

10-701: Machine Learning Project Proposal

Sandeep Subramanian (sandeeps) and Manu Bhargava Reddy Nannuri (mnannuri)

October 1, 2014

Using Hero Impact Metrics to Improve Dota 2 Match Outcome Predictions

1. **Background** : Dota (Defense of the Ancients) 2 is an MMO (Massively Multiplayer Online) game developed by Valve [1] and released as a free to play to the public in 2013. Dota 2 has become the most actively played game on Steam [2], a platform of online games, with daily peaks of over 800,000 concurrent players [3].

A game of Dota 2 consists of 2 competing teams (Radiant vs Dire). Each team has 5 players who each get to pick a "hero" for themselves. The most common game format in Dota, the "all pick" mode, allows each player to chose from a pool of 108 heroes. Each hero has a set of abilities unique to the hero. The sole objective of each team is to destroy the other team's ancient structure while protecting their own. Each player is awarded gold and experience for killing NPCs and other players in the game. Heroes gain levels based on the experience earned and players can use the gold earned to buy items that make them stronger.

2. **Questions that hoped to be answered by our research** : *This projects hopes to improve the performance of extant approaches to predicting the outcome of Dota 2 matches.* Extant approaches to predicting the outcome of matches involve using solely the heroes that are picked by either team [4]. The approach uses binary features that indicates whether a player picks a particular hero or not. The Dota2API provides a wide range of statistics about each completed match. We hope to incorporate not just information about which heroes played against each other but also the impact they had in the game into our prediction engine.

3. **Dataset** : Valve released a Dota 2 match history WebAPI to extract match history and match details in an XML or JSON format. The WebAPI can be queried using a Steam API key that every Steam registered user has. The retrieved values required to be normalized. We will explore a simple min-max normalization and Z score normalization.

4. **Validation of Results** : Our system uses a supervised learning approach. The WebAPI results contains the result (Radiant or Dire win) for each match. We will tune our parameters using k fold cross-validation and test the performance by partitioning our dataset in to 80% training and 20% testing data.

5. **First steps and Milestones** : The first steps we will take is to hand-engineer our features and come up with a concrete metric for the "impact" a hero had in a game. We will also

are there any limits to how much data you can query, say, per day?

I think you can go into a little more detail. Will this info be available during test/evaluation? Would that make the comparison to [4] unfair?

I see, so you chose to use logistic regression since that is what the baseline uses. But I think you can do more for the midterm report. Look into automatically learning features and other learning algorithms.

explore the possibility of incorporating metrics that quantify synergistic or antagonistic relationships between heroes. The milestone we will set for the midway report is to come up with a baseline approach toward classification using logistic regression. Our subsequent milestones will be to explore various ML algorithms one after another, that hope to capture synergistic and antagonistic relationships between heroes.

6. **Minimum Project Goals for a successful project** : The minimum project goals would be to match the results produced in [4].
7. **Stretchable Goals** : There are Dota 2 replay parsers that exist which parse out specific events during the course of the game. For example, it can record the exact time at which a specific hero killed another, movement tracking of different heroes across the map. Incorporating temporal information into our approach could boost performance.

References

- [1] Steam, <http://store.steampowered.com/>.
- [2] Valve, <http://www.valvesoftware.com/>.
- [3] Hing, David (June 2, 2014) <http://www.bit-tech.net/news/gaming/2014/06/02/dota-2-prize-pool-passes-8m/1>.
- [4] Kevin Conley, Daniel Perry *How Does He Saw Me? A Recommendation Engine for Picking Heroes in Dota 2* CS 229 Project at Stanford University

↓
this is cool, but see note above on what kind of info is available at training and testing.

— I think you can do more for the midterm report.