Abstract:

League of legends is the most played video game in the world. Every match consists of two opposing teams (blue and red) each with 5 players that try to destroy the enemy’s base (or nexus). Typically, matches last between 25-35 minutes. In addition, each game is extremely unique; there are a multitude of factors that can influence the outcome of a game. These include player mentality, objectives around the map, team compositions, communication and many more. With the increasing accessibility of data analytics tools, it is now possible to make predictions on which team is more likely to win a game. This project will be focusing on predicting the outcome of a match based on a list of numerical attributes extracted from approximately 10,000 games. The goal is to build and compare 4 different models based on the following algorithms: logistic regression, Naives-Bayes, random forest and knn neighbours. Afterwards, we will be looking at which attributes are most important when it comes to winning games and how to optimize the models we built. By the end of this project, we will have a deeper understanding of the factors that can influence the outcome of a game and will be able to recommend which decisions are the best to maximize the chances of winning for a team.

1. Describing the dataset

The dataset being used can be found at <https://www.kaggle.com/bobbyscience/league-of-legends-diamond-ranked-games-10-min>.

It consists of 38 features (19 per team) alongside the column blueWins which is our target value for building our models. It also has a column gameID showing that each game is unique.   
These attributes were measured at the 10-minute mark of each game, also considered the early game. Consequentially, the models we build will provide insights into which decisions to make prior to this point in the game. They will also give recommendations on which factors are the most important in the early game and how to maximize chances of winning.

Here is a quick summary of each attribute present within the dataset. I will only explain what they represent for a single team (the blue team) to avoid redundance. This is because these attributes mean the same for the blue team as they do for the red team.

The attributes can be divided into the following main categories:

-gameID: the unique ID for each game

-blueWins: the target value, 1 if the blue team won the game and 0 if the red team won the game

-Wards (includes blueWardsPlaced, blueWardsDestroyed): wards are trinkets placed around the map to grant information on where the enemy team is situated. A team can place a ward in an area, and it will grant them information around the area. The opposing team can then destroy the ward to remove this ability. blueWardsPlaced is the number of wards placed by the blue team and blueWardsDestroyed is the number of red wards destroyed by the blue team.

-Kills/Assists (includes blueFirstBlood, blueKills, blueAssists): the number of kills and assists for the blue team. FirstBlood represents the first kill in a game: blueFirstBlood is equal to 1 if the blue team got the first kill and 0 if the red team did.

-blueDeaths: the number of deaths for a team

-Elite Monsters (includes blueEliteMonsters, blueDragons, blueHeralds): dragons and the Herald are two elite monsters present within the game. blueEliteMonsters is the count of how many elite monsters were defeated. If the blue team has defeated a dragon and a herald, blueEliteMonsters would be 2.

-blueTowersDestroyed: this is the number of towers destroyed by the blue team. If the red team has lost 2 towers, it is because the blue team has destroyed them and, in this case, blueTowersDestroyed would be 2.

- blueTotalGold: the total amount of gold gathered by the blue team. Gold is the main currency in the game and is used to buy items in order to increase the strength of the player.

-Level (includes blueAvgLevel, blueTotalExperience): the average level of each member within the blue team as well as the total experience by the blue team. Experience is used to learn new gameplay prospects and advance the development of the game.

-Creep Score CS (includes blueTotalMinionsKilled, blueTotalJungleMinionsKilled): minions (or creeps) are units present within the game that grant gold when defeated. They are the primary source of income and are found everywhere around the map.

-blueGoldDiff, blueExperienceDiff: these are difference metrics, but we will be creating our own difference metrics so these ones will be removed from the dataset

-blueCSPerMin, blueGoldPerMin: like above, these will be dropped from the dataset.

1. Explaining the constraints on the dataset

When doing the exploratory data analysis and finding the correlation matrix, we notice that a lot of attributes have very strong positive and negative correlations. For example: blueKills is perfectly inversely correlated to redDeaths. This is because the dataset is symmetrical: it contains data for both blue and red team for the same game. If a blue team member gets a kill on the red team, the blueKills counter goes up by 1 and so does the redDeaths counter.   
To run most predictive classification models, high correlations are not beneficial at all. They lead to high variance and end up obscuring the weights attributed to each attribute. Usually, we keep one of the two highly correlated features. To fix this, we can simply create a list of new metrics that take the difference between red and blue team. For example, instead of including blueWardsPlace and redWardsPlaced, we introduce a new metric called wardDiff that represents the difference between both prior attributes.

