Electronic Sports: Winner Prediction

Hao Ding

Department of Statistics

University of Connecticut

October 31, 2023

Abstract

Here is the abstract.

1 Introduction

PUBG (PlayerUnknown's Battlegrounds). PUBG is a popular online multiplayer battle royale game developed and published by PUBG Corporation, a subsidiary of South Korean video game company Bluehole. It was inspired by the Japanese film "Battle Royale" and released in March 2017.

In PUBG, up to 100 players parachute onto an island and scavenge for weapons and equipment to eliminate others while avoiding getting eliminated themselves. The game features a shrinking play area that forces players into closer proximity, creating intense encounters. The last player or team standing wins the match.

PUBG gained immense popularity shortly after its release, contributing to the rise of the battle royale genre in gaming. The game is available on various platforms, including PC, Xbox, PlayStation, and mobile devices. PUBG Corporation has also organized esports events, such as the PUBG Global Championship, where professional players and teams compete for significant prizes

Our goal is to analyze this dataset to figure out which specific factors or behaviors lead to more kills or longer game life, which leads to player wins.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

To cite a reference, here are examples. Xie (2015) did something ... Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

A lot of work has been done (e.g., Xie, 2015). Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris. Some parametric bootstrap sample size approach was proposed by Dwivedi et al. (2017).

The rest of the paper is organized as follows. The data will be presented in Section 2. The methods are described in Section 3. The results are reported in Section 4. A discussion concludes in Section 5.

2 Data

Data from the PUBG Developer API, it is a official game data. There are over 65,000 games recorded in this database.

Assists - Number of enemy players this player damaged that were killed by teammates.

Boosts - Number of boost items used. Revives - Number of times this player revived team-mates. Heals - Number of healing items used. Damage Dealt - Total damage dealt. Note: Self inflicted damage is subtracted. TeamKills - Number of times this player killed a team-mate. Kill Place - Ranking in match of number of enemy players killed. kills - Number of enemy players killed. Ride Distance - Total distance traveled in vehicles measured in meters. Swim Distance - Total distance traveled by swimming measured in meters. Walk Distance - Total distance traveled on foot measured in meters. Weapons Acquired - Number of weapons picked up. Match Duration - Duration of match in seconds. Win Place Perc - This is a percentile winning placement, where 1 corresponds to 1st place, and 0 corresponds to last place in the match.

3 Methods

Like we said before, correlation analysis, Feature importance from tree-based modes and recursive feature elimination are main method to select data to reduce the dataset size. Then, we split the data with training and testing sets. Choose appropriate machine learning models. Given that this is a predictive task with multiple features, ensemble methods like Random Forest or Gradient Boosting might be effective. Additionally, deep learning models like neural networks could be explored. we should train the chosen model using the training set.

4 Results

Table 1 summarizes some example draws from some distributions. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie

Table 1: This is my first table.

normal	Poisson	gamma
-0.110	4	2.401
0.116	4	3.529
-0.828	9	2.112
-0.066	6	11.104
0.219	3	4.815
0.303	5	2.188
0.544	0	8.050
-2.617	8	3.646
0.747	1	5.178
-1.103	4	3.043

vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Figure 2 shows Correlation between each factors so that we can select the important factor.

5 Discussion

We can choose several models, and compared with the accuracy of those models, and we will choose the highest accuracy model as our final model to predict whether the players or teams can win. So, we need to consider more detail for different modeling methods that will be lots of work. Second point is ensure that we have a sufficiently large and diverse dataset training a robust model. Since this is a video game, PUBG is a dynamic game with constant updates and changes. The model might need periodic updates to adapt to the evolving game environment.

If it is possible, we should work closely with PUBG experts to gain insights into the specific dynamics of the game that may not be immediately evident from the data.

