Trouncing in Dota 2: An Investigation of Blowout Matches

Markos Viggiato, Cor-Paul Bezemer
Electrical and Computer Engineering Department
University of Alberta
{viggiato,bezemer}@ualberta.ca

Abstract

With an increasing popularity, Multiplayer Online Battle Arena games where two teams compete against each other, such as Dota 2, play a major role in esports tournaments, attracting millions of spectators. Some matches (so-called blowout matches) end extremely quickly or have a very large difference in scores. Understanding which factors lead to a victory in a blowout match is useful knowledge for players who wish to improve their chances of winning and for improving the accuracy of recommendation systems for heroes. In this paper, we perform a comparative study between blowout and regular matches. We study 55,287 past professional Dota 2 matches to (1) investigate how accurately we can predict victory using only pre-match features and (2) explain the factors that are correlated with the victory. We investigate three machine learning algorithms and find that Gradient Boosting Machines (XGBoost) perform best with an Area Under the Curve (AUC) of up to 0.86. Our results show that the experience of the player with the picked hero has a different importance for blowout and regular matches. Also, hero attributes are more important for blowouts with a large score difference. Based on our results, we suggest that players (1) pick heroes with which they achieved a high performance in previous matches to increase their chances of winning and (2) focus on heroes' attributes such as intelligence to win with a large score difference.

Introduction

The gaming industry has become a multi-billion dollar industry in recent years, experiencing a sharp growth and an expected revenue of $196 billion dollars by 2022 (Webb 2019). Esports (or electronic sports) is an organized form of competitively playing video games, which has played a major role within the gaming industry. Multiplayer Online Battle Arena (MOBA) games, where two teams compete against each other, is a popular genre in esports, with tournaments that offer multi-million dollar prize pools and are watched by millions of spectators (Schubert, Drachen, and Mahlmann 2016; Block et al. 2018). Examples of very popular MOBA games are *Defense of the Ancients 2* (Dota 2) and *League of Legends* (LoL).

Huge amounts of data have been generated from esports, allowing us to extract important insights, which is often referred to as game analytics (El-Nasr, Drachen, and Canossa 2016). A vast body of work has investigated different game aspects, such as game outcome prediction (Ravari, Bakkes, and Spronck 2016; Ravari et al. 2017; Makarov et al. 2017), recommendation systems (Hanke and Chaimowicz 2017; Looi et al. 2018), automatic extraction of game events (Luo, Guzdial, and Riedl 2019), and team encounters (Schubert, Drachen, and Mahlmann 2016).

In this work, we focus on the Dota 2 game, where each team of 5 players must choose one side (Radiant or Dire). By inspecting professional Dota 2 matches, we note that some of them end very quickly or have a very large difference in teams' final scores. We refer to these types of matches as *time blowout matches* and *score blowout matches* (vs. *regular matches*). Although a blowout match might be seen as "expected" (e.g., because one team is considered stronger than the other one), it might also indicate imbalances in the game's gameplay, which can, for example, favor players who chose one team or another (Radiant or Dire) (Gopya 2020). Understanding how blowouts differ from regular matches can be useful knowledge for players who wish to increase their chances of winning.

In this paper, we perform a comparative analysis between blowout and regular matches with regard to the following two aspects: (1) the performance of win prediction models and (2) the explanation of which factors are correlated with a victory. We study 55,287 past Dota 2 professional matches and seek to answer the following research questions (RQs):

RQ1: How well can we predict victory in blowout and regular Dota 2 matches?

We first investigate whether we can find high-performing models to predict victory in different types of Dota 2 matches before identifying the factors associated with victory. We found that XGBoost provides the best performance in blowout and regular matches, with an Area Under the Curve (AUC) of up to 0.86.

RQ2: Which factors are correlated with victory in blowout and regular Dota 2 matches?

Comparing which factors are associated with the victory in blowout and regular matches is important as it can help players focus on specific aspects that increase their chance of winning a match and support new recommendation systems for heroes. Our models show that the up-to-date win rate of the players is an important factor for victory in blowout and regular matches. However, only for score blowouts (matches with a large score difference), heroes' features (e.g., hero's role and the intelligence attribute) are important factors.

Our study makes three major contributions:

- We provide a practical, high-achieving model to predict victory in Dota 2 using only pre-match information (i.e., information available right before the match begins).
- We identified the most important factors that are correlated with victory in blowout and regular matches in Dota 2.
- We provide access to the data analysis code[1] and the data[2] with the up-to-date changelog and historical attribute values of each hero, historical statistics of heroes and players, and the computed features.

Background and Related Work

In this section, we provide a background on the Dota 2 gameplay and outline prior work on Dota 2.

Dota 2 Gameplay. Dota (*Defence of the Ancients*) 2 is an action Real-Time Strategy (RTS) game,[3] sometimes referred to as a Multiplayer Online Battle Arena (MOBA) game because it combines elements from the RTS genre with tower defense elements (Rondina 2018). Dota 2 is the successor of *Defense of the Ancients*, a mod for *Warcraft 3*.