Other attributes we can easily remove are blueCSPerMin and blueGoldPerMin. blueCSPerMin is equal to the creep score divided by minutes and blueGoldPerMin is equal to blueTotalGold divided by minutes. Since this information is already present in the dataset, we have no need for these duplicate attributes.

We are also removing the blueEliteMonsters attribute from the dataset as it is simply the sum of blueDragons and blueHeralds. Both dragons and Heralds are elite monsters within the game and provide different team bonuses.

The dataset has no missing values, but there are some outliers present for certain attributes. As an example, let’s take a closer look at the shape of the data for the attribute 'blueWardsPlaced'. First, we look for outliers by creating a box whisper.

A picture containing table

Description automatically generated

There are some outliers present past 200+ blue wards placed. Let’s take a further look and see how the data is distributed.

Chart

Description automatically generated

We notice that even past 100 there is very sparse distribution.

When put into context it becomes apparent why these outliers need to be removed. In LoL, it is practically impossible for a team playing normally to have even 100 vision score by the 10-minute mark. These outliers stem from situations where a player will completely give up on the current game and start placing an absurd number of wards. They do this to express not caring anymore. This also happens with another attribute: blueDeaths. These players will die as many times as possible, and the resulting data will end up with an underlying bias.   
Furthermore, A player that spams wards or dies repeatedly is most likely in a game already lost. This is because these actions are considered toxic, and toxicity has a direct negative impact on team performance. In the Literature Review section of this project, I refer to a paper titled *Effects of individual toxic behavior on team performance in League of Legends* (Monge and O’brien, 2021)that confirms this.

Because of this, it is valid to remove these outliers to avoid introducing bias in the data.

Using python, we can figure out how many games are removed from the dataset that have over 100 wards placed.

A screenshot of a computer

Description automatically generated with medium confidence

108 games in total, meaning we only lose 1.1% of the total games.

This method will also be done for the games with abnormally high values in blueDeaths.

Finally, we make sure that the distribution of the class values is balanced. We create the following histogram to confirm that there is indeed nearly the same number of wins and losses.

Chart, bar chart

Description automatically generated

1. Overall methodology

Here is a tentative list depicting the proposed methodology. It is very crude and subject to change, but nevertheless presents a good foundation for the project.

1) Exploratory data analysis (find outliers, check for missing data, identify correlations, etc.)

2) Adjusting the dataset to prepare it for the model (creating the diff attributes, removing redundant attributes…)

3) Data Standardization

4) Constructing the models (one for each of the following: LR, Naives-Bayes, Random Forest and KNN neighbours)

5) Analysis of the models

6) Adjusting the models based on the analysis (which features are the most important, how to increase accuracy/precision/recall etc.)

7) Extracting insights from the adjusted models

8) Conclusion

The code for this project will be posted in the following repository:

<https://github.com/JeanLeclerc1/CIND820Capstone.git>

1. Literature review:

The literature revolving LoL mostly falls into two main classes:

Analytics: these revolve around insights gathered from quantitative, numerical data often pulled from the LoL API. This project is part of this branch.

Behavior Studies: these aim at understanding the behavior of players and how it affects the game environment.

<https://www-sciencedirect-com.ezproxy.lib.ryerson.ca/science/article/pii/S0378873318301229>

Player-centric networks in League of Legends is a peer-reviewed paper that aims to find and define the different types of players that are present throughout League of Legends. It was published in 2018 on ScienceDirect, a trustworthy and peer-reviewed source for scientific research. Written by Marçal Mora-Cantallops and Miguel-ÁngelSicilia, it describes the creation of a social network that connects hundreds of thousands of players. Each relationship between 2 players is attributed a weight that is influenced by the number of times they ended up in a game together. Afterwards, indicators are defined. These indicators exist to help characterize the underlying structure and act as a basis for the clustering algorithm. To find the best indicators, the authors assume node relevancy and cohesion as the most relevant structural features within the dataset. More specifically, since League of Legends is a social game, the focus of the indicators was placed on how often a player would play within their core community (friends), the quality of the structure of the communities, the sizes of these communities and the differences found between them. Finally, they use an affinity propagation algorithm to cluster the resulting nodes and build the network.

