
Amrin Sandhar - 220211447

CSC 180 - 01 Intelligent Systems
Professor Chen

Final Project Proposal

4th April 2022

Paper Title:

Dota 2 Match Winner Prediction

Project Type:

B - Application Oriented

Team Members:

Amrin Sandhar - 220211447

Project Description:

This project will use match data from a sample of Dota 2 games to predict the winner of matches based on game mode, game type, and hero selection. A hero is one of many playable characters, each with different strengths, weaknesses, and attributes.

Motivation:

In a competitive game like Dota 2, players want to maximize their odds of winning through any reasonable means possible. Predicting team winners based on team composition can reveal potential synergistic or antagonistic combinations of heroes.

Professional teams, competing for large prize pools, may want to use such prediction when forming strategies for composition in stage matches.

Background:

In Dota 2, ten players play against each other in a 5 vs 5 format. Each player may choose a hero to play as. Each hero is unique, and each hero may only be chosen once within a match (i.e., there cannot be two instances of the same hero within one game).

Heroes each have unique “kits”, which describes their statistics (i.e., health, movement speed, etc...) as well as abilities. As such, certain compositions of heroes will have greater intrinsic synergy, as some heroes may fill in for the weaknesses of others, or multiply the strengths of others.

Due to the great number of heroes (113 at the time of capture of this dataset), there are many potential team compositions - approximately 140 million ($_{113}C_5$).

Data Set:

[Link To Dataset](#)

This dataset contains information on:

1. The winner of each match
2. The cluster ID (relative location of each server set)
3. Game mode (all pick, turbo mode, etc...)
4. Game type (ranked, unranked)
5. Heroes picked

System / Algorithmic Design:

In this project, a **Fully-Connected Neural Network** and **Convolutional Neural Network** will be used, both **without transfer learning**.

Logistic regression will be used, as this is a classification problem.

Evaluation Plan:

Data will be split into a **training set** and a **testing set**.

Due to the classification nature of this problem, **precision, recall, F-measure, confusion matrices, and ROC curves** will be used to evaluate our results.

Task Division:

All work will be completed by Amrin, as he will be working alone.