**INFSCI 2160 DATA MINING**

**FINAL PROJECT PROPOSAL**

1. Title: PUBG Finish Placement Prediction

Team members:

Jing Pang jip45@pitt.edu

Tian Xue tix20@pitt.edu

Yanbing Yang yay76@pitt.edu

1. Problems you have chosen:

For the final project, we have selected the PUBG finish placement prediction problem on Kaggle. This project aims to try to predict the final placement of a player/squad in percentage.

1. Why the problem is interesting to you:

We are interested in this project because this game is getting more and more spotlight around the world; it is close to our life experience. Many conditions may affect a result of a single round, and even more complicated for the final placement. We are quite curious about what can make a player or a squad get a higher ranking at the end of the game. So we chose this topic and tried to make some data mining analysis on it. The results of our study considered as an excellent reference for the professional esports teams and common players for fun. Players can try to apply the features in our results to improve their final placement in a single game.

1. Approach/general project plan you intend to take:

For this project, we will try to implement our dataset in programming R. We will use what we learned from the class to analyze the data and built a prediction model based on current information. In the beginning, we will explore the data at a statistics view, and prepare the data in cognitive status. In this project, we will try several algorithms, including Random Forest, XGBoost and so on, to implement with the data.

1. What kind of data you plan to use:

We plan to use the data from Kaggle that contains 28 variables that describe PUBG game stats. Each row includes one player’s post-game stats indicated by its player ID. The data comes from matches of all types: solos, duos, squads, and custom; there is no guarantee of there being 100 players per match, nor at most four players per group. A PUBG that has up to 100 players start in each match. Players can be on teams which get ranked at the end of the game based on how many other teams are still alive when they are eliminated.