Marçal Mora-Cantallops and Miguel-ÁngelSicilia ended up with four main clusters. The first cluster (C1) corresponds to the “team player”, or the individual that plays with the same group and this group plays together. C2 is the individual that plays often with a couple of friends, but also plays with other groups (although less frequently). C3 is nearly the same, however the player engages in solo matches instead of with other groups. C4 is the solo individual, or the one that likes to play most games alone.  
Even more interestingly, the authors then compared these clusters to the rank of the player. They learnt that the higher the rank of a player, the likelier it is that this player is a solo player or exhibits solo tendencies. On the contrary, the team and group players all had ranks that were relatively low. Here, the authors explain that there are two main factors contributing to this. Firstly, higher ranked games have many more restrictions in terms of who you are allowed to play with. Secondly, playing in groups and with friends is a more casual experience because there are no expectations to win. Thus, group players have lower ranks than solo ones.

This paper provides great context for our project. The dataset we have chosen was created based on high diamond ranked games. Diamond is the highest of the traditional ranks in League of Legends, preceded only by Master and Challenger which are special divisions reserved exclusively for the best players in a region. Thus, we can theorize that a large majority of the games in our dataset were games that opposed solo players (or C4) against each other. This is a useful insight because when it comes to predicting how to win a game, numerical attributes are not everything. As mentioned in the abstract, there are other factors that influence the outcome of a game. Each player is human, and thus there is also an entirely social aspect to winning. With this paper we can establish that a lot of the interactions between players that are part of the same team are interactions between complete strangers. However, we cannot make assumptions on the quality and substance of these interactions and how they affect game outcomes. Evaluating player interaction and behavior is outside the scope of the project, but it remains important to mention this as a limitation of our model.

Marçal Mora-Cantallops and Miguel-ÁngelSicilia have also written other papers about League of Legends. In 2017 they had a journal entry titled “Exploring player experience in ranked League of Legends” published in Taylor&Francis Online, a reputable international publisher. This shows prior experience with the topic and further lends to their credibility.

<https://www-tandfonline-com.ezproxy.lib.ryerson.ca/doi/full/10.1080/15213269.2020.1868322>

This paper explores the relationship between toxic behaviour in League of Legends and team performance. It was published on January 7th, 2021, by C. K. Monge and T. C. O’Brien in Taylor & Francis Online, a website that offers peer-reviewed journals and articles.

In this study, the authors investigate if toxic individuals can have a negative impact on overall team performance in League of Legends. They introduce a randomly assigned team member called a “confederate” and have them act according to toxic behavior. Afterwards, a two-group multivariate analysis of variance (MANOVA) is performed using team metrics such as gold, objectives, etc. Using this method, Monge and O’brien validate that there is indeed a causal impact of negative behavior on overall team performance. When a player issues toxic communications, the rest of the team performs much worse.   
Furthermore, Monge and O’brien explain that while toxic behavior does bring down team performance, the other way around is also true. Other behavioral studies (Breuer et al., 2015; Kasumovic & Kuznekoff, 2015) have demonstrated that poor performance also brings toxic behavior. They describe a broader cyclical relationship between team performance and toxic behavior.

Even though studying player behavior is outside the scope of the project, this paper tells us that an ideal predictive model for winning would have to include human behavior.

