



# Predicting match outcome in DOTA 2

Team 3:

Hanwen Jiang

Renqiu Chen



# Overview

- . Use cases
- . Methodology
- . Data sources
- . Sprint cycle
- . implementation and codebase
- . Acceptance Criteria
- . Project goal

# Some brief introduction!

Some of u may not heard about DOTA 2...well this slide is prepared for u!!

One of the well-known MOBA video game, basically is a 5 vs 5 player game, whom destroy the base of other team will win!

Want to learn more about this game?

[https://en.wikipedia.org/wiki/Dota\\_2](https://en.wikipedia.org/wiki/Dota_2)





# Use cases

## 1. Administrator:

System will train the ML-model and select the best model to make the prediction

System will give the binary result (win or lose) or winning possibilities of the matches in a short period of time

System will prompt user to enter the specific hero on each side

## 2. User:

User can get the possible result of the match by entering the battle information

(e.g. specific hero)



# Methodology

Select the best features from the dataset and use Spark to build our model.

Train our model and choose the one with the highest accuracy to do the prediction

Quantify a set of factors in a given battle which affect the winning possibilities

Do data preprocessing, such as regularization and interpolation simulation



# Data sources

We will use the datasets coming from Dota 2 Matches data on kaggle website

Link: <https://www.kaggle.com/devinanzelmo/dota-2-matches>

Datasets most likely we will be using..

hero\_names.csv 113 rows

players.csv 500K rows



# Sprint Cycles

---Week 9 (11.4-11.10): do more research and learn new stuff and set up the environment

---Week 10(11.10-11.17): officially start, import csv file, parse file and select features

---Week 11(11.17-11.24): start build up our ML models and apply our computing algorithm

---Week 12(11.24-12.1): train our model, and apply “Play” framework as our front-end UI

---Week 13(12.1-12.9): testing our model, optimizing and do final coding to achieve what we expected



# Implementation and our codebase

Data pre-processing

Use library to build up our ML model and write our own computing algorithms

We will mostly use scala to develop this project, and use “PLAY” which is a scala framework as our UI to show the demo. If further programming language is added, we will add it to our final presentation!

Where is our codebase?

[https://github.com/jianghanw/CSYE7200\\_Team\\_Project\\_Team3](https://github.com/jianghanw/CSYE7200_Team_Project_Team3)





# Acceptance Criteria

Normally, accuracy and response time are two main factors to measure

Accuracy should be  $\geq 60$  percent. (if winning possibilities was given, the accuracy should be at least 90 percent)

The time for system to train the model should be no more than 25min.

The time for the model predict the match result should be no more than 2s



# Project Goal

1. Our own ML-model will not only predict the match result successfully but also achieve a high accuracy
2. Hopefully, the execution time will be as least as possible
3. Lastly, this is the first time we use scala to write a prediction project, we will try our best to apply we have learned in this project, and learn other new stuffs from this experience!



**Thank you!**