

Heroes of the Storm Win/Loss Predictor

Edward Chiou

Introduction

Heroes of the Storm, or HotS for short, is a 5v5 multiplayer online battle arena (MOBA) game developed by Blizzard Entertainment. It is the most recent MOBA to surface, following in the footsteps of Riot Games' League of Legends and Valve Corporation's DotA 2. It entered its technical alpha testing phase on March 13, 2014 and its technical closed beta testing phase on January 13, 2015. Heroes of the Storm was officially released as of today, June 2, 2015.

The main objective of Heroes of the Storm, like its competitors, is to destroy the opponents' primary structure known as the Nexus. In order to do so, each player selects his/her own hero to control to engage the opposing team in combat. Heroes become stronger as they gain experience from destroying key enemy structures, killing enemy minions, clearing out neutral mercenary camps, and killing enemy heroes. Every hero has a unique set of four abilities and a trait. These abilities scale as the hero gains experience, allowing the hero to become more powerful over time.

There are three primary game modes: Quick Match, Hero League, and Team League. In Quick Match, players pick their hero before they are matched with other players. In Hero League, ten players are matched together before each player drafts his or her own hero in a top-down order from highest in-game rating to lowest in-game rating. For Team League, five players must create a team roster before entering. One main difference between Hero/Team League and Quick Match is that no hero can be

selected by both teams in Hero/Team League; essentially, duplicates across teams are only allowed in Quick Match. Players are matched via a hidden rating system that is adjusted for every victory and loss. This system is suspected to be similar to Elo or Microsoft's TrueSkill ranking system, but Blizzard Entertainment has yet to come forth with their exact methodology.

In Heroes of the Storm, a player must choose their hero carefully if they wish to succeed. In Quick Match, this often means picking heroes that have a high, consistent win rate or a hero that fulfills a much needed role. In Hero League and Team League, drafting strategy is crucial in securing the best possible heroes for your team while denying the opposition certain key heroes. We can see from this that Heroes of the Storm lends itself very well to a classic predictive modeling task where we seek to use the skill ratings of each player, their respective hero selections, and the overall team compositions to determine which team emerges victorious.

Dataset

For this predictive model I am using a set of data that was released on May 24th, 2015 by /u/barett77, a well-known Reddit user who owns the domain hotlogs.com (from which this dataset was procured), a third-party service that allows players to build their own personalized profile and match history by uploading their replays from the game. This particular dataset consists of 600,000 games collected over ten consecutive

days with the following qualities:

- Each game belongs to the “Quick Match”, “Hero League”, or “Team League” modes.
- Each game consists of ten actual human players and their hero selections (player identification scrubbed out for privacy reasons).
- There were no restraints on player skill. Games ranged from low skill to high skill. (More on skill ratings later).
- There are certain games that lack recorded skill ratings.

I split the dataset in half; 300,000 games for the training set, 300,000 games for the test set.

Additional Background

Since Blizzard Entertainment does not provide the public with game data, it is entirely up to third-party services like hotslogs.com to collect and process game data that is manually uploaded by players. The key feature of hotslogs.com is the player ranking system; hotslogs.com attempts to rank players by leveraging Microsoft’s TrueSkill Ranking System (used to matchmaking players on Xbox Live). This feature, despite being a fairly decent tool for tracking player progress, has several major flaws:

- Not all games have been uploaded to the service. For players who have uploaded a significant amount of their

games, this is less of an issue; for players with very few games uploaded, their calculated TrueSkill rating could be off by a large amount.

- There are players who attempt to inflate their TrueSkill rating. Since hotslogs.com relies entirely on player uploads, players can choose to upload any set of games. Some players have been known to only upload their wins, which skews their TrueSkill rating upwards.
- The TrueSkill ranking system does not necessarily reflect Heroes of the Storm’s actual ranking system. There is currently no way of knowing what ranking system is actually in use unless Blizzard Entertainment confirms it.

Exploratory Analysis

I performed some very basic exploratory analysis mainly to confirm whether or not the dataset fell in line with my expectations. The first task I performed was simply counting the win rate for each hero for all 600,000 games. The results of this task are shown below.

conclusions:

Hero Name	Hero Win Rate
Abathur	0.4484
Anub'arak	0.6044
Arthas	0.4901
Azmodan	0.5389
Brightwing	0.5594
Chen	0.5009
Diablo	0.4643
E.T.C	0.4662
Falstad	0.4969
Gazlowe	0.5129
Illidan	0.5015
Jaina	0.5107
Kael'thas	0.4984
Kerrigan	0.5088
Li Li	0.5408
Malfurion	0.4585
Muradin	0.5028
Murky	0.5598
Nazeebo	0.5536
Nova	0.5191
Raynor	0.4870
Rehgar	0.4863
Hammer	0.4390
Sonya	0.4227
Stitches	0.4293
Sylvanas	0.5193
Tassadar	0.4451
TLV	0.5958
Thrall	0.5160
Tychus	0.4705
Tyrael	0.5358
Tyrande	0.4881
Uther	0.4950
Valla	0.4955
Zagara	0.5134
Zeratul	0.4988

I compared this table to the hotlogs data and came to the following

- There are discrepancies between these numbers and the actual realized win rate for certain heroes. For instance, Kael'thas, the most recent hero released before June 2nd, has a win rate of less than 50%, which is considered below par. In reality, he has maintained above a 50% win rate after May 17th and is considered to be one of the strongest heroes in game. In conclusion, the time period in which this data was sampled may not have accurately reflected the potential of each hero.
- The win rates are clumped together very densely, which is an indicator that hero win rate might not be a strong indicator of victory or loss for any given team. Most win rates lie within a 42% to 55% range. Statistically this range might be enough to work with, but this data was sampled right after a huge patch release i.e. there were significant changes to the game that caused win rates to fluctuate significantly.

