Statistical Models for Predicting Results in Professional League of Legends

Robbie Jadowski[1] and Stuart Cunningham[1,2][0000−0002−5348−7700]

[1] Department of Computing and Mathematics, Manchester Metropolitan University, Manchester, M1 5GD, UK ROBBIE.JADOWSKI@stu.mmu.ac.uk
[2] Centre for Advanced Computational Science, Manchester Metropolitan University, Manchester, M1 5GD, UK s.cunningham@mmu.ac.uk

Abstract. The esports industry has seen enormous growth in popularity. With increased viewership and revenue, further investment has been made to improve professional players' competitive strength. The modern esports team is a hierarchical business fuelled by investors and sponsorship. This paper is focused on the professional competitions in League of Legends esports. In existing real-world sports such as football or baseball, there is great attention paid to statistic driven analysis of the competition, and these stats are used to quantify player and team performance. These statistics hold significant value for competitive improvement, the gambling industry, and market influence within the esports industry.
This paper presents an analysis of data and metrics gathered from professional games during 2020 in several League of Legends international competitions. The objective was to build a predictive model through the combination of existing data analysis and machine learning that can rate team and player performance. The best performing model was able to correctly predict 67% of 306 games. Results indicate that while it is possible to predict the outcome of a competitive League of Legends game, to do so with a higher degree of accuracy would require substantially more data and contextual information.

Keywords: esports · League of Legends · machine learning · regression.

1 Introduction

Statistical analysis of sports, or sports analytics, has become an increasingly popular method for recruitment and strategising in modern sport and competition. The popularisation of sports analytics is often attributed to Billy Beane, who famously achieved great success as the general manager of the Oakland Athletics baseball team using a data-driven approach to evaluate and recruit players on a much lower budget than competing teams. Other teams took note of this approach and went on to achieve success through data-based decision making. This success was noticed by executives and owners of teams in other professional sports leagues, to the point where practically all modern sporting organisations now recruit analytic experts or entire departments dedicated to sports analytics [12].

The convenient nature of statistics allows managers and coaches to identify a player's strengths and weaknesses at a glance, without having to spectate each game the players compete in. The same data can used by gambling organisations to determine probability and assign odds to certain outcomes.

For example, football statistics have evolved to include automated sensing technology that can track player position, movement and other observations from fixed and mobile cameras and sensors. Several professional statistical analysis firms offer data and analysis to professional teams as a product, providing context to the data collected and helping teams make tactical decisions [2].

Since League of Legends (LoL) is a video game, an abundance of statistics can be gathered automatically as they are tracked by the game itself. The wealth of data available provides many opportunities to perform analytics on the game. Most of the existing forms of public analytics involving LoL is used by journalists and fans to make comparisons and fuel narratives. Other organisations provide LoL teams with a paid product package to enhance in-house analysis and supplement coaching.

The aim of this research is to build a statistical model using metrics from this data that can accurately rate team and player performance, with the intention of predicting the outcome of games featuring those players and teams in future games.

2 League of Legends

League of Legends was released in October 2009, and in the years since its release, it has developed a competitive infrastructure across multiple regions that rivals that of traditional sports [8]. Each region's competitive league features franchised teams that compete against each other in weekly broadcasts that regularly draw thousands of viewers and annual inter-regional championships that have drawn 44 million peak concurrent viewers during grand finals [21]. The events feature grand finals in venues such as the Staples Center, selling out the venue within 1 hour of tickets being available [22], and the Beijing National Stadium, catering to live audiences in their thousands.

LoL is a team-based strategy game where two competing teams of 5 players aim to destroy their opponents base, canonically named the *Nexus*. Each game of League of Legends takes place on the same map, known as *Summoner's Rift*. Summoner's Rift is split into three lanes, commonly known as Top, middle and Bottom. These lanes form a path that leads from one team's base to the other. The two sides of Summoner's Rift, referred to as 'Blue Side' and 'Red Side' are separated by a River that runs from top lane to bottom lane, and the area in-between the lanes is known collectively as the Jungle. Blue team's base and nexus is situated in the bottom-left of the map, while red team's base and nexus is in the top-right. A representation of the map is shown in Figure 1

Players select one of over 140 champions to control in order to complete the objective and each possesses abilities that aid in combat, navigating the environment or supporting their team. Each player fulfils a different role for the

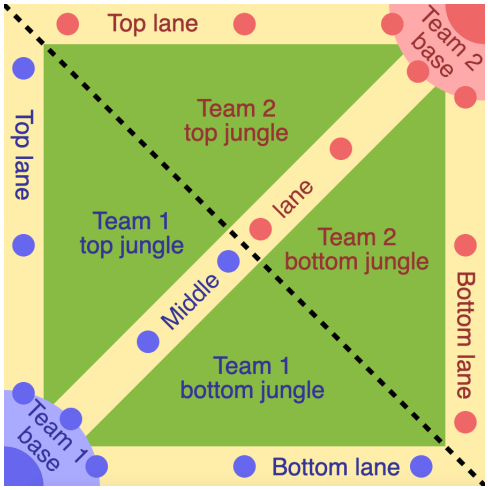

Fig. 1. Simplified Version of the Summoner's Rift Map. Original PNG version by Raizin, SVG rework by Sameboat licensed under CC BY-SA 3.0 [17]

team, much like the different positions in a football team. The roles featured in LoL are: Top Laner; Jungler; Mid Laner; Bot Laner; and Support. Each corresponds to the area or lane of the map that the player will operate in the opening of a game, with the Support player often partnering with the Bot Laner. These roles traditionally feature a typical character archetype, though there are exceptions and champions that buck the trend.

For a team to reach and destroy the enemy team's nexus, they must overcome a series of AI controlled structures known as Turrets. These structures are very difficult to destroy without assistance, which is usually provided by the waves of AI controlled minions that spawn periodically from a team's base. These minions will follow a lane's path to the enemy base until they run into the opposing team's champions, minions or turrets. Players must aid their minions in their advance in order to take down Turrets and reach the opposing teams base, while defending their own Turrets from the opponents.

The map also features neutral objectives, Dragon, Rift Herald and Baron Nashor. These neutral monsters can be defeated by a team to obtain permanent and temporary bonuses, ranging from additional movement speed, a percentage increase in ability power, or buffs to friendly minions to aid in sieging the opponents base.

Due to the asymmetrical nature of the map, granting the blue team easier access to the area that Baron Nashor spawns, combined with the pre-game champion draft where blue side can choose their first champion before the red team, there is a debate that blue side has an inherent advantage compared to the red team. Similar to the home advantage often seen in traditional sports. This advantage will be explored when analysing the data from competitive games and considered when making predictions if such an advantage exists.

3 Background

The use of player rankings in LoL is recognised as being an important feature of the game for individuals as well as to ensure the competitive edge of the game [11], which may arguably extend to system of team rankings and statistics. Previous work has examined the effect that the ability of LoL players working together in teams, and the presence of female gender players, has in being able to predict the competitive performance of those teams, however this relies upon individual measures being taken from players, such as measures of collective intelligence, gender, and so forth, that are not intrinsic to the LoL game statistics and so require additional information gather to take place [10]. Unsurprisingly, much existing research tends to point towards the influence that individual players, and their ability to form effective teams, can have on game outcomes [5, 4]. However, in terms of win prediction, it has been shown that for other Multi-player Online Battle Arena games in *professional contexts*, accuracy rates of up to 85% are possible [9].

4 Dataset and Preparation

4.1 Dataset Source

This report is focused on seven competitive leagues in LoL: the LEC (Europe); LCS (North America); LCS Academy (North America); LCK (South Korea); PCS (Southeast Asia); CBLOL (Brazil); and TCL (Turkey). While the Chinese league is the largest and perhaps most dominant region, there is insufficient data for each individual game available, and so it is excluded from the analysis. By using the data from every competitive game played during the 2020 spring split from 24th January 2020 to 2nd March 2020, we aim to predict the outcome of games that take place in the 2020 summer split. Each training dataset was validated using 10-fold cross validation. There were a total of 306 games in the 2020 Summer Split dataset used for testing.

There are several independent analysts who create content and collect data of competitive LoL to enable community-driven analysis and discussion. The data used in this report was obtained from an independent analyst, Tim Sevenhuysen, who runs the website oracleselixir.com [19].

