Business Analytics Project Proposal

Group 7：Annan Chen/Jiepeng Lian/Jing Xiao/Qingyuan Ren/ Zequan Zhang

1. Introduction

In our project we want to analyse the matches data of League of Legends and try to explore value-added information from figures during the game to the outcome prediction.

League of Legends is a PC-side MOBA game that is popular all over the world. Similar to many MOBA games, players can accumulate levels and money by killing monsters, minions and opponent heroes in the game, and finally winning the game by destroying the enemy base. Players can get satisfaction through various operations in the game. The outcome of the game is the result that every player cares about. League of Legends, as the leader of this type of game, has also established its own unique game mechanism and system during years of development.

This game not only has a mass user base but also develops rich chains which contain tremendous business value. The professional competitions take place all over the world every year with a cornucopia of regional leagues including LCS, LEC, LCK, LPL, etc. The winners of the competition could receive up to millions of dollars. Besides, there are also various betting companies or live stream activities closely following the detailed results of each competition. The outstanding teams or players also attract various advertisement endorsement and commercial activities.Therefore, analysing the matches data to understand this game more and even predict the outcome at an early stage could create great value.

2. Data description

We have two candidates for the dataset that would be further evaluated in our project.

- The first datasets (https://oracleselixir.com/tools/downloads) include in-game features for all professional games during 2014 to 2020. Compared to the data in Kaggle, it contains individual level performance features like gold difference, experience, creep score, vision score, kill, assist, etc. The data also include features for the player that is in the same position on the enemy team.
- The second dataset - competitive match (Esport) (https://www.kaggle.com/chuckephron/leagueoflegends) is different from which in ranked matches. This dataset from Kaggle contains all competitive matches between 2015-2017, including NALCS, EULCS, LCK, LMS, and CBLoL leagues as well as the World Championship and Mid-Season Invitational tournaments. The detailed features match data are: Seasons, teams, gamelength, gold difference, kills, ban & pick (Might be the most important for the team's strategy), etc.
- Compared to the data source from *oracleselixir.com*, this dataset is easier to preprocess but includes less game seasons. Also, one drawback is that the dataset doesn't take LPL into account - which is the Pro League in Mainland China, also the winning Region for 2018 and 2019 World Championship.

3. Methods and Plans

With our dataset, we will apply some data visualization to see the statistical characteristics. Through this process, we get a sense of the data that may be helpful for our prediction. By analyzing the "gold" data, for example, we can obtain the gold data for each position of the two teams and compare how much the win team accumulated in a match and how much a lost team gained. Further analysis, like statistical hypothesis tests, can be done to check the significance to look for more connections.

We also need some preprocess. Like, transforming some variables to numeric and get one-hot vector. Also, since the dataset is large and there are lots of features, some dimensionality reduction tech can be done . We plan to use the PCA we learned from the class to deal with it and improve computation efficiency. We will also check other methods to see the results.

Currently, we plan to predict the result based on a logistic classification model. Certainly, other models will be tested if they tend to get a better performance.

Figure 1: A typical MOBA game map

Ø The terrain of the entire map is shown on the left, showing a certain symmetry. The lower and upper right are separated by the middle river.
Ø The diamond pattern in the picture is a crystal/base on one side. Destroying the opponent's crystal means winning.
Ø The 11 circular patterns in the picture are defensive towers. The crystal can only be destroyed after the defensive towers are destroyed.
Ø In the direction of the three arrows in the figure, it will automatically attack the opposite side according to the time interval. These three routes are called upper/middle/lower roads.
Ø The 3 square patterns in the picture are barracks. After destroying the opposing barracks, one side will be strengthened.

Ø The triangle in the figure represents the monster. The two big triangles in the river channel are special bosses known as Big/Little Dragon.
Ø In the double play game, 5 members each control different characters/heroes to move freely on the map. Different heroes have different characteristics and skills.
Ø Killing wild monsters, minions, enemy heroes, destroying enemy defense towers will increase the hero's experience
Experience and money. Upgrade can strengthen skills, and money can buy equipment.
Ø More powerful heroes can win the game more easily.

Appendix: https://linemovement.com/2020/07/09/how-to-bet-on-league-of-legends-from-lcs-to-lpl/