Past this point, I did not do much standard exploratory analysis. Instead I extrapolated from several external sources, such as my own experience playing the game as well as those of full-time Twitch.tv streamers. For a point of reference, I have played approximately 700 games of Heroes of the Storm and have a win rate of ~58% in Quick Match only.

Common things that I noticed in games that would lead to a win are:

- The more players that have are in a party i.e. premade group have a higher chance of winning than those who aren't. Players in parties are more likely to communicate with one another because they know each other beforehand. The dataset doesn't provide any way of identifying player clusters, so I ignore this.
- The team with the higher cumulative skill rating is more likely to win. The higher a player's skill rating, the more likely they are to contribute to a team's victory than to a loss.
- Balanced team compositions are necessary for winning. A team with too many of any category (warrior, assassin, support, specialist) is less likely to win than a balanced team. Blizzard has implemented matchmaking rules that prevent a team from becoming too imbalanced; therefore I believe that this is something worth looking into.

Methodology

My overall methodology was fairly simple – use a classification system to predict whether a team would win or lose based on the features of the game. I chose to use logistic regression because other predictive models for games like Heroes of the Storm have used logistic

regression to a certain degree of success (see Dota2cp).

My first step was to use logistic regression to account for each individual hero drafted by a team. There are 36 heroes total, so I created a feature vector with 36 dimensions, one for each hero.

Model A

$$X_i = \begin{cases} 1 & \text{if the team picked } i\text{th hero,} \\ 0 & \text{otherwise} \end{cases}$$
$$Y = \begin{cases} 1 & \text{if the team won,} \\ 0 & \text{otherwise} \end{cases}$$

My next step was to use logistic regression to account for the overall composition of each team. There are four defined roles by Blizzard i.e. the warrior role, the assassin role, the support role, and the specialist role (in this specific order).

Model B

$$X_i = \begin{cases} 1 & \text{if the team had the } i\text{th role} \\ 0 & \text{otherwise} \end{cases}$$
$$Y = \begin{cases} 1 & \text{if the team won,} \\ 0 & \text{otherwise} \end{cases}$$

My last step was to use logistic regression to account for the difference in cumulative skill rating between the two teams in each match. I summed up the skill ratings for the first team in each match and

subtracted the sum of the second team's skill rating from the firsts.

Model C

$$X = \left\{ \begin{array}{l} 1 \text{ if first team had higher} \\ \text{overall skill rating} \\ 0 \text{ otherwise} \end{array} \right\}$$
$$Y = \left\{ \begin{array}{l} 1 \text{ if the first team won,} \\ 0 \text{ otherwise} \end{array} \right\}$$

I used the LogisticRegression function available from the sklearn library to fit the data according to these models and ran them each individually.

Results

Model	Accuracy
A	0.5748
B	0.5915
C	0.5326

As seen from the results, Model B (team composition) has the highest accuracy out of the three models. I attempted to combine all three models in all combinations possible, but I was unable to achieve a predictor that performed better than 59%.

As I mentioned before, there have been attempts to predict win/loss using logistic regression for other similar games such as League of Legends and DotA 2. Many of the datasets and experiments seen thus far are different in the following areas:

- No skill rating used. Many of these datasets did not have access to rating information at

the time of experimentation.

- Much more selective when it comes to data gathering. Some experiments limit data to high skill games, other limit to a certain game mode, etc.
- Many of the state-of-the-art methods used to predict win/loss for multiplayer online battle arena games seem to use a graph model to detect networks among heroes. This is most likely the key element that allows for a decent prediction rate (>63%).

Conclusion & Thoughts

Overall, none of the models are significantly better for predicting win/loss compared to guessing at random. There are several possible reasons, some of which have been explained above:

- Skewed/inaccurate skill rating measurements
- Win/loss criteria are different at each level of play. Perhaps low skill games are more dependent on factors such as team composition, whereas high skill games are more dependent on factors such as which players are grouped together, which players are the most experienced, etc.
- There are seven maps in Heroes of the Storm. Hero win rates might vary from map to map.

- Hero pairings might be more significant than team compositions. Certain heroes synergize very well with others, regardless of imbalance in the team composition.
- The dataset was sampled over too short of a time period and at a bad time (right after patch release). A better dataset would have been 600,000 games played at least one week after patch release for players to familiarize themselves with any new changes and heroes.