A Dota 2 match consists of a battle between two teams (Radiant and Dire). Each team is composed of 5 players, each one controlling a unique character (hero). The ultimate goal of the game is to destroy the opponent team's ancient, which is the main structure in Dota 2. The idea is that each team should move along three lanes to reach the enemy's ancient while facing battles along the way with different creatures and having to destroy the opponent team's towers beside the battles with opponent's heroes and other creatures. The score of each team corresponds to the death count of that team, i.e., the number of times all of a team's entities, including its heroes and non-playable characters, killed an opponent's character. The Dota 2 match can be played in different modes, which affect the way the players pick heroes. For instance, the *All Pick* mode allows a player to pick a hero from the entire pool of available heroes, while in the *Captains* mode the team's captain picks the heroes for the team and bans heroes (which cannot be picked by any team). The *Captains* mode is the standard mode for tournaments (i.e., professional matches).

Each Dota 2 hero has attributes and specific abilities. Primarily, heroes are categorized based on three attributes: strength, agility, and intelligence. Each attribute has a base value, with which the hero starts the match, and a gain factor, which is by how much the base values increase as the hero levels up. Abilities refer to unique spells that heroes can use, which can be in different forms, such as to damage the opponent or help allies (Demediuk et al. 2019). Abilities can be developed along the match as the player gains experience points and gold (Eggert et al. 2015; Drachen et al. 2014). Thus, as the player advances levels, their hero abilities can be improved or even new abilities might appear (Drachen et al. 2014). Each hero has one or more roles in the match and the player should be aware of the chosen hero's role to make the most effective use of it. There is a total of 9 roles, such as: *Carry* (heroes that have the greatest increase in power throughout the match, being responsible, many times, for the team victory); *Support* (heroes that are responsible for supporting their partners by keeping them alive and allowing them to earn more experience points and gold); and *Durable* (heroes that are able to resist to a lot of damage from the enemy and usually have large amounts of health generation) (Semenov et al. 2016).

Dota 2 studies. Several studies addressed different aspects of Dota 2. Katona et al. (2019) built a deep neural network to predict a hero's death in a Dota 2 match within a window of 5 seconds using gameplay features and professional/semi-professional matches. Their findings show the model has a precision of 0.377 and a recall of 0.725 when predicting the death of any of the 10 players within the next 5 seconds. Luo, Guzdial, and Riedl (2019) proposed an accessible method to extract events from Dota 2 gameplay videos with a Convolution Neural Network (CNN). Using techniques such as transfer learning, zero-shot and network pruning, the method is capable of extracting 10 events, such as the use of the *Black King Bar* item and tower destructions. Demediuk et al. (2019) provided a method to classify and label individual roles for each hero in Dota 2 using non-performance metrics of the types: map movement, resource priority, and ability prioritisation. Hanke and Chaimowicz (2017) proposed a recommendation system to support hero selection. The authors used association rules to suggest heroes and evaluated the system with a neural network capable of predicting the winner team. Their recommendation system presented 74.9% success rate. Looi et al. (2018) also proposed a recommendation system but for Dota 2 items. The system is based on commonly used purchasing strategies. The authors used 3 recommender systems, based on rules, logistic regression, and logistic regression enhanced with clustering. While the first two systems achieved accuracies between 66.5% and 87.1% for highly-skilled players, the third system had a slight improvement compared to the second system.

These prior works focused on recommendation systems, clustering techniques to identify hero roles, automatic extraction of events and hero death prediction within a limited time window. In our work, we focus on predicting the winning team of Dota 2 matches using only pre-match information and prediction explanation by indicating the most important features for the prediction. Furthermore, to the best of our knowledge, prior work on video games has not investigated how blowout matches differ from regular matches in terms of win prediction performance and important features.

Game Outcome Prediction. Predicting the outcome of competitive games, such as Dota 2 and Starcraft, is a popular research topic given the complexity of such games.

[1] https://github.com/asgaardlab/dota2-prediction-models

[2] https://doi.org/10.5281/zenodo.3890315

[3] https://dota2.gamepedia.com/Dota_2

Prior works provided prediction models for different types of games. Ravari, Bakkes, and Spronck (2016) built several models to predict the winner in StarCraft. The authors used replays from StarCraft competitions and computed time-dependent and time-independent features for the models. Their findings show that time-independent features worsen the prediction performance compared to using time-dependent features and economic aspects of StarCraft matches are the strongest predictors for victory. Ravari et al. (2017) investigated the outcome prediction for Destiny, an online Multiplayer First-Person Shooter game. The authors developed two types of models: one based on the game mode and one which is mode-independent. Specific-mode models performed better and the authors showed that, in player(s) versus player(s) mode, score-per-life, kill-death ratