Molde Dotaklubb - Doracle

Gruppe 8

av

Olav Ausland Onstad, Thomas Brinck-Mortensen og Sondre Horpestad

i

IKT201

IKT110 - Artificial Intelligence Architecture

Veiledet av Sondre Glimsdal

Fakultet for teknologi og realfag
Universitetet i Agder

Grimstad, desember 2022

Innhold

1 Introduction

Molde Dotaklubb has tasked us with improving their capabilities for analysis in the drafting phase. In order to do this, they have requested the creation of an interactive webpage that can help them improve their performance. We use multiple machine-learning models to predict which team is going to win as well as which hero a team should pick based on already chosen ones. The project also delivers a statistical analysis of each champion.

2 Theoretical background

2.1 Dataset

We got delivered a 7GB zipped file as our dataset, the decompressed size is around 32GB. We used a python library called **zipfile**, this way python takes care of the reading without us needing to decompress it.

2.2 Tensorflow

Tensorflow was used for the prediction of which team had the highest probability of winning after the drafting phase. This solution used a 2 layer MLP with an input layer as well as an output layer. In the first layer we used a size of 272 perceptrons, each one representing a hero, the same one was used in the second layer. If we could use more than 2 layer, we might would consider using a dropout layer to prevent our model to overfit. Overall we got a 58.5% accuracy on unseen data, our test / validation split was $80 - 20$.

2.2.1 Data / Input

For the input to our team win prediction model, we used a input shape of (272,) which represents the champions using **One-Hot encoding**.

2.3 SKLearn

The way SKLearn were used, were in combination with some data we had prepossessed and predict result of the data. The data it were trained on contained the combination of pairs of heroes and their win rate dependent on team. It were done withe relu as the activation's method and sgd as the way to solve it. The reason sdg were used were because after trying the different methods this gave the best results. It were also used to predict the win rate of two full teams, but this result were pore in comparison to TensorFlow.

2.3.1 Data / Input

3 Implementation

3.1 What hero is the most picked?

We found out that the hero with id_{14} was the hero with the highest pick rate, which corresponds to the hero **Pudge**. This hero had a pick rate of 0.2613 which means that he was present in 26.13% of the games in our dataset.

3.2 What hero has the highest win rate?

After running an analysis of our given dataset, we concluded that the hero with id_{61} **Broodmother** had the highest win rate of 63.48%. For this result, we looped through each game and associated a hero id with wins and losses.

3.3 Is there an advantage to playing Dire or Radiant?

The dataset we were given represented a clear advantage of playing on the radiant side, overall the radiant team had a win rate of 57.46% whilst the dire team had only a win rate of 42.53%.

3.3.1 What hero is most affected by the side?

The dataset showed two heroes were affected equally as much by side, these were 108 & 61. The difference between when the hero was on their best and worst side was a flat difference of 29% in their win ratio.

The answer was found by finding each hero's average win rate based on side, and then the largest difference these two percentages.

3.4 What hero has the highest impact on the game? (Define impact yourself).

We defined the hero with the most impact to be the hero with the highest win rate. We can circle back to the answer in avsnitt 3.2, where we got the hero with id_{61} which corresponds **Broodmother**.

3.5 What hero has the longest games?

The dataset given shows that the hero with the longest game is the hero with id_{34} which corresponds to **Tinker**. This hero had an average game length of 1970 seconds, which corresponds to 32.83 minutes.

3.6 What hero has the shortest games?

The dataset given shows that the hero with the shortest game is the hero with id_{61} which again corresponds to the hero **Broodmother**. This hero had an average game length of 1710 seconds, which corresponds to 28.5 minutes.

3.7 What pair of heroes are the best?

From the dataset, we found that the best pair of heroes are; id_{77} **Lycan** & id_{103}**Elder titan**. These pairs of heroes had the highest win rate of 75% over a total of 228 games.

3.8 What team (of 5 heroes) is the strongest (Define best yourself)?

We defined the best team based on the win rate. The strongest team by our definition was the combination of $[id_{14}, id_{19}, id_{65}, id_{78}, id_{103}]$**Pudge ,Tiny ,Batrider ,Brewmaster and Elder titan** which had a win rate of 90.91% over 11 games.

3.9 How can Molde Dotaklubb use the webpage to improve?

Molde Dotaklubb can take use of our webpage to get information about different champions. You can pick any champion to get the best partner, as well as winning percentage and pick rate.

Molde Dotaklubb can also take use of our state of art deep machine learning algorithm to predict their win chance based on full teams.