The training data featured 882 games of data. Each game includes 12 rows. One row for each player (10 players) and one row for each team (two teams). In order to make the data usable it was separated data, into two subsets: the raw data of per game averages of each team (Table 1); and the raw data of per game averages of each player. Player statistics comprise of: *Position; Games Played; Win Percentage; Counter-Pick Rate; Total Kills; Total Deaths; Total Assists; Total Kill/Death/Assist Ratio; Kill Participation; Kill Share; Average Share of Team's Deaths; First Blood Rate; Average Gold Difference at 10 minutes; Average Experience Difference at 10 minutes; Average Creep Score Difference at 10 minutes; Average Monsters + Minions killed per minute; Average Share of Team's Total Creep Score post-15-minutes; Average Damage to Champions per*

minute; Damage Share; Average Earned Gold per minute; Gold Share; Average Wards Placed per minute; and *Average Wards Cleared per minute*. The players are separated by their role in the team, since different metrics can be more important to specific roles.

Table 1. Metrics and opposite metrics in the team statistics subset

Metric	**Opposite**
Kills	Deaths
Gold at 15	Opponent Gold at 15
XP at 15	Opponent XP at 15
CS at 15	Opponent CS at 15
Towers	Opponent Towers
Dragons	Opponent Dragons
Vision Score per Minute	Opponent Vision Score per Minute
Kills per Minute	Opponent Kills per Minute
Damage per Minute	Opponent Damage per Minute
Barons	Opponent Barons
Heralds	Opponent Heralds
Inhibitors	Opponent Inhibitors
Wins	Losses

4.2 Performance Measures

Pythagorean Expectation Pythagorean Expectation (PE) is used to calculate the expected total wins for a competitor over a number of games. George William 'Bill' James, known for his approach to analysing professional baseball using data and statistics, developed the formula to predict a baseball team's win percentage from the observed number of runs scored and runs allowed during a given baseball season. James is widely recognised for coining the term *Sabermetrics*. This term is a combination of the acronym SABR (Society for American Baseball Research) and the word metrics. Sabermetrics has become widely accepted as a useful baseball evaluation tool [3]. It is argued that the PE was the impetus for baseball's Sabermetricians movement, where, most notably, the Oakland Athletics adopted statistical principles that revolutionised their approach to baseball team management [12].

$$W = \frac{S^2}{S^2 + A^2} = \frac{1}{1 + (A/S)^2} \tag{1}$$

In the original formula, W is the win percentage, S is the observed number of runs scored, and A is the observed number of runs allowed. James initially used an exponent of 2, inspiring the use of Pythagorean in the formula's name. The formula has since been studied to identify the optimal exponent value for accurate predictions. Different exponents can be calculated for each team in

order to more accurately predict win percentages, and methods to find those exponents, such as the Pythagenpat formula, have been developed

$$n = \frac{S + A}{G}^{0.287} \tag{2}$$

where n is the exponent, and G is the total number of games. Though originally used for baseball, the simple concept of an offensive and defensive stat forming the foundation of the PE formula means that it can be applied to other sports [13, 15].

For LoL there are several metrics that can be used in an application of PE. The most obvious one would be kills and deaths. While the win condition of LoL is not having a higher margin of kills than the other team, it is an obvious metric that usually indicates the more dominant team. Another alternative would be turrets destroyed vs turrets lost. The planned model for rating teams will be calculating an overall offensive and defensive rating for each team, so these ratings can also serve as the values used in the PE formula.

Log5 Once the values of the PE formula for each team are known, we can use another formula to estimate the probability of one team beating another. James also devised Log5, a formula that uses two teams' winning percentages to calculate head-to-head match up probabilities [14].

$$pA, B = \frac{pA - pA \times pB}{pA + pB - 2 \times pA \times pB} \tag{3}$$

The Log5 formula considers the winning percentage of team A (pA) and team B (pB) and returns the percentage chance that team A beats team B. From which we can easily calculate the chance that team B beats team A. We can experiment using this formula with the values obtained from PE and compare them to predictions from logistic regression models to see if it offers better or worse performance.

Strength of Schedule If two teams have an equal record in sports, it can be challenging to determine which one could technically be considered the better team. One way of determining this is to assess the strength of the schedule for each team. Strength of Schedule (SOS) refers to the strength of the opponents a team has faced, compared to others [6].

