Predicting Competitive Ranking from Individual Statistics in Bungie's *Destiny 2*

Jackie Li, Jonathan Lu
UC Berkeley Data Science 100
Spring 2019

I. Abstract

This project looks to explore the relationship between personal performance statistics and the competitive ranking system "Glory" in the online multiplayer first-person shooter *Destiny 2*, developed by Bungie. 1601 players were selected through convenience sampling by pulling unique players from an author's past 200 of casual player vs. player matches. Our focus is to find particular personal statistics, if any, that could be used to predict a player's Glory. After validating a relationship between these statistics and Glory, we found higher personal statistics meant a player was *more likely* to be higher Glory, but no particular statistic or combination of statistics served as a better predictor than a regression model using every available personal statistic as features. We conclude that personal statistics from casual game modes cannot serve as good predictors of Glory, but that the aforementioned statistics can be used to predict more general trends in a players' Glory placements.

II. Introduction

Destiny 2, like many other first-person shooters, has a player vs. player (PvP) mode called the Crucible. The Crucible has casual game modes including a playlist called Quickplay (QP) as well as a competitive playlist. In QP, there are two teams composed of 6 players each; in competitive, there are two teams composed of 4 players each. While Quickplay does not use skill-based matchmaking (SBMM) to populate lobbies, competitive uses a skill ranking system called Glory to match players against each other. Glory ranges from 0 to 5500, and contains 6 ranks: Guardian (0 - 199), Brave (200 - 1049), Heroic (1050 - 2099), Fabled (2100 - 3499), and Mythic (3500 - 5500). Each player begins with 0 Glory and gain Glory by winning a match and lose Glory by losing one. The amount of Glory gained or lost is fixed for each rank, with gains decreasing as one moves up in ranks. In addition, more Glory is gained after consecutive wins, reaching a maximum at 5 wins in a row.

Ever since the introduction of the competitive playlist and the Glory system in early 2018, many players have expressed concern and dissatisfaction over its effectiveness. Unlike Elo-based systems such as Microsoft's TrueSkill, Glory does not take into account rank differences between players or even teams for matches. Because of this, a team with an average Glory in the low Mythic range (~3500) that loses to a team in the high Mythic range (~5400) would lose the same amount of Glory per player as they would had they lost to a team that is around their own Glory level. The flaw in this system is obvious, but we felt that Glory was still a valid enough metric for player skill to see a trend in player statistics for each rank. More specifically, we hypothesized that it is possible to use a player's personal performance metrics to predict their Glory rank.

Bungie, the developers of *Destiny 2*, tracks detailed individual statistics for players and offers them publicly in a highly detailed API. Available metrics include player activity histories, weapons and gear used, and most importantly for this project, personal performance metrics. In particular, we focused on a few metrics. Our reasoning for focusing on these particular metrics is due to the nature of the competitive playlist vs. Quickplay. Because Quickplay is not matchmade by skill, i.e. game lobbies are not populated based on proximity in skill level, a highly skilled player is much more likely to do well in game compared to a below average or average player. Assuming that Glory is an accurate reflection of player skill, it follows that a player with high Glory is more likely to do well than a player with lower Glory. We felt the following metrics were likely to be associated with either play skill or effectiveness:

- **Total Kills**: the total number of enemy players killed over a player's lifetime. A player with a higher kill count is more likely to have spent more time in the Crucible.
- **Kill/Death Ratio (KDR)**: the ratio of a player's total kills divided by their total deaths. This is an objective metric for player effectiveness.
- **Efficiency:** the ratio of a player's (kills + assists) divided by their number of deaths. An assist is when a player helps another player defeat an enemy but not land the final blow.
- **Average Score Per Life**: the average score earned by a player each life before they are killed. A score is assigned for player action that is beneficial for their team. For example, killing an enemy player, capturing an objective, denying an enemy an objective all have scores associated with them.
- **Score Per Game:** the average score earned by a player each game.
- **Orbs Dropped Per Game:** the average amount of orbs dropped by a player per game. Orbs of Light are spawned by players when they kill enemies using their Super ability or an upgraded weapon.
- **Win/Loss ratio:** the ratio of a player's total wins divided by their total losses.
- **Precision Kills Per Game:** the average number of precision kills by a player each game. A precision kill is a kill on an enemy player where the final blow is a headshot. As the name suggests, these kills require more precision.
- **Combat Rating:** an internal metric used by Bungie to reflect a player's overall skill.

III. Data

We collected a large sample of players by first finding the activity summaries of an author's last 200 games played. From these activity summaries, we collected IDs of every unique player. We could not collect a truly random sample because a list of players populating the Crucible is not available through Bungie's API. While our sample is not a truly random sample from the player population, we felt it would be representative of the population, as matchmaking in the quickplay game mode is dependent only on connection, with no element of skill taken into account. Therefore, our sample is a random sample of players with a reasonable connection to the Western United States, which is likely representative of a full random sample. A total of 1601 players was sampled through this method.

