CS 484 – Data Mining
Jacob Boisseau and Andrea Perez Isla
Final Project

PROJECT PROPOSAL

Video games have become more important over the years thanks to the advancement of technology and graphics that makes them more realistic and playable. In terms of sales, they even surpass movies, television and music. Focusing on PC games, battle-royale style videogames stand out this year, for example, Playerunknown's Battlegrounds (PUBG). It starts with 100 players dropped onto an island, where they have to survive and eliminate other players with different weapons while the play land shrinks progressively.

The main objective of the project is to predict in which position is going to finish a player in a round as a percentage (1 for first place, 0 for last). We have available the data divided in two parts, training set and test set, and a sample of the submission at kaggle's website.

For this purpose, we have this trainset we have mentioned before from Kaggle with more than 4.4 million instances and 29 features (for instance, weapons acquired or rank, among others). We will be dealing with a few challenges, at first sight. The most important is that we will be working with a very large data size.

To make the prediction, we are planning to use KNN and clustering or similar, depending on the accuracy of the final output. Furthermore, to evaluate the method used we can split the training set to see how precise are the results. For this reason, we need to find out which features are really significant and how they play into getting the first play at a game.