Applications of Machine Learning in Dota 2: Literature Review and Practical Knowledge Sharing

Aleksandr Semenov[1], Peter Romov[2,3], Kirill Neklyudov[2,3], Daniil Yashkov[2,3], and Daniil Kireev[4]

[1] International Laboratory for Applied Network Research, National Research University Higher School of Economics, Moscow, Russia
[2] Yandex Data Factory, Moscow, Russia
[3] Moscow Institute of Physics and Technology, Moscow, Russia
[4] Moscow State University, Moscow, Russia
avsemenov@hse.ru, peter@romov.ru, k.necludov@gmail.com, daniil.yashkov@phystech.edu, dager.kd@gmail.com

Abstract. We present the review of recent applications of Machine Learning in Dota 2. It includes prediction of the winning team from the drafting stage of the game, calculating optimal jungling paths, predict the result of teamfights, recommendataion engine for the draft, and detection of in-game roles with emphasis on win prediction from team composition data. Besides that we discuss our own experience with making Dota 2 Machine Learning hachathon and Kaggle competitions.

Keywords: Dota 2, MOBA, eSports, Machine Learning

1 Introduction to Dota 2 game rules and mechanics

DotA 2 is an online multiplayer video game and its first part, DotA (after “Defense of the Ancients”) created a new genre, called Multiplayer Online Battle Arena (MOBA). It is played by two teams, called Radiant and Dire which consist of five players each. The main goal of the game is to destroy other team's “Ancient”, located at the opposite corners of the map.

Each of the players choose one hero to play with from a pool of 113 heroes in the drafting stage of the game (sometimes called 'picks'). Each hero has a set of features that define his role in the team and playstyle. Among these features there are his basic attribute (Strength, Agility or Intelligence) and unique set of 4 (or for some heroes even more) skills. These features allow each hero to fill several roles in the team, such as “damage dealer” (hero, whose role is to attack the enemies in the fight), “healer” (hero, who mostly heals and otherwise helps his teammates), “caster” (hero, who mostly relies on his spells) etc.

1.1 Previous Research on Dota 2

The first article mentioning Dota 2 was qualitative and analyzed correlation of leadership styles of players with roles in the game they choose to play [1]. In the first quantitative research of Dota 2, authors analyzed cooperation within teams, national compositions of players, role distribution of heroes and other stats based on information from Dota 2 web forums [2].

Rioult et al. [3] analyzed topological patterns of DotA teams based on area, inertia, diameter, distance and other features derived from their positions and movements of the players around the map to identify which of them are related with winning or losing the game. Drachen et al. used Neural Networks and Genetic Algorithms to analyze and optimize patterns of heroes' movements on the map in DotA [4].

Another direction of research is encounter detection and fight results prediction. Yang et al. [5] applied graph theory to identify patterns in combat hence analyzing teams' tactics and predict fight results with them with 80% accuracy on test data. Schubert et al. [6] build up on this approach and took into account range of attack and spells for each hero to make a better algorithm for encounter detection and team performance evaluation.

Finally, another branch of research in Dota 2 is detection and classification of heros' roles and positions in the game. Gao et. al [7] used Logistic Regression and Random Forest for that purpose and managed to detect hero roles with 75% accuracy for hero ids for both public and professional games and 85 % and 90% accuracy for hero positions respectively. Eggert et al. [8] continued this work and got even better results with 96.15% test accuracy with Logistic Regression.

1.2 Game Outcome Prediction

Most popular topic in applications of Machine Learning to Dota 2 is win probability prediction from team drafts. Conley & Perry were the first to demonstrate the importance of information from the draft stage of the game with Logistic Regression and k-Nearest Neighbors (kNN) [9]. They trained a Logistic Regression classifier on 18,000 examples and obtained 69.8% test accuracy. kNN with custom weights for neighbors and distance metrics with 2-fold cross-validation on 20,000 matches got 67.43% accuracy on cross-validation and 70% accuracy on 50,000 matches in the test set.

Although their work was the first to show the importance of draft alone, the interaction among heroes within and between teams were hard to capture with such a simplistic approach. Agarwala & Pearce tried to take that into account including the interactions among heroes into the logistic regression model [10]. To define a role of each hero and model their interactions they used PCA analysis of the heroes' statistics (kills, deaths, gold per minute etc.). However, their results showed inefficiency of such approach, because it got them only 57% accuracy while the model without interactions got 62% accuracy. Its worth noticing that although the PCA-based models couldn't match predictive accuracy of logistic regression, the composition of teams they suggested looked more balanced and reasonable from the game's point of view. Besides that, they tried to find meaningful strategies with K-Means clustering on end-game statistics but could not find clusters which means that no patterns of gameplay could be detected on their data.

Another approach to that problem of modeling heroes' interactions was proposed by Kuangyan Song et al. [11]. They took 6,000 matches and manually added 50 combinations of 2 heroes to the features set and used forward stepwise regression for feature selection. With 10-fold CV for Logistic Regression on 3,000 matches they got 54% accuracy. They concluded that only addition of particular heroes improves the model while the others might cause the prediction go wrong.

Kalyanaraman was the first to implicitly introduced the roles of the heroes as a feature in the model of win prediction [12]. Author took 30,426 matches filtered by Match Making Ranking (MMR) to select only skilled players and used an ensemble of Genetic Algorithms and Logistic Regression on 220 matches. Logistic Regression alone obtained accuracy of 69.42% and an ensemble with Genetic Algorighm and Logistic Regression approached 74.1% accuracy on the test set. Although it is the highest result among all the articles in the review, lack of AUROC information and the small sample of matches, chosen for the Genetic Algorithm, hampers its reliability.

