CAN WE PREDICT THE OUTCOME OF A

LEAGUE OF LEGENDS

GAME IN THE FIRST 10 MINUTES?

MEGIN VAN HERK,
SEMESTER 4

REPORT DOCUMENT.

The demonstration materials report and/or present the approach, process or outcomes of your work in a methodologically sound way targeted on the client.

OVERVIEW

WHY

When you think you suck at a video game. you just keep losing, cant seem to win. it might motivate you just to quit. But what if you knew exactly what you needed to do to be able to win. That is why this research is needed. to motivate people to keep playing league of legends. Improve themselves. Improve the game.

PRODUCT

The end project would be something similar to an app within the game, Telling you your chances of winning at the 10 minute mark.

GOALS

The goal of the project is to help players improve themselves. Based on their own and other players games. Showing during or after a game a section in which a player can improve or a section they have mastered, can really motivate someone to do better and keep playing.

League player looking to improve their win rate, is my target user. It would be the best fit for beginner or intermediate players. Since well experienced people don't need to know their win rate, and since the accuracy of the program isn't 100%, experienced players probably know better themselves. Such as in the esports scheme.

MACHINE LEARNING

ALGORITHMS

We tested many algorithms to find the best ones for this project. Including:

- KNN, nearest neighbor
- Logistic regression
- SVM, sigmoid kernel

VARIABLES

League of legends is a big game with allot of different factor impacting the outcome. Choosing the right one to perform the algorithm on is very important

The below given graph shows The red team's kills compared to the blue team's kills. (0=lost, 1=wins). You can see at some point when the blue team gets a certain amount of kills they stop losing.
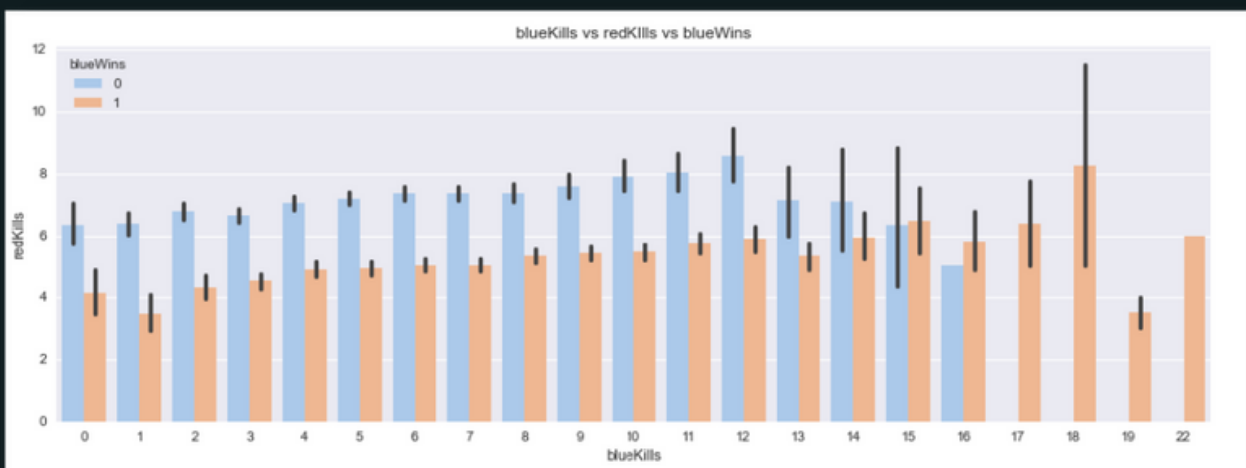
Kills have a big impact on the game the graph below shows, kills and deaths compared. The blue color means losing and green winning.
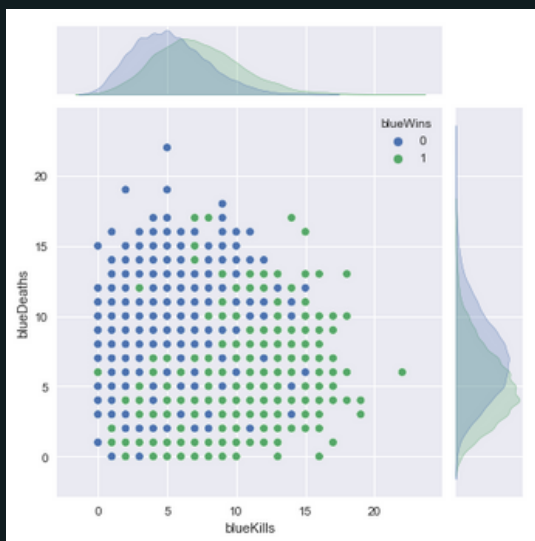
Here you can clearly see the line between win and los, exluding the outlier ofcourse. The more deaths and less kills the more likely you are to loose. and the more kills and less deaths the more likely you are to win.

I think these variables are best to perform our machine learning on, since they seem to impact the game the most. From the data, we can see that the outcome of a game is binary: the winner is denoted by either a 0 (blue team) or a 1 (red team). Using the data from different aspects of the game, we want to be able to predict who will win the game.

