Decoding Team Composition in MOBA Games
A Learning Approach

Jiachen Li   Yilin He   Chengliang Lian

Introduction

What's MOBA: Multiplayer online battle arena is a rising force in online games. Millions of players participate in MOBA games such as LoL and DotA 2 daily.

Game Rule: Each player picks a unique character, and 2 teams of 5 players are pitted against each other. A big factor in game result is the combination of characters picked in each team.

Problem: There are more than 100 characters in game, so an interesting problem is how to form a good team composition that is more likely to win.

Objective

1. Predict game result based on character selection by each team;
2. Interpret the parameters learnt by our ranking model to see which character may be important for a good team;
3. Compare the performance and characteristic of different machine learning algorithms on this learning problem;
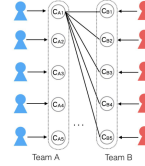
Model

- Two teams in game: T={A, B};
- Each team contains 5 unique characters;
- N: number of different characters in game.

Character Ranking Model

Feature vector: $\mathbf{x} = \{x_1, x_2, \ldots, x_N\}$
where

$$x_i = \begin{cases} 1 & \text{if } i \in \mathcal{A}, \\ -1 & \text{if } i \in \mathcal{B}, \\ 0 & \text{otherwise.} \end{cases}$$

Ranking Model with Prior

Character comparing matrix $\mathbf{C}_{N \times N}$
$$c_{i,j} = \mathbb{P}(\text{ i wins when played against j })$$

We model team A by
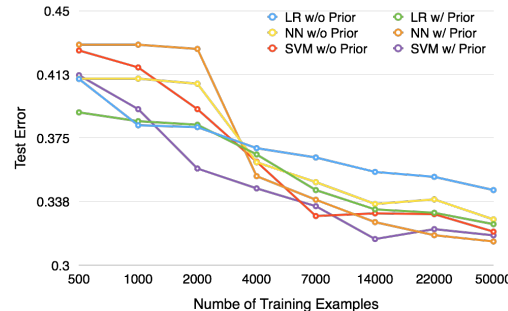$$\mathbf{a} = \{a_1, a_2, \ldots, a_N\} \quad \mathbf{a}' = \{a'_1, a'_2, \ldots, a'_N\}$$

$$a_i = \begin{cases} 1 & if \ i \in A \\ 0 & otherwise \end{cases} \quad a'_i = \begin{cases} \prod_{j \in \overline{\mathcal{A}}} c_{i,j} & if \ i \in A \\ 0 & otherwise \end{cases}$$

and model team B using $\mathbf{b}$ and $\mathbf{b}'$ obtained from same formula.

Modified Feature Vector: $\mathbf{x} = \{\mathbf{a}, \mathbf{b}, \mathbf{a}', \mathbf{b}'\}$

Analysis 1: Algorithm Comparison

Analysis 2: Prediction and Parameter

Game Result Prediction:

| Algorithm	DotA 2		LoL	
Accuracy (%)	w/o prior	w/ prior	w/o prior	w/ prior
Logistic Regression	65.57	67.58	55.90	55.36
SVM (Gaussian Kernel)	68.02	68.24	55.97	56.45
NN w/ pre-training	67.30	**68.60**	56.10	**56.80**

* DotA 2 and LoL have 50,000 and 15,000 training samples respectively.
* Our prior improves the accuracy of prediction
* Neural Networks model achieves the highest performance
* The performance on DotA 2 Dataset is more reliable.

Model Parameter Interpretation:
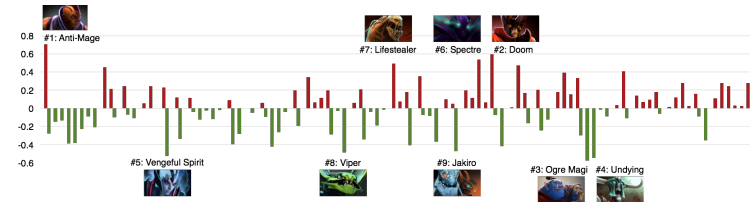
The weights of characters learnt in the ranking model for DotA 2

Conclusions

1. Team composition has inherent connection with game results, which can be interpreted by our models.
2. The prior knowledge can improve the learning performance. And as the training samples increase, all the algorithms can achieve similar results.
3. The prediction for game results could not be perfect due to the balance design principle in MOBA games.

Acknowledgement