Calculation of SOS involves comparing the combined winning percentages of each team's opponents against their own record or adjusting statistics by adding or subtracting based on an opponent's record. Assessing a team's strength of schedule can lead to interesting insights, where a bad team who appear strong on paper, may have only played against weaker teams, and a good team with worse statistics may have only played against stronger teams. Since the LoL teams that place higher in the rankings during the spring split round robin phase progress to the spring split playoffs, they end up playing more games

against tougher opponents than other teams. A team's per game average stats might be lower than a worse team, simply because they had to play more games against stronger teams.

This work will be taking SOS into account when analysing the data set, since many teams in LoL do not play against each other the same amount of times over the course of a split. This results in a method of adjusting a team's stats based on the strength of their opponents, with the goal of identifying a team's strength of schedule and building a more accurate representation of a team's overall strength.

To calculate a team's adjusted total, the metric M for a team T is

$$AdjTotal_{MT} = \sum_{i=1}^{N}(OppStat_i - AvgStat_M - SideAdv_{MT}) \tag{4}$$

where N is the number of games featuring the selected team, $OppStat_i$ is the opponent's opposite raw stat in row i, $AvgStat_M$ is the overall league average stat for metric M, and $SideAdv_{MT}$ is the average advantage/disadvantage for metric M on team T's side of the map.

The adjustment to the chosen metric is made by dividing $AdjTotal$ by the number of games a team has played and subtracting that from $RawStat$

$$AdjustedStat_{MT} = RawStat_{MT} - \frac{AdjTotal_{MT}}{TotalGames_T} \tag{5}$$

where $RawStat_{MT}$ is the raw per-game average stat for metric M for team T and $TotalGames_T$ is the total amount of games played by team T.

Using this information, one can calculate what a team's adjusted stats would be for each metric and compare them to their actual performance. If a team's adjusted stats are lower than their actual performance, this would indicate that the level of their opponents was worse in that metric and vice versa.

5 Evaluation

5.1 Team Ratings

Side Advantage Before devising and evaluating a model, it is important to determine if the dataset is balanced or not. In this case, whether the side of the map a team starts on provides any advantage. While most physical sports feature a home advantage due to familiar locations, less travel and playing in front of their own fans, LoL takes place in a virtual environment. Therefore, no significant difference or advantage for either team should be discovered. Despite this, there are major differences between starting on either side of the map that could provide an advantage to a team.

It may be argued that the blue side of the map holds an inherent advantage due to several factors. These include the asymmetrical geometry of Summoner's Rift and the isometric point-of-view favouring the blue side of the map. Most importantly, the pick/ban phase strategy of a team is often dictated by the side of

the map the team is going to playing. Data suggests that this side advantage does exist. In 2017, professional League of Legends games saw a period where blue side had a win rate of 64%. So much so that the developers of LoL, have sought to balance this advantage through various balance updates, such as making dragons a more lucrative objective.

The dataset used in this study includes 882 games, of which blue side won 477. This equates to a 54.08% win rate for blue side. A chi-square test suggests that the side of the map does have an impact on a team's chances of winning $\chi^2(1, 882) = 5.878, p = 0.015$. This infers that blue wins are expected to be more prevalent in the dataset, causing a slight imbalance.

Metric Selection Using all available metrics in a prediction model can be detrimental to its performance and prediction accuracy. Using a point-biserial correlation coefficient (PBCC) [23] calculation for each metric, identified which metrics strongly correlate with the result of a game. Two sets of calculations were carried out on each metric, one for the true stats of each game, and one for the per game averages of each team in each game. The tables were split into calculations for blue side and red side and ordered by the highest averages PBCC (converted to absolute value).

