

PUBG GAME DATA ANALYSIS AND PREDICTION

YANG CAO

ABSTRACT. With the rise of e-sports games, data analysis plays an increasingly important role. PUBG has taken the world by storm with its diverse playability. As a shooting game, players can move, supply, and shoot in the map to obtain higher terms. This paper uses different models to predict the final ranking based on different game data. Finally, the linear regression model was used to obtain 0.015303007019988265 test MSE.

Contents

Date: (None).

2020 *Mathematics Subject Classification.* Artificial Intelligence.

Key words and phrases. Machine Learning, data visualization, prediction.

Thanks to all members of TULIP...

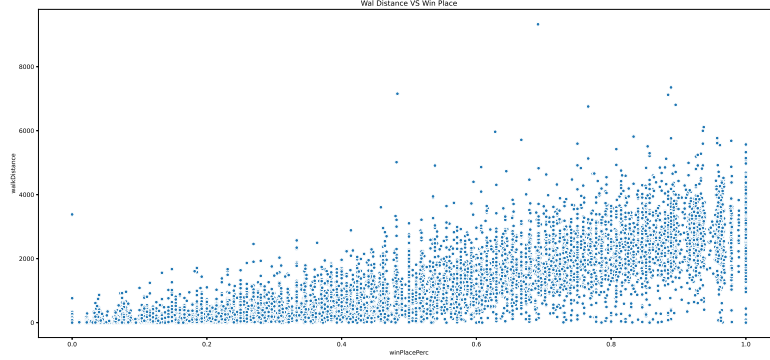


FIGURE 1. walking distance VS Win Place

1. INTRODUCTION

In a PUBG game, up to 100 players start in each match (matchId). Players can be on teams (groupId) which get ranked at the end of the game (winPlacePerc) based on how many other teams are still alive when they are eliminated. In game, players can pick up different munitions, revive downed-but-not-out (knocked) teammates, drive vehicles, swim, run, shoot, and experience all of the consequences – such as falling too far or running themselves over and eliminating themselves. Different game behaviors will lead to different final rankings, so the main purpose is to build a model to predicts players' finishing placement based on their final stats, on a scale from 1 (first place) to 0 (last place).

In this paper, we show that how to use the different game actions to predict the final win places. It can help game player to get higher rank and help game data analyst to get higher correlation actions to help preffessional player to get better win places in the competiton.

The remainder of this paper is structured as follows: In section 2, data visualizaiton will show the different game type proportion and relationship between two attributes. In Section 3, use linear regression and decision tree to build model and predict.

2. DATA VISUALIZAITON

- Figure 1 shows the different game type proportion of all matched games. Most players choose to play squad fpp and duo fpp. So if players want to play with more players.
- Figure 2 shows long walking distance always means higher rank because if a player move a lot and survive, he always can get a higher rank. For players, if want to get higher win places, they need to keep moving.



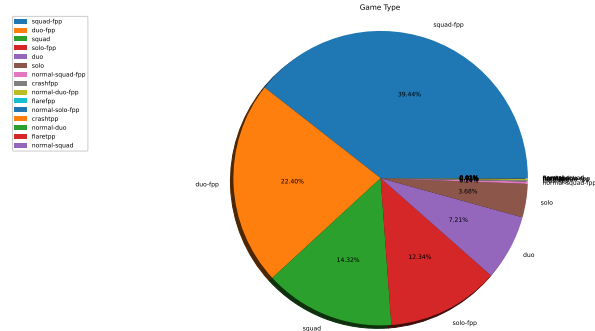


FIGURE 2. game type proportion

TABLE 1. Precision Comparison

	train MSE	test MSE
linear regression	0.01564124116618947	0.015303007019988265
decision tree	0.01859312767771217	0.17048691321544765

3. EXPERIMENT AND ANALYSIS

- Base on table 1, we can see that linear regression model get lower MSE value, it means that it has a higher accuracy and lower error value between predictions and actuals.

4. CONCLUSIONS

- Both training and testing data shows that linear model get lower mean square error value.
- Most players choose to play squad-fpp and duo-fpp
- More walking distance always can bring higher win place.

List of Todos

(A. 2) SCHOOL OF INFORMATION TECHNOLOGY, DEAKIN UNIVERSITY, GEELONG, VIC 3216,
AUSTRALIA

Email address, A. 2: Yang Cao@deakin.edu.au