Another attempt to include interaction among heroes was done by Kinkade & Lim took 62,000 matches with "very high" skill level without leavers and game duration at least 10 minutes [13] and divided it into 52,000 matches for training, 5,000 for testing and 5,000 for validation. On this data they tried Logistic Regression and Random Forest with such feature as pairwise winrate for Radiant and Dire. The feature in theory could capture such relationships as matchup, synergy and countering and each of them increased the quality of the model up to 72.9%. Logistic Regression and Random Forest on picks data only got 72.9% test accuracy for Logistic Regression and overfitted Random Forest which gave them only 67% test accuracy after tuning parameters. It is worth mentioning that their baseline, which included highest combined individual win rate for the heroes, had 63% accuracy.

Several authors expanded the scope of win prediction from draft information to other sources of data from the game. Johansson & Wikstrom trained Random Forest on the information from the game (such as amount of gold for each hero, his kills, deaths assists for each minute etc.) which had 82.23% accuracy at the five minute point [14]. Although such accuracy seem to be very high, that fact that it is based on data from the game events makes its use limited, because it demand real-time data to be practically useful.

The key results from papers described in this sections are summarized in the following table.

Table 1. Key results for Dota 2 win predictions from drafts

Reference	Model	Training (+ Validation) set size	Validation set size	Test Accuracy
Conley & Perry 2013	Logistic Regression	56691	5669	69,80%
	kNN	50000	6691	67,43%
Agarwal & Pearce 2014	Logistic Regression	40000	4000	62,00%
	PCA	40000	4000	57,00%
Song et al. 2015	Logistic Regression	6000	600	58,00%
Kalyanaraman 2014	Logistic Regression	18500	1500	69,42%
	Genetic Algorithm & Logistic Regression	220		74,10%
Kinkade & Lim 2015	Logistic Regression	62000	5000	72,90%
	Random Forest	62000	5000	67,00%

2 Our Projects on Machine Learning in Dota 2

Based on the previous research we have conducted several projects in that field which we would like to describe and discuss at the workshop:

- mining and preparing of large and consistent datasets of DotA 2 matches for creating, testing and comparing Machine Learning algorithms [1];
- paper, introducing Factorization Machines for the task of game outcome prediction, which was presented at the 5th conference on Analysis of Images, Social Networks, and Texts (AIST 2016) [2];
- In-class Kaggle competition for Machine Learning course at Coursera [3];
- hackathon for real-time prediction of the winner during the Dota 2 Shanghai Major [4].

We are eagerly looking forward to share our experience from these projects with other participants of the workshop.

References

1. T. Nuangjumnonga and H. Mitomo, "Leadership development through online gaming," in *19th ITS Biennial Conference: : Moving Forward with Future Technologies: Opening a Platform for All*, (Bangkok), pp. 1–24, 2012.
2. N. Pobiedina and J. Neidhardt, "On successful team formation," tech. rep., 2013.
3. F. Rioult, J.-P. Métivier, B. Helleu, N. Scelles, and C. Durand, "Mining Tracks of Competitive Video Games," *AASRI Procedia*, vol. 8, no. Secs, pp. 82–87, 2014.
4. A. Drachen, M. Yancey, J. Maguire, D. Chu, I. Y. Wang, T. Mahlmann, M. Schubert, and D. Klabajan, "Skill-based differences in spatio-temporal team behaviour in defence of the Ancients 2 (DotA 2)," *Games Media Entertainment (GEM), 2014 IEEE*, vol. 2, no. DotA 2, pp. 1–8, 2014.
5. P. Yang, B. Harrison, and D. L. Roberts, "Identifying Patterns in Combat that are Predictive of Success in MOBA Games," in *Proceedings of Foundations of Digital Games*, (Miami, Florida), pp. 1–8, 2014.
6. M. Schubert, A. Drachen, and T. Mahlmann, "Esports Analytics Through Encounter Detection," in *MIT SLOAN Sports Analytics Conference*, pp. 1 – 18, 2016.
7. L. Gao, J. Judd, D. Wong, and J. Lowder, "Classifying Dota 2 Hero Characters Based on Play Style and Performance," 2013.
8. C. Eggert, M. Herrlich, J. Smeddinck, and R. Malaka, "Classification of Player Roles in the Team-Based Multi-player Game Dota 2," in *Entertainment Computing - ICEC 2015* (K. Chorianopoulos, M. Divitini, J. Baalsrud Hauge, L. Jaccheri, and R. Malaka, eds.), vol. 9353 of *Lecture Notes in Computer Science*, (Cham), pp. 112–125, Springer International Publishing, 2015.
9. K. Conley and D. Perry, "How Does He Saw Me? A Recommendation Engine for Picking Heroes in Dota 2," tech. rep., 2013.

[1] `http://dotascience.com/papers/aist2016`
[2] In print. Draft available at `http://dotascience.com/papers/aist2016/aist2016-ml-dota2-drafts_preprint.pdf`
[3] `https://inclass.kaggle.com/c/dota-2-win-probability-prediction`
[4] `http://dotascience.com`

10. A. Agarwala and M. Pearce, “Learning Dota 2 Team Compositions,” tech. rep., Stanford University, 2014.
11. K. Song, T. Zhang, and C. Ma, “Predicting the winning side of DotA2,” tech. rep., Stanford University, 2015.
12. K. Kalyanaraman, “To win or not to win? A prediction model to determine the outcome of a DotA2 match,” tech. rep., University of California San Diego, 2014.
13. N. Kinkade, L. Jolla, and K. Lim, “DOTA 2 Win Prediction,” tech. rep., University of California San Diego, 2015.
14. F. Johansson, J. Wikström, and F. Johansson, “Result Prediction by Mining Replays in Dota 2,” Master’s thesis, Blekinge Institute of Technology, 2015.