Table 2. Offensive metrics for forming team rating

Metric	True PBCC Abs. (Blue/Red)	Average PBCC Abs. (Blue/Red)
Towers	0.887 / 0.891	0.319 / 0.309
Inhibitors	0.734 / 0.755	0.293 / 0.277
Kills	0.644 / 0.731	0.258 / 0.267
KPM	0.664 / 0.735	0.263 /0.273

Table 3. Defensive metrics for forming team rating

Metric	True PBCC Abs. (Blue/Red)	Average PBCC Abs. (Blue/Red)
Opponent Towers	0.891 / 0.887	0.316 / 0.323
Opponent Inhibitors	0.755 / 0.734	0.285 / 0.290
Opponent Dragons	0.672 / 0.636	0.255 / 0.290
Opponent Barons	0.674 / 0.574	0.274 / 0.280

We selected the top 8 metrics from the red and blue teams because all 8 metrics scored above 0.5 absolute true PBCC and 0.25 absolute averages PBCC. They can also be evenly split into offensive (shown in Table 2) and defensive (Table 3) metrics, which will form the basis of offensive and defensive team ratings. The coefficient values can be used to calculate a weighting for each metric

Table 4. Weightings for Offensive and Defensive metrics

Metric	Tactic	Weight
Kills	Offensive	23.24%
KPM	Offensive	23.75%
Towers	Offensive	27.78%
Inhibitors	Offensive	25.23%
Opponent Barons	Defensive	24.75%
Opponent Dragons	Defensive	22.98%
Opponent Towers	Defensive	27.51%
Opponent Inhibitors	Defensive	24.77%

when producing a team rating. Another prediction model can also be formed by using these metrics as features, meaning that the results can be compared to the prediction models using all available metrics.

Normalization and Team Ratings In forming team ratings, Z-score normalization was selected over min-max as Z-score does a better job at handling outliers and will grant a team a higher value if they are drastically better in a particular metric, rather than pushing all other teams to be within a smaller range of each other.

Following normalization, the next step was to determine the weight of each metric. Weighting was calculated using the PBCCs used earlier to select the most relevant features. The weights were separated into offensive and defensive and calculated by summing the mean coefficients for each metric for both blue and red team and then calculating the percentage each mean coefficient contributes, shown in Table 4.

After calculating the weights, an offensive and defensive rating were formed using the sum of each normalized metric multiplied by its weight. This creates two new metrics, the offensive rating and the defensive rating. Figure 2 displays each team in terms of their offensive and defensive ratings, creating a visualization of where a team's strengths lie in their play style. These metrics can be considered opposites, lending themselves to being used in a Pythagorean expectation formula.

Pythagorean Expectation: Exponent To determine the most accurate PE exponent for the offensive and defensive ratings, we iteratively evaluated exponents between 0 and 10 and calculated the Mean Absolute Error (MAE) from each team's predicted y and actual x win percentages.

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} \tag{6}$$

This is done with the intention of finding the PE exponent value that minimises the MAE. The values of the defensive rating were inverted and each added to a constant of 5, since the formula relies on a lower, positive value, defensive

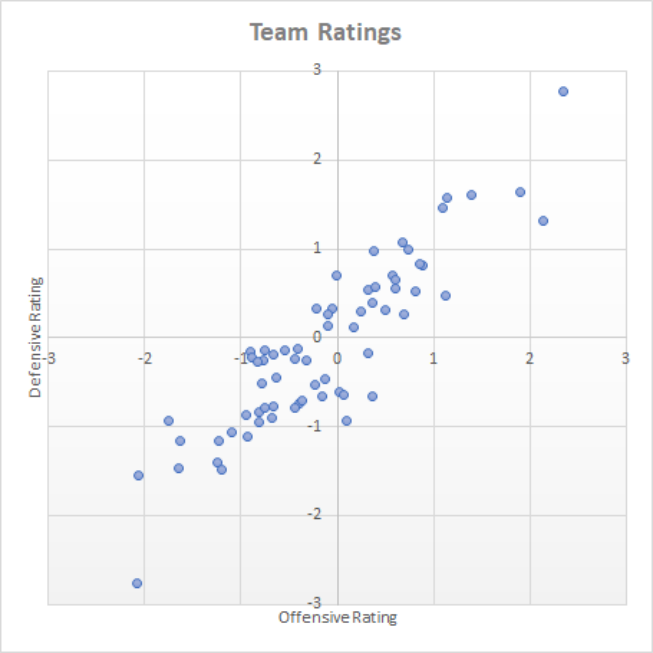

Fig. 2. Offensive and defensive ratings for all teams.

stat being a reflection of a team's ability. We found a value of 1.82 the most accurate single exponent to use for this dataset, with MAE of 0.0397. The MAE values for this exponent range are shown in Figure 3.

