



github.com/cinspire93



sqrchen93@gmail.com



linkedin.com/in/cinspire93



DotA 2 Match Prediction

Chen Chen

Capstone Project

Introduction

- **Imagine a forty-minute game whose outcome can be predicted reliably at the ten-minute mark, this would not be a fun game. In fact, it is an UNBALANCED game.**
- **In the gaming industry, game balance is a very real design problem game developers have to face. An unbalanced game will fail to impress players and lose a lot of money.**

Introduction

- **DotA 2, or Defense of the Ancients 2, is a popular Multiplayer Online Battle Arena game (MOBA), in which ten players are split into two teams of five, called the Radiant and the Dire, who then try to destroy each others' Ancient.**
- **It is currently the second most popular MOBA game, trailing League of Legends and netting over \$200 million in yearly revenue. Both players and developers, unsurprisingly, take game balance very seriously.**

Motivations

- **My main goal, is to investigate whether DotA 2 is balanced by looking at the first ten minutes of any given match.**

Motivations

- However, there are two ramifications from my main objective:
 - First, with a good prediction model, I can offer solid strategic advice to players looking to maximize their win rates.
 - Second, there are in-game wagering systems that could benefit quite a lot from an accurate prediction model.
- We will come back to them.

Feature Engineering

- Quick inspection of the data revealed the absence of any useful in-place features. I had to engineer every one of them using domain knowledge.

Feature Engineering

- **1st Engineered Feature—Hero Selection:**



Feature Engineering

- **1st Engineered Feature—Hero Selection:**
 - **Heroes in DotA 2 do not just look different, they also work in many different ways. Therefore, it follows intuitively that certain hero compositions will affect match outcomes more than others.**

Feature Engineering

- **2nd Engineered Feature—Net Worth at the ten-minute mark:**



Feature Engineering

- **2nd Engineered Feature—Net Worth at the ten-minute mark:**
 - **Net worth, or the total amount of gold (in both item values and currency) a hero has, is a key economic indicator. Heroes with higher net worth can purchase better items, which amplifies their powers and makes them more effective in winning the match.**

Feature Engineering

- Other features I have engineered include net death counts from team fights within the first ten minutes, hero role compositions, and the interaction between a hero's role and its net worth.
- Sadly, they did not increase my model's accuracy much. I believe that the high complexity of hero selection and net worth information exhausted the information space within the first ten minutes of any match, leaving little to learn.

Findings and Results

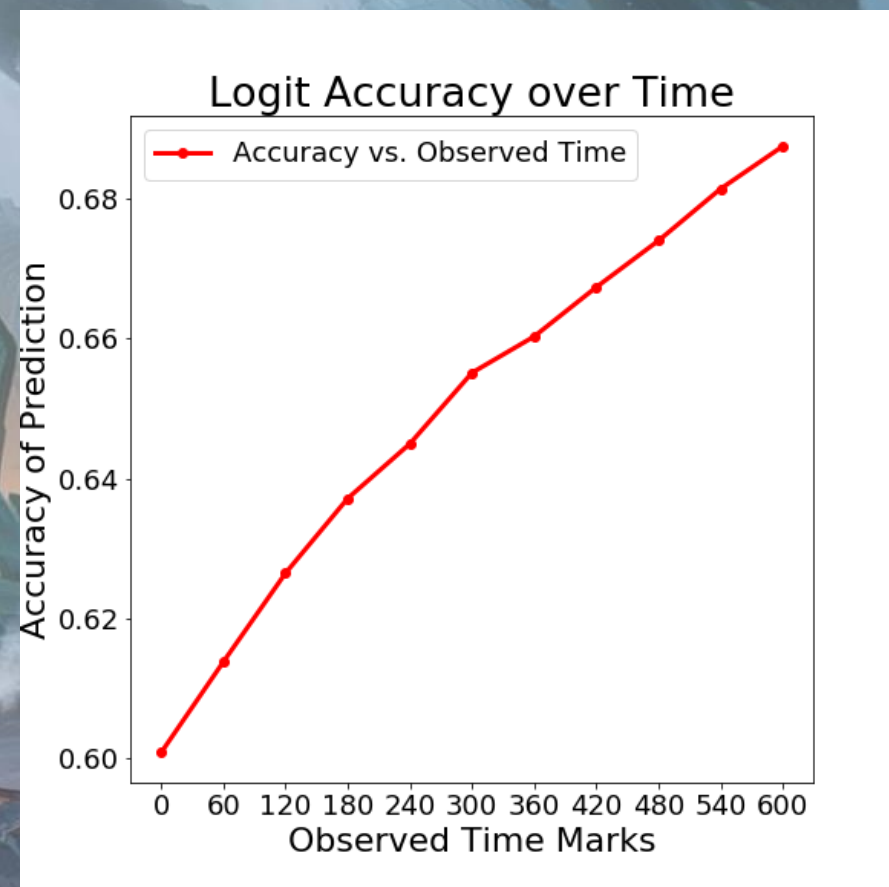
- I applied a Logistic Regression model, using hero selection and net worth as my input matrix, and achieved a 68.5% accuracy in match outcome prediction.
- A near 70% accuracy implies that DotA 2 is terribly balanced. However, I decided to dig in a bit more so as to not fool myself with the result.

