**League of Legends: The Importance of the First Fifteen Minutes**

Alan J Donahue

Bellevue University

DSC 550: Data Mining

Aaron Lee

June 4, 2022

**Introduction**

League of Legends (LoL) is a multiplayer online battle arena (MOBA) created by Riot Games in 2009. The game consists of two teams of five players that each play a character called a champion. As the game progresses, the players need to level up their champions by collecting gold and experience points and purchasing items. There are two sides of the map in a game and the point of the game is to destroy the enemy “base” by destroying their nexus.

League of Legends became so successful that Riot Games created an esports ecosystem for teams and player to compete. The very first professional tournaments started in 2011. Over the course of many years, Riot Games has created different leagues based on regions. There are now a total of 12 Tier One Professional leagues, which include any where between 8 to 17 teams in each league. Each league has a spring and summer split, which determines which teams from the league will compete at the international events.

Over the course of almost 11 years of esports international events, the North American region, League Championship Series (LCS), has never one an international event. Most wins have come from either the South Korean region, League of Legends Champions Korea (LCK), or the Chinese region, League of Legends Pro League (LPL).

Since teams from the LCS have continued to be a disappointment when it comes to international events, there is a growing concern with how the teams are utilizing their time during their training and scrimming (practicing against other teams). The LCK and LPL teams are known for taking their training very seriously. One of the scrim techniques they have is a mode they call “Blitz.” Blitz is when two teams face each other and only play the first 15 minutes of the game. After the first 15 minutes, they end the game, take notes, and go again, repeating this process over several hours. They conduct this training because they believe that the first 15 minutes of the game are vital to the outcome of the game.

Now, it is known that North American teams push back against such a scrim method. In fact, many Korean players and coaches have ended up moving to the LCS to play for North American teams. They have tried to implement such methods and usually end up leaving the team because the other players don’t want to conduct Blitz training.

This is a huge issue for LCS teams and players. The fans have been upset by their performances and have gone online to state they are lazy, don’t want to win, etc. On the business side of the house, the LCS teams and players’ brands are starting to take a hit for not wanting to enforce training like Blitz. Additionally, winning international events comes with a monetary prize as well. So, businesses and players are losing out on money because their brands are taking a hit, fans don’t want to follow them anymore for consistently losing, and they aren’t winning the prize pools from tournaments.

It is my hypothesis that I believe the first 15 minutes of a game can determine the outcome of the rest of the game. I would pitch this problem and solution to several different levels of stakeholders. First, there are CEOs of the teams who are heavily involved. I would pitch the problem in terms of funding and monetary gain for them. If it is true, training for the first 15 minutes can mean more wins and success in international events, which can lead to more sponsorships and fans. Next, I would pitch it to the coaches who still want to improve the region’s training. I would pitch it to them in terms of winning. Finally, I would pitch it to the players who are trying to make changes amongst their own teams. I would pitch it in terms of winning and brand opportunities.

I acquired my data set from the website of one of the leading data scientists for the LCS. He acquired the data from multiple sources including match history pages, the official League of Legends website, and the Riot Games API.

**Exploratory Data Analysis**

My first step in the project was to conduct exploratory data analysis (EDA). While conducting EDA, I went in with a few assumptions. My first assumption was that gold difference between two teams at the 15-minute marker is incredibly important to the game. When a team has a huge lead in gold, they can purchase more items which in turn means their champions are stronger than the other team’s champions. Below are the two graphs for gold difference at 15 minutes for the winners and the losers.

Figure 1:

Chart, histogram

Description automatically generated

Figure 2:

Chart, histogram

Description automatically generated

Looking at the two graphs, it is obvious that in most cases, winners tend to have a gold lead while losers tend to have a gold deficit. However, there are still many situations where the winners either don’t have a gold lead at 15 minutes or they have a gold deficit. Same goes for the losers. There are games where the losers are tied in gold, or they even have a gold lead. In my opinion, it doesn’t provide enough evidence to support the fact that the first 15 minutes will determine the rest of the game.

My next assumption was that depending on what side of the map a team was on could sway the outcome of the game. In LoL, there are two sides that a team could potentially start the game on: the red side or the blue side. Which side a team is on will determine the order for the picks and ban phase of the game where teams choose their champions and ban other champions. There are many coaches in different leagues who have blamed their teams lose on the side of the map the team starts on. Below is the pie chart that breaks down the wins between the red side and the blue side.

Figure 3:

Chart, pie chart

Description automatically generated

The pie chart above demonstrates that the side selection doesn’t influence the outcome of a game because it is almost a 50/50 split between the number of wins per side.

The next assumption I had was that the amount of turret plates a team knocks down can help influence the outcome of the game. In a game, each team has 10 turrets on the map. For the first 14 minutes, each turret has 5 turret plates which gives the opposite team gold when they destroy them. My assumption was that the more turret plates knocked down, the more gold a team has which can snowball into a lead. Below are the two graphs for destroyed turret plates split between the winners and the losers.

