Using Apache Spark to predict finish placement in PlayerUnknown's Battlegrounds game

Pathan Faisal Khan, Letterkenny Institute of Technology

Abstract—

Index Terms—

I. INTRODUCTION

PLAYERUNKNOWN'S Battlegrounds (PUBG) [1] is an online shooter multiplayer game of battle royale genre [2] inspired by the Japanese film "Battle Royale" [3]. The game is playable on Personal Computers (PC), Sony's PlayStation, Micorsoft's Xbox and mobile devices (Android and iOS). It was developed and published in late 2017 by South Korean video game company Bluehole's subsidiary PUBG Corporation. Since it's launch, PUBG has seen exponential growth. It crossed the 1 million players mark within 48 hours of its release on consoles and a total of 30 million copies were sold for both PC and consoles just within a few days of its release. It has gained quite a good traction in Asian countries especially in China and India on to its mobile version with India standing at 116 million downloads which makes 21% of its total worldwide downloads followed by China with 108 million downloads which accounted for 19% while the USA stood at 8% with 42 million downloads [4]. With these many players on its platform, it generated vast amounts of data.

The game's most popular mode, classic mode, allows up to 100 players to parachute on an area of up to 8 x 8 kilometers island on which players will fight with each other in teams of maximum 4 players, the team to survive at the last is winner. The game world depending on the map selected has different terrains with sea, rivers, deserts, forests, snow, roads and bridges. Buildings and towns are spread across the map which has weapons and other essential items for a player to fight and survive in the game.

This technical report is going to explore a publicly available dataset [5] of the popular game. At the time of writing this report, the dataset is hosted on a popular online community for data scientists and machine learning practitioners, Kaggle [6]. The code for this report has been open-sourced and hosted on Github [7].

The primary goal of this technical report is predicting percentile of a player's finishing position where 0 denotes last position and 1 stands for 1st position in a match. The report has also answered the following questions with visualizations based on analysis of the data:

1) Average time a player spends in a match.

2) Total number of kills done by users when playing alone versus when playing in a team.
3) How better a player performs (in terms of rankPoints, winPlace, winPoints) when playing alone versus when playing in a team?

Organization of the report is formatted in the following way. The paper will start with a brief overview of the program paradigm in Section **II**. Later on, the algorithms used for prediction will be covered in Section **III**. The implementation plan which describes the data pre-processing stage as well as model creation will be discussed in Section **IV**. Section **V** will discuss the observations of the analysis answering the questions mentioned earlier. This section will also compare the used algorithms. The paper will finally conclude with the observations from this report in Section **VI**.

II. SYSTEM OVERVIEW

III. ALGORITHMS

IV. IMPLEMENTATION

V. EXPERIMENTATION RESULTS AND DISCUSSIONS

VI. CONCLUSION

REFERENCES

[1] PlayerUnknown's battleground- this is a battle. [Online]. Available: https://www.pubg.com/
[2] "Battle royale game," page Version ID: 930356754. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Battle_royale_game&oldid=930356754
[3] K. Fukasaku, "Battle royale," page Version ID: 931632941. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Battle_Royale_(film)&oldid=931632941
[4] A. McAloon. Now out of early access, battlegrounds crosses 30m players. [Online]. Available: https://bit.ly/2PbIlps
[5] PUBG finish placement prediction (kernels only). [Online]. Available: https://kaggle.com/c/pubg-finish-placement-prediction
[6] Kaggle: Your home for data science. [Online]. Available: https://www.kaggle.com/
[7] P. F. Khan, "faisal3325/pubg_prediction." [Online]. Available: https://github.com/faisal3325/pubg_prediction