RESULTS

our target variable is if we win or lose, this variable is stored in both "blueWins" and "redWins". if blue wins the game it will be "blueWins=1, redWins = 0". You need only use 1, if you know blueWins= 0 then you can presume redWins = 1. so we only use "blueWins" as our target variable. Running the most impactful variables through the algorithm gave us these results:

APPROACH 1: NEAREST NEIGBOR

The k-nearest neighbors (KNN) algorithm is a simple, supervised machine learning algorithm that can be used to solve both classification and regression problems. Here we got a result of 0.6502024291497975 accuracy.
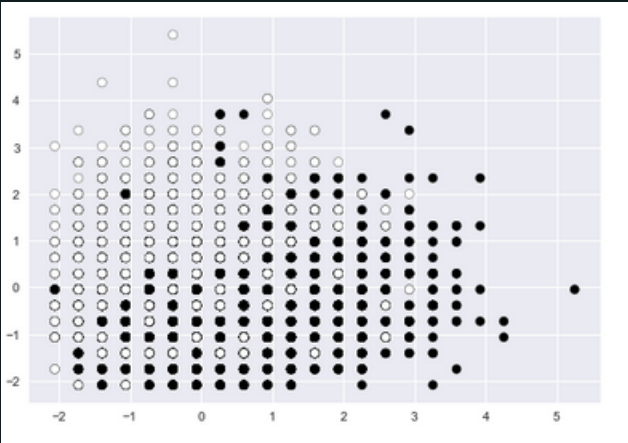
APPROACH 2: LOGISTIC REGRESSION

Because of the binary nature of the outcome we want to predict, we will use logistic regression to produce an output that will always be in between 0 and 1, and will allow it to correctly predict who will win.
1.0 here means you win the game. 0.0 means you lost. Its interesting to see we never touch 0.0 here, this means there is always a slight chance of you winning thegame no matter if you get 0 kills
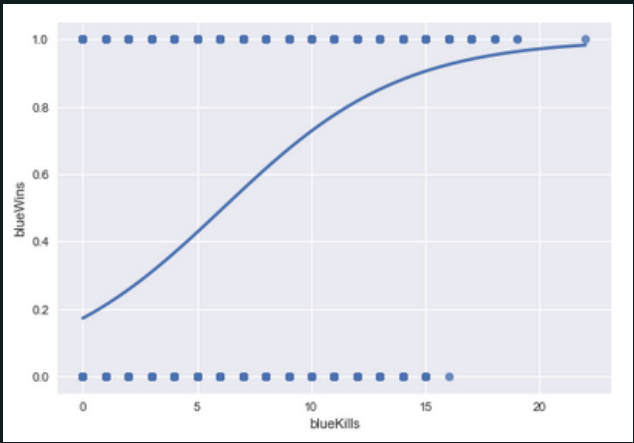
	precision	recall	f1-score	support
0	0.68	0.74	0.71	1623
1	0.72	0.66	0.69	1638
accuracy			0.70	3261
macro avg	0.70	0.70	0.70	3261
weighted avg	0.70	0.70	0.70	3261

All the scores are similar to one another. Recall has the highest with a 0.74. High recall means that the algorithm returns most of the relevant results. So the range of our accuracy is between 0.66 to 0.74, but generally 0.7 on average. Logistic regression is the highest algorithm accuracy score so far. In context to our domain, we can predict if you're going to win your game in the first 10 minutes with an accuracy of 70%

APPROACH 3: SVM

Sigmoid kernel
is mostly preferred for neural networks. So
binary, which in our case, 'blueWins' is either 1
or 0, so binary applies. The sigmoid function
takes any real-valued input and maps it to a
real number in the range (0, 1).

We can think of this almost like saying

"if the value we map to output near 1, the blue
team wins if it maps to output near 0, the
blue team loses".
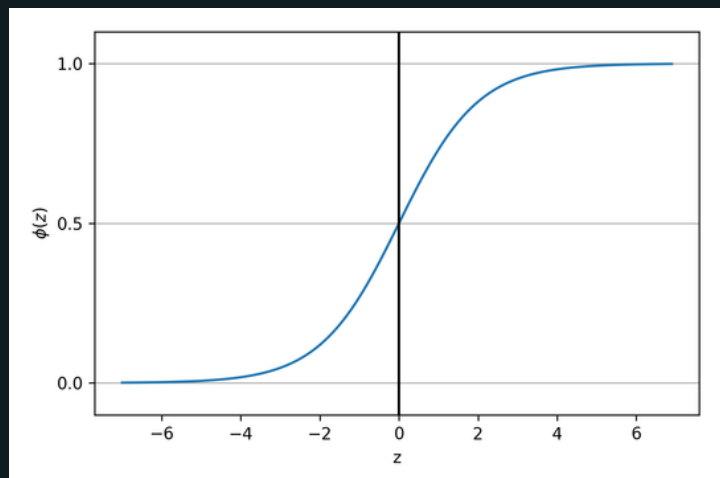
	precision	recall	f1-score	support
0	0.61	0.59	0.60	1250
1	0.59	0.61	0.60	1220
accuracy			0.60	2470
macro avg	0.60	0.60	0.60	2470
weighted avg	0.60	0.60	0.60	2470

So the range of our accuracy is between 0.59 to 0.61 but generally 0.6 on average. This
means we can predict the outcome of a league of legends game with a 60% accuracy

CONCLUSION

So it is possible to predict the outcome of a game within 10 minutes, but only
with an accucary of 70%, wich cannot always be trusted. 100% is not possible
in this domain, or atleast not at the 10 minute mark. There will always be a
chance for the tables to turn.