<https://web-p-ebscohost-com.ezproxy.lib.ryerson.ca/ehost/detail/detail?vid=0&sid=031c1eb9-9a02-42fe-a8fd-fd0108d63034%40redis&bdata=JnNpdGU9ZWhvc3QtbGl2ZQ%3d%3d#AN=148210763&db=s3h>

When it comes to the quantitative side of the analysis surrounding Lol, the paper *Smart kills and worthless deaths: eSports analytics for League of Legends* shows us a great example of the benefits provided by defining new and advanced metrics for a predictive model. Written by Philip Z. Maymin in 2019. It was published on EBSCO, a highly regarded scholarly source that offers fully peer-reviewed journals. The author is a professor of analytics involved in sports analytics as well as finance algorithms. He has a Ph.D in Finance from the University of Chicago, a Master’s in Applied Math and a Bachelor’s in Computer Science both from Harvard.

In the paper, Maymin talks about traditional metrics for evaluating sports and compares them to a set of advanced metrics that he has defined. Using these “smart” attributes, he aims to answer the following question: how does individual performance relate to the team’s odds of winning? He starts by using computer vision on the LoL spectator client to extract various metrics from over 150 000 games. Afterwards, a logistic regression model is designed that predicts whether a team has won or lost the game. His model is based on minutes elapsed, kills for both (red and blue) teams, towers for both teams and monsters defeated for both teams. Some of these metrics are similar to the ones we use in our project, although we are concerning ourselves with the difference between both teams instead. Maymin goes further and defines a new set of metrics that he calls “smart” metrics. These include smart kills (kills that advance the game and increase the chance of winning), worthless deaths (deaths that had no real purpose and decreased the individual’s chances of winning), time management (map coverage) and many others. With these advanced metrics Maymin is able to create a decision tree that boasts an accuracy of 80% and a false positive rate of only 13%.

This paper is a great example of the effect that defining new and insightful metrics has on creating a successful model. Using basic attributes, Maymin is not able to predict the result of individual performance on the outcome of games. However, once he has defined a clear set of advanced metrics, he is able to predict the specific impact of an individual on the result of a game.

The author had access to very powerful resources and a lot of big data to build his model upon. The dataset we use for the project is much smaller and less detailed. If we were to carry over the main concept from Maymin and try to invent “smart” metrics in this situation, there wouldn’t be enough data in the first place. However, by using the differences between red-blue attributes as our metrics, one could call this a much more simplified version of the Maymin smart metric. Nevertheless, it will be interesting to compare models and find their differences.

Maymin also does a good job at highlighting the disconnect between traditional sports and eSports. Indeed, in eSports it is much easier to gather and evaluate data. There exists an incredibly large number of games and the data collection is much more accessible. This facilitates the implementation of machine learning algorithms and data analytic methods. Meanwhile, in traditional sports, data gathering is more difficult and sparser resulting in higher difficulty to create insightful models.

https://web-p-ebscohost-com.ezproxy.lib.ryerson.ca/ehost/detail/detail?vid=0&sid=b483cece-7848-436b-b648-605650370ad2%40redis&bdata=JnNpdGU9ZWhvc3QtbGl2ZQ%3d%3d#AN=148210763&db=s3h

This paper was published by Shinyoung Kim, Dohyeon Kim, HyungGeun Ahn and Byeongtae Ahn in an international journal titled *Multimedia Tools and Applications* in August of 2020. This journal is intended for academics who are involved in multimedia system research. It offers high quality peer-reviewed research. However, this research was done with the goal of creating a new service in mind. Thus, there is a product involved.

The paper highlights a need for an analytic tool that goes beyond the displaying of simple data. The authors mention that most analytics websites concerning LoL focus on showing data for immediate consumption and don’t grant deeper insights or further analysis. This is because these services rely only on data extracted from the public API. Consequently, the primary goal of the paper was to invent a tool that would excel at extracting raw data from various games and complementing it with the one from the API. To achieve this, they used a combination of video processing, data from the client and data from the public API.

However, they also build a model that tries to characterize the different playstyles of LoL players. Using knn and principal component analysis (PCA) they classify players as either cooperative or aggressive. Unfortunately, they were unable to correlate playstyle to in-game actions due to lack of research. Still, it remains noteworthy that they were able to make this distinction based on multiple sources of data.