Figure 4:

Chart, histogram

Description automatically generated

Figure 5:

Chart, histogram

Description automatically generated

Looking at the two graphs, it looks like the team that destroys multiple turret plates have a greater chance of winning the game. I believe that this does potentially support the hypothesis that the first 15 minutes are important to the outcome of the game.

My final assumption I had while conducting EDA was that the kills, deaths, and assists (KDA) in the game will influence the outcome of the game. It makes sense that if a team has more kills and assists and less deaths, that team should be more likely to win the game. Below are the three kernel density estimation graphs covering KDA.

Figure 6:

A picture containing shape

Description automatically generated

Figure 7:

A picture containing histogram

Description automatically generated

Figure 8:

Logo

Description automatically generated

The graphs displayed show that the winning teams tend to have more kills and assists and less deaths than their opponents at the 15-minute mark. This supports the hypothesis that the first 15 minutes can determine the outcome of the game.

**Data Preparation**

After acquiring a base line knowledge of the data, I moved on to properly prepare the data. The first step in preparing the data for a model was to remove all the unnecessary columns in the data set. These columns, like url, year, date, etc., didn’t have relevance in determining if the first 15 minutes were important to the outcome of the game. Any of the columns labeled opp\_\* meant that it was the opponent's stats being tracked against that team. I removed those columns because that data is already reflected in their own respective team row.

The next step was to eliminate all the null values in the data set. First, I removed all columns that were only 100% null values as I couldn’t do anything with that data. Second, I noticed a pattern of nulls in some columns. In specific columns, there were a total of 1738 null values. To fix this, I removed the null values from one column, and it removed the rest of the nulls values from the other columns. Finally, there were only a few null values left over. I took a deep dive into the data to figure out where the null values were located. The null values fell under the different dragon columns and the turret plates column. I replaced the null values with zeros because that was what was meant to be there.

The final number of columns are displayed in the table below.

Table 1:

Graphical user interface

Description automatically generated with medium confidence

After handling the null values, I created dummy variables for the side column. Lastly, once everything was cleaned, I split the data into training and target sets with a train/target split ratio of 80%/20%.

**Model Building and Evaluation**

I chose the decision tree classifier as my model. For this project, I was less interested in predicting the outcome of a game and more interested in how important each variable was. Predicting the outcome of the game is not beneficial because the game is already finished by the time the data is recorded. There is benefit in knowing which variables are more important because that data can be used to refine a team’s focus. Additionally, from a business perspective, it is easier to get the message across to less tech savvy individuals using a decision tree.

I created multiple decision tree models: a regular decision tree, a decision tree with hyperparameter tuning, a decision tree with hyperparameter tuning that didn’t have towers or inhibitors in the model, and finally a decision tree with hyperparameter tuning that only included the variables that occur before the 15-minute mark.

Since I am interested more in the variables than the ability for the model to predict the outcome of a game, I used accuracy score as my metric for model evaluation.

**Conclusion**

In conclusion, the analysis and the model building shows there is potential that the first 15 minutes of a match can sway the outcome of the game. The decision tree classifier model (model A) with hyperparameter tuning had a 95.73% accuracy. The decision tree classifier model (model B) with only variables that occurred before the 15-minute mark had a 72.51% accuracy. The fact that using only the first 15 minutes of data can predict with 72% accuracy the outcome of a game definitely supports the hypothesis.

These models wouldn’t be deployed in the traditional sense. It wouldn’t be added to a pipeline to predict outcomes. Instead, it would be used to present the data to the CEOs, coaches, and players. In this case, the models are ready to be presented to use as supporting evidence that the first 15 minutes of a match are important to the outcome of the game.

Based off the models and analysis, it is my recommendation that the LCS teams need to start incorporating Blitz mode into their training. I believe teams that will do this will win more times than not and potentially have international success. This in turn will boost the brands of the players, coaches, and teams, which will of course, bring a monetary boost to them all.

Outside of the current models, I think exploring the top 5 most important aspects of a game, not just the first 15 minutes, would be incredibly beneficial to the LCS. This information would again help teams tailor their practices to be more beneficial.

# References

*List of League of Legends leagues and tournaments*. (2022, June 4). Retrieved from Wikipedia: https://en.wikipedia.org/wiki/List\_of\_League\_of\_Legends\_leagues\_and\_tournaments

*Portal: Tournaments/International*. (2022, June 4). Retrieved from LoL Fandom: https://lol.fandom.com/wiki/Portal:Tournaments/International

Sevenhuysen, T. (2022, June 4). *Match Data Downloads.* Retrieved from Oracles Elixir: https://oracleselixir.com/tools/downloads