that we are using we will be pulling from set 13. For a deeper understanding of how the game works, there are team compositions, item builds, and augments that also add more complexity into the game. This model will quantify each composition's overall performance, this means that given a certain comp, what place will it achieve on average given the star level, cost, traits of the resulting board. For example, a 2 star 5 cost champion, a very high cap addition to a team, will on average boost the placement of the player significantly.

### **2. Data sources that will be used – A reference to any data sets utilized in the project**

As Teamfight Tactics is a game with constant updates, where each new set introduces completely different units and traits, we had to use the most up-to-date data for the specific set we are analyzing (Set 13). Since we could not find any thorough datasets for this set, we opted to get data directly stored by Riot Games using the RiotAPI. The RiotAPI allows us to have detailed and robust match data to train our models on. We created the dataset by making calls to the API, creating a dataframe out of the match data, and exporting it as a .CSV file. For each match, the data consists of the placement, the player level, the gold left available, the match length, the activated traits,

and the player's units along with their levels. A sample of this data can be found within the .CSV file attached to the Github repository.

### **3. List of high-level methods, techniques and/or technologies that you are considering using.**

For this project, after cleaning up data from separate game accounts totaling to over 5000+ games, 40000+ units, and 35000+ traits, we will evaluate five key machine learning models and their results: Test RMSE, Random Forest RMSE, XGBoost RMSE, CatBoost Test RMSE, and Transformer Test RMSE. In order to predict the dependent variable (the final placement), we are thinking of training a neural network; the layers used will depend on the features being fed. We could also use clustering techniques to identify compositions that are similar to each other and visualize that. Lastly, we will use exploratory data analysis to identify the most winning traits and the best units to play with them.

### **4. Products to be delivered – what are the primary deliverables for the project?**

#### **This is what we will be grading**

First will be a cleaned up CSV file that is mostly accurate containing the TFT data (units, win rates, team comps, traits, etc.). This CSV will be read from our python script which has been sculpted to fit the CSV format that we had to download from Riot API. This included cleaning up Champion names, Augment names, and Trait names for over 100+ champions and traits.

Secondly we will have a trained model that can predict the placement of a composition. This model will be our best model that has the highest statistical probability of accuracy. We will test and hone this model to be the most accurate out of everything we have tried. Thirdly, we will include our Jupyter Notebook folder, this will include our previous trials, mistakes, and successes. This will include many different models ranging from a simple Dense ANN to TabTransformers to Gradient Trees. This will also include our trials on ETL, EDA, model training, and overall evaluation. Fourthly, we will be delivering a concise PDF report that summarizes the Project requirement sections.