Also we should emphasize a important factor, while machine learning models can provide valuable insights, predicting the outcome of a PUBG involves various unpredictable factors, including player strategies, teamwork, and unexpected events during matches. Always inter-
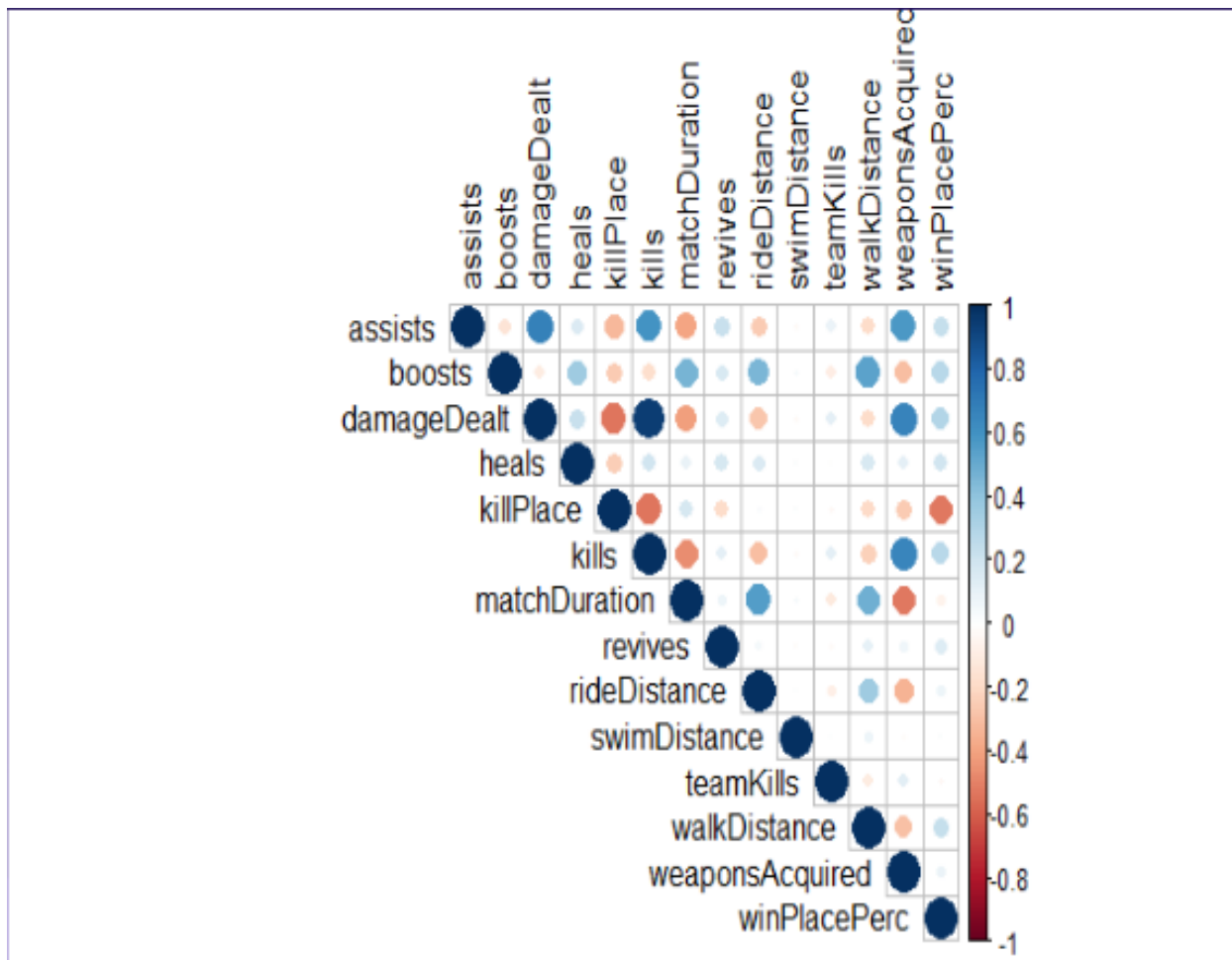
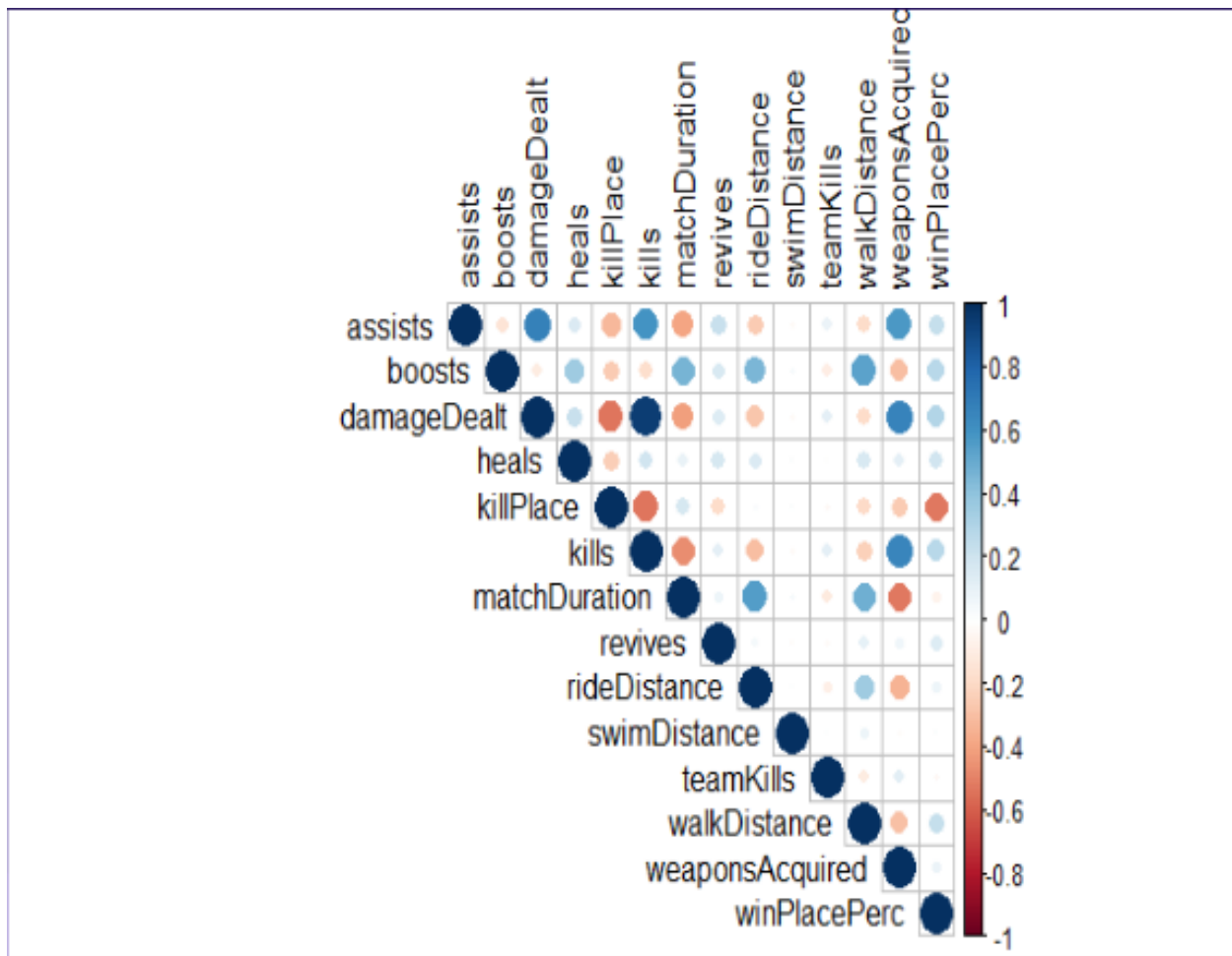Figure 1: The Correlation between each Factor

Figure 2: The Correlation between each Factor

pret the model's predictions cautiously and use them as a supplementary tool rather than a definitive source for predicting winners.

References

A. K. Dwivedi, I. Mallawaarachchi and L. A. Alvarado (2017). "Analysis of small sample size studies using nonparametric bootstrap test with pooled resampling method." *Statistics in Medicine* **36**, 2187–2205.

S. Wild, G. Roglic, A. Green, R. Sicree and H. King (2004). "Global prevalence of diabetes: Estimates for the year 2000 and projections for 2030." *Diabetes Care* **27**, 1047–1053.

Y. Xie (2015). *Dynamic Documents with R and knitr*. Boca Raton, Florida: Chapman and Hall/CRC, 2nd edition. ISBN 978-1498716963.