To collect a large sample of players' stats, we made API requests by players' unique identifiers and collected them in a dataframe indexed by the player's unique ID. We additionally found players' Glory by querying the player's progression stats from the API and collecting the player's Glory from the response. More specifically, for each player in our collection of player IDs, we queried the player's Glory as a character progression, and then queried their stats if they have a non-zero Glory rank. If their progression data was unavailable (indicative of a private profile) or they had a Glory of 0 (indicative of not having a competitive rank), we removed the player from our sample.

Our unfiltered set of features is the set of Quickplay stats available through the Bungie API through the GetHistoricalStats endpoint. These stats include both quantity stats and efficiency stats (per-game averages, as denoted by the suffix "_pga". Our final set of features is the original set of features with stats that return 0 for all players (due to being inapplicable to quickplay game modes) removed. 1180 players remained in the sample after cleaning.

IV. Method

We visually explored the relationships between the individual metrics we felt may be specifically associated with Glory and confirmed that there is a strong association between them. Using this information, we tried numerous combinations of features (including only efficiency stats and only career quantity stats), but we found that the best results were generally found by including all non-zero features.

To build our models to predict Glory and Glory-related categorizations, we used the Scikit-learn library to build several models.

First, we utilized linear regression to attempt to predict players' specific Glory ranks, training the model on a training set of data, and testing it on an independent testing set.

We then posed a classification problem, converting a player's Glory rank to their rank bracket as defined by the game. We investigated our ability to predict this by training a logistic regression classifier and testing its accuracy.

We then distilled our classification problem to a binary classification. Because the only tangible rewards to the Glory system are obtained at Fabled (aside from one exception obtained at 5450 Glory, nearly the maximum rank), many users reach Fabled and choose to stop playing the competitive mode, as they've obtained all the rewards realistically attainable to them. Therefore, this segments the competitive playerbase into two categories - players that have not yet reached Fabled, and those that have. We used these categories as our classifications, and trained both a logistic regression model and decision tree classifier based on these classifications.

V. Summary of Results

Fig 1. Accuracy of Classification Models on Random Test Sets

It is apparent that there is an association between each metric and Glory, with data points with higher metrics noticeably trending towards higher Glory. Out of the nine metrics we explored, Combat Rating and Score Per Game appeared to have the strongest association with Glory.

Despite using different combinations of features such as only efficiency statistics, only quantity statistics, only nine features mentioned previously, we found that using all features available through the API yield the best results for a prediction model.

The linear regression model that predicts a player's numerical Glory rating yielded a root mean squared error of about 1100. After classification into 5 categorical ranks, a logistic regression model yielded an accuracy of about 49%. Finally, a logistic regression model after binary classification into Fabled/Not Fabled yielded an accuracy of 83%, by far the best result.

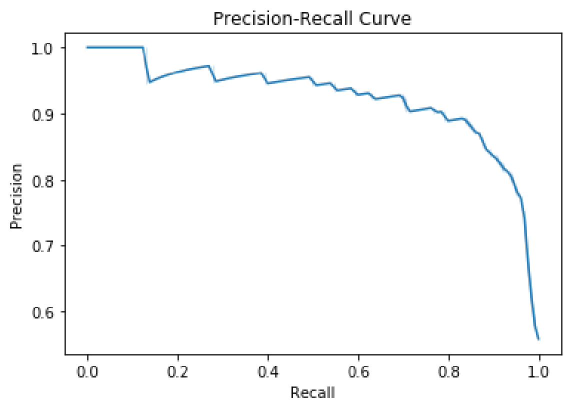

Fig 2. Precision-Recall Curve of Binary Logistic Regression Classifier

VI. Discussion

There are a number of factors that we felt could be negatively affecting our results. An assumption we made to test our hypothesis was that Glory was a valid metric to reflect player skill. Many factors complicate this assumption, the largest of which is its inherent flaw of being blind to Glory differences between teams. Most popular skill ranking algorithms for video games are based on existing ranking systems such as Elo, Glicko-2, and TrueSkill, and therefore take into account differences in skill between two teams when adjusting their rankings based on results, while Glory does not. While the competitive matchmaker attempts to match and balance teams by Glory, this is only possible to limited degrees, so matched teams may be mismatched, and a lower Glory team can lose the same number of points when losing to a high Glory team as they do when they lose to a team of similar Glory to themselves.

Additionally, the competitive mode is often regarded as unpopular among *Destiny 2* players, so we felt that players in the population may not have been motivated to spend the time to attain a Glory rank proportional to their actual skill level. Our binary classification partially controlled for that by separating players into two categories, only one of which had attained the primary rank players are incentivized to reach via tangible in-game rewards. Unsurprisingly, that model's predictions were much more accurate.

Finally, players have recently been known to utilize a matchmaking exploit to match into a game with an empty enemy team, providing free ranking points. This adds significant outliers that interfere with and circumvent our model, as players with low stats that cheat in this way can reach high Glory ranks that do not remotely represent their skill.

VII. Conclusion

Considering results from our three different prediction models, we cannot confirm our hypothesis that statistics from casual game modes can be used to accurately predict Glory. They can be used, however, to relatively accurately predict and classify players into two general categories, below Fabled and above Fabled. Flaws in Bungie's competitive matchmaking and ranking systems contributed to these results, but it leaves open the possibility for Bungie to improve upon them in the future.