5.2 Player Ratings

Focusing on the performance of an entire team to predict results can be flawed for several reasons. In competitive LoL, teams may use substitute players to take the place of another player in a certain role. There is also the case of players transferring to a different team between each split. Teams will try to sign new players to replace under-performing ones, or more successful teams might attract the best players from lesser teams. While some teams tend to maintain a certain level of dominance despite changing their roster, this is usually down to the team's infrastructure and coaching. Most teams will notice a certain change in performance even by changing just one member of their roster.

Predicting future results only by a team's combined results could lead to problems if that team changes its roster. In this case, rating each player may result in more accurate predictions. By assigning each player their own rating, a modular overall team rating can be formed. The process for creating player ratings is similar to the process of creating team ratings described in the previous sub-section.

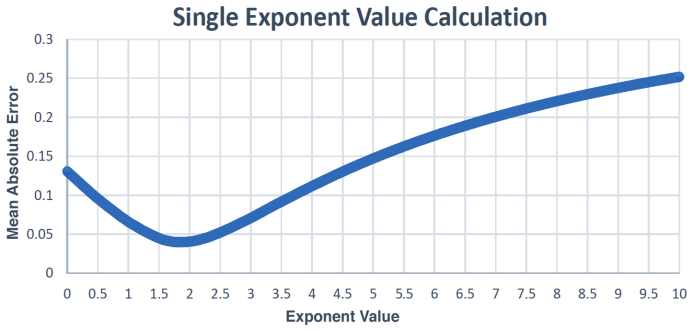

Fig. 3. Pythagorean Expectation: Exponent Value Calculation.

Rather than choosing the same metrics for every player type in LoL, there is reason to consider the difference in each class a player can assume, and what aspects of the game are important for that role. For selection of player metrics, the same process was followed as for team metrics, namely selecting the strongest correlation coefficients and weighting them accordingly. These are displayed in Table 5, noting the acronyms: Average Dominance Factor (Dom F); average Damage dealt to champions Per Minute (DPM); average Gold Difference between a player and their opponent in their respective role at the 10-minute mark (GD10); average Kills, Deaths and Assists ratio (KDA); average Creep Score Difference between a player and their opponent in their respective role at the 10-minute mark (CSD10); and average experience points (XP) Difference between a player and their opponent in their respective role at the 10-minute mark (XPD10).

After selecting the metrics for each role, the next task was to arrive at an overall rating. The player statistics dataset is not suitable for an offensive and defensive metric split, so each player will only have one rating based on the stated metrics. After calculating each player's rating, another model can be set up using each team's individual player ratings as a feature. Therefore, if a player is swapped out for a different one in a game, the rating will adjust to match the new player, affecting prediction outcome.

5.3 Performance Evaluation

Summary of Approaches A total of seven sets of features and approaches were evaluated to identify the one resulting in the best prediction outcomes. These approaches were: (1) the *un-adjusted per game metrics* per team (UT); (2) the *adjusted per game metrics* per teams (AT); (3) the eight *weighted metrics selected by their PBCC* scores per team (WT); (4) the calculated *offensive rating and defensive rating* per team (OD); (5) a *player rating* for each player in both teams (PR); (6) *actual win rate percentages* of each team (WP); and (7) the *expected win percentage* calculated using the Pythagorean expectation formula

Table 5. Player Metrics and Weightings

Metric	**PBCC**	**Weight**
Top Laner		
Win %	0.330	41.32%
Dom F	0.234	29.36%
Assists	0.234	29.33%
Jungle		
Win %	0.339	32.04%
Dom F	0.284	26.86%
GD10	0.161	15.20%
XPD10	0.148	14.02%
CSD10	0.126	11.88%
Mid Laner		
Win %	0.335	37.12%
Dom F	0.270	29.95%
DPM	0.149	16.49%
GD10	0.148	16.44%
Bot Laner		
Win %	0.344	29.36%
KDA	0.266	22.71%
Kills	0.232	19.82%
DPM	0.175	14.89%
GD10	0.155	13.21%
Support		
Win %	0.345	29.25%
Dom F	0.263	22.32%
Assists	0.238	20.20%
GD10	0.170	14.43%
XPD10	0.163	13.81%

for both teams (PE). Approaches 1 to 5 made use of logistic regression to predict game outcomes and 6 to 7 made use of the Log5 formula for prediction.