The main takeaway from this paper is about assembling a more proper picture when tackling a problem. Here, the data comes from varying sources. However, when combined, this data can provide a more valuable insight than just the sum of its parts and doing so properly can result in a more complete solution to a problem. This is valuable to know for prospects regarding this project. We know that it is possible to add more layers to our model as we include new sources of data. If done accordingly, the models will be more accurate. This is an avenue for exploration down the line.

One thing to note is that in their literature review, the authors mention Maymin’s *Smart kills and worthless deaths: eSports analytics for League of Legends.* They claim he did not find a relation between individual and team performance. However, as we know, Maymin did find a relation but only by using his self-defined advanced metrics.

<https://ieee-cog.org/2021/assets/papers/paper_292.pdf>

*Feature Analysis to League of Legends Victory Prediction on the Picks and Bans Phase* is a paper written by Lincoln Magalhaes Costa, Rafael Gomes Mantovani, Francisco Carlos Monteiro Souza and Geraldo Xexeo. It was published August of 2021 in the 2021 IEEE Conference on Games (CoG). This conference is part of IEEE Xplore, a platform that showcases different academic works. IEEE CoG claim their papers as being double peer reviewed.

In this paper, the authors describe a predictive model that is given purely pre-game information. This includes the character (or champion) chosen by the player, their win rate percentage (WR), their number of games played, and various other metrics found outside the game. The goal of the model is to predict the outcome of the game based on the information available prior to starting. They find that the most important features are the ones related to previous game performance. Mainly win rate and champion KDA (kill/death ratio). This makes sense because features like these are compiled across many games. To calculate win rate, you need at least 2 games. Thus, these features act as a strong indicator for predicting match outcome based solely on pre-game information. Due to the binary state of this classification problem, the authors used Area Under Curve (AUC) as a performance metric for their model. They obtain 0.97 AUC using Random Forest and Linear Regression. This means the classifier is highly capable at distinguishing between a win or a loss.

This was the first study to validate that the pre-game stage of LoL games has an effect on winning. The difference between this study and our project is that our project takes data obtained during a game instead of prior to it, specifically the first 10 minutes. This means that Lincoln Costa’s model will do a good job at explaining the dynamics of champion select (also known as the pre-game stage). Meanwhile, our model will provide insights on the early stage of a game.

It is important to note that some attributes share similarities between both projects. For example, our project has features such as kills, deaths and assists which are all attributes that compose the KDA feature in Costa’s paper. The difference is that KDA is aggregated over many games and given as an average. Our features are also team-wise whereas in the paper they are only for individuals. Given these differences, I still expect our findings to be in line with the results of this paper. Mainly kills, assists and deaths being the most important attributes.

Graphical user interface, timeline

Description automatically generated

Mora-Cantallops, M., & Sicilia, M.A. (2018). Player-centric networks in League of Legends. *Social Networks, 55,* 149-159. doi: [https://doi.org/10.1016/j.socnet.2018.06.002](https://doi-org.ezproxy.lib.ryerson.ca/10.1016/j.socnet.2018.06.002)

Monge, C. K., & O’Brien, T.C. (2021). Effects of individual toxic behavior on team performance in League of Legends. *Media Psychology, 25(1),* 82-105. doi: <https://doi-org.ezproxy.lib.ryerson.ca/10.1080/15213269.2020.1868322>

Maymin, P. Z. (2019). Smart kills and worthless deaths: eSports analytics for League of Legends. *Journal of Quantitative Analysis in Sports, 17(1).* doi: <https://doi.org/10.1515/jqas-2019-0096>

Kim, S., Kim, D., Ahn, H., & Ahn, B. (2021). Implementation of user playstyle coaching using video processing and statistical methods in league of legends. *Multimedia Tools and Applications, 80,* 34189–34201. Retrieved from <https://link-springer-com.ezproxy.lib.ryerson.ca/article/10.1007/s11042-020-09413-4>

Costa, L. M., Mantovani, R. F., Souza, F. C. M., & Xexeo, G. (2021). Feature Analysis to League of Legends Victory Prediction on the Picks and Bans Phase. *2021 IEEE Conference on Games (CoG).* Retrieved from <https://ieee-cog.org/2021/assets/papers/paper_292.pdf>