Findings and Results

- The net worth information is time-dependent, but offers only a sliced view of a running match. A less stationary statistic should be more telling about the balance of any given match.
- What if I vary the time at which I collect the net worth information, retrain my model and check how its accuracy varies over time?

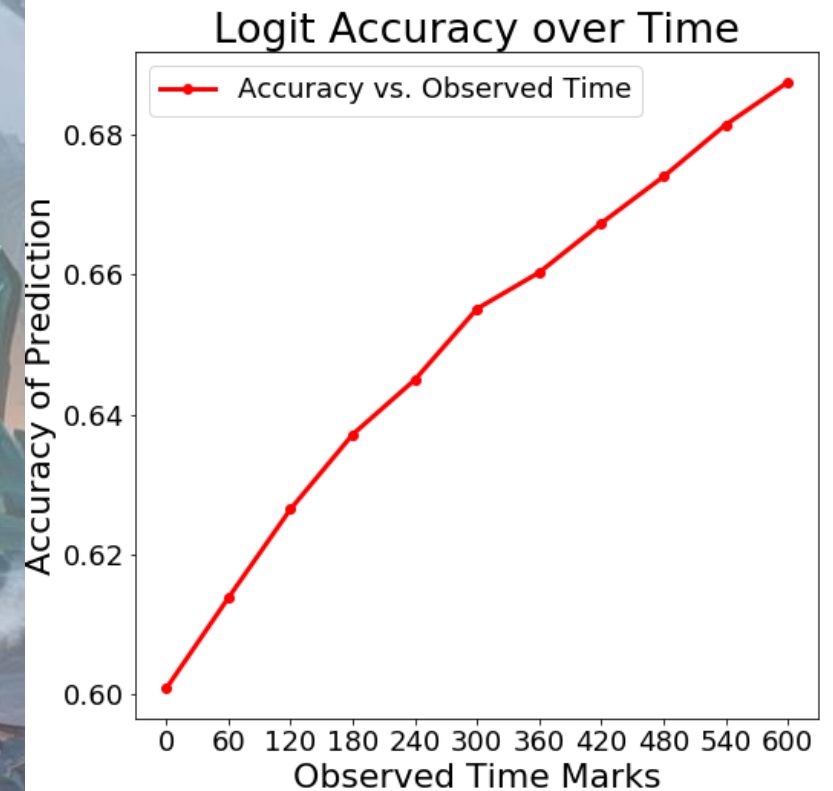
Findings and Results

In fact, what I observed was a near linear relationship between the accuracy of my model and the time passed. This tells me that there are no points in time, within the first ten minutes, that offers significant insight into the outcome of the game. In other words, the game is balanced with respect to time, because early games do not decide outcomes.



Findings and Results

Notice, however, that my model's accuracy did not start at 50% at time 0. This means that hero selection also matters. In fact, a bad strategic decision in hero drafting might as well cost you the game!




Conclusions

- **So in conclusion:**
 - **The game is balanced with respect to time, but not so much heroes. However, I believe that out-drafting your opponent in any match is what makes a good player. That tiny imbalanced incentive actually rewards those who really understand the game.**

Conclusions

- **On the other hand:**
 - **The Logistic Regression model performed fairly well in terms of accuracy. A wagering machine can definitely built to exploit this advantage. Betters better be happy!**



Thank you!
Questions?
Reach out to me!



github.com/cinspire93



sqrchen93@gmail.com



linkedin.com/in/cinspire93