Performance Metrics and Results The following metrics were used to measure performance of the approaches: Classification Accuracy (CA) [18]; F1 Score (F1) [1]; Area Under the Curve (AUC) [7]; Mathews Correlation Coefficient (MCC) [1]; Log Loss (LL) [18].

Following training of the logistic regression models and calculation of the Log5 outcomes, the results were obtained for each approach using the test data set from the 2020 Summer Split, as shown in Table 6, where the highest performing outcome for each metric is highlighted in bold.

Table 6. Evaluation Results (Summer 2020)

Model	CA	F1 (Blue)	F1 (Red)	AUC	MCC	LogLoss
UT	0.618	0.670	0.545	0.649	0.217	0.544
AT	0.621	0.669	0.557	0.657	0.226	0.528
WT	0.631	0.685	0.553	0.654	0.242	0.489
OD	0.634	0.689	0.556	0.639	0.248	0.516
PR	**0.673**	**0.724**	**0.600**	**0.666**	**0.329**	**0.413**
WR	0.592	0.654	0.512	0.621	0.168	0.509
PE	0.637	0.691	0.561	0.638	0.255	0.491

The *Player Rating* model scores best in each performance metric, especially MCC, while all models suffered lower F1 scores for predicting Red Wins than predicting Blue Wins. This indicates that the models have more difficulty identifying if the red team wins, and seems resistant to predict this, despite having taken the blue side advantage into account during stat adjustments for the models. Prediction performance of wins for the *Player Rating* model is illustrated in Figure 4.

6 Conclusions and Future Work

The Player Rating approach achieved a significant classification accuracy of 67.3% classification when predicting 306 games from the 2020 Summer Split, which is significantly better than chance $\chi^2(1, 306) = 11.560, p < 0.001$. Compare this to results in the 2015/2016 English Football Premier league season through logistic regression, where 69.5% accuracy was achieved [16] after iterating on earlier work that achieved accuracy of 51.06% predicting the 2011/2012 season [20]. As this result is a first iteration, it stands to reason that improvements are possible. The findings may also have utility in player scouting, where a player may be performing better than their competitors, but is on a worse performing team.

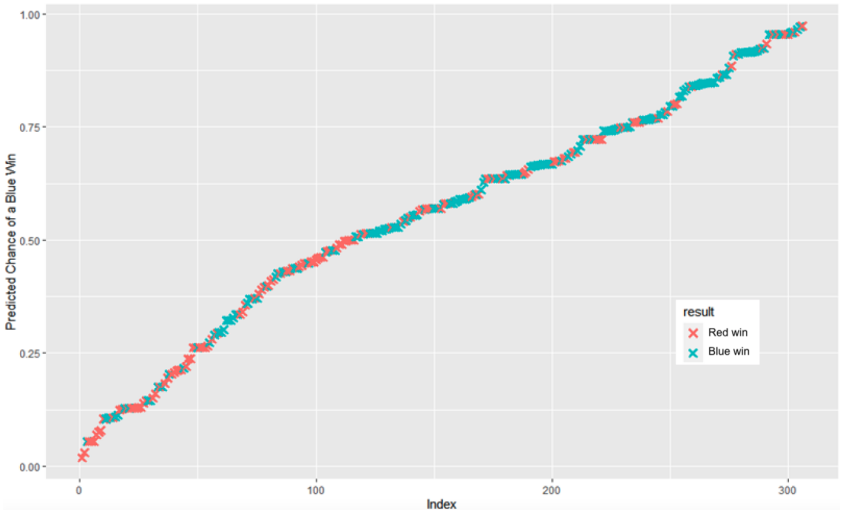

Fig. 4. Player Rating model prediction results.

The Win Rate approach scored worst in all metrics other than LogLoss. This confirms assumptions can't be made based on the previous results of teams alone, but further investigation into their actual performance in other game metrics reveals that they will better influence future results.

Since the approaches used per game averages rather than game by game data, they were unlikely to achieve a 90%+ classification accuracy, due to the inevitability of upsets. Even during the closing periods of a LoL game, the outcome can be highly volatile due to the nature of the game.

Future work should include a way to update a model after each game is played and weight more recent games higher than older when calculating a team's strength, eventually forgetting those games as they become irrelevant. For additional experimentation, a combination of team ratings and player ratings would likely be ideal. Due to the small team size in LoL, a roster change can have massive implications on the future performance of a team. There is also precedent for dominant teams falling, even without roster changes.

References

1. Chicco, D., Jurman, G.: The advantages of the matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation. BMC genomics **21**(1), 1–13 (2020)
2. Cintia, P., Giannotti, F., Pappalardo, L., Pedreschi, D., Malvaldi, M.: The harsh rule of the goals: Data-driven performance indicators for football teams. In: 2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA). pp. 1–10. IEEE (2015)
3. Costa, G.B., Huber, M.R., Saccoman, J.T.: Understanding sabermetrics: An introduction to the science of baseball statistics. McFarland (2019)

4. Costa, L.M., Souza, A.C.C., Souza, F.C.M.: An approach for team composition in league of legends using genetic algorithm. In: 2019 18th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames). pp. 52–61. IEEE (2019)
5. Do, T.D., Dylan, S.Y., Anwer, S., Wang, S.I.: Using collaborative filtering to recommend champions in league of legends. In: 2020 IEEE Conference on Games (CoG). pp. 650–653. IEEE (2020)
6. Fearnhead, P., Taylor, B.M.: Calculating strength of schedule, and choosing teams for march madness. The American Statistician **64**(2), 108–115 (2010)
7. Fogarty, J., Baker, R.S., Hudson, S.E.: Case studies in the use of roc curve analysis for sensor-based estimates in human computer interaction. In: Proceedings of Graphics Interface 2005. pp. 129–136 (2005)
8. Games, R.: League of Legends. Riot Games, Garena, Santa Monica, CA, USA (2009)
9. Hodge, V.J., Devlin, S.M., Sephton, N.J., Block, F.O., Cowling, P.I., Drachen, A.: Win prediction in multi-player esports: Live professional match prediction. IEEE Transactions on Games (2019)
10. Kim, Y.J., Engel, D., Woolley, A.W., Lin, J.Y.T., McArthur, N., Malone, T.W.: What makes a strong team? using collective intelligence to predict team performance in league of legends. In: Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing. pp. 2316–2329 (2017)
11. Kou, Y., Gui, X., Kow, Y.M.: Ranking practices and distinction in league of legends. In: Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play. pp. 4–9 (2016)
12. Lewis, M.: Moneyball: The art of winning an unfair game. WW Norton & Company (2004)
13. Morey, D.: Stats basketball scoreboard. STATS Basketball Scoreboard pp. 1–288 (1993)
14. Morey, L.C., Cohen, M.A.: Bias in the log5 estimation of outcome of batter/pitcher matchups, and an alternative. Journal of Sports Analytics **1**(1), 65–76 (2015)
15. Oliver, D.: Basketball on paper: rules and tools for performance analysis. Potomac Books, Inc. (2004)
16. Prasetio, D., et al.: Predicting football match results with logistic regression. In: 2016 International Conference On Advanced Informatics: Concepts, Theory And Application (ICAICTA). pp. 1–5. IEEE (2016)
17. Raizin, Sameboat: Simplified version of the summoner's rift map. CC BY-SA 3.0 (https://creativecommons.org/licenses/by-sa/3.0/) (2013), https://commons.wikimedia.org/w/index.php?curid=29443207
18. Saleh, H.: Machine Learning Fundamentals: Use Python and scikit-learn to get up and running with the hottest developments in machine learning. Packt Publishing Ltd (2018)
19. Sevenhuysen, T.: Oracle's elixir - lol esports stats (2021), https://oracleselixir.com
20. Snyder, J.: What actually wins soccer matches: Prediction of the 2011-2012 premier league for fun and profit. Thesis, University of Washington, WA: Department of Computer Science (2013)
21. Staff, L.E.: 2019 world championship hits record viewership. https://nexus.leagueoflegends.com/en-us/2019/12/2019-world-championship-hits-record-viewership/, accessed: 2021-03-26
22. Tassi, P.: League of legends finals sells out la's staples center in an hour. Forbes (2013)
23. Tate, R.F.: Correlation between a discrete and a continuous variable. point-biserial correlation. The Annals of mathematical statistics **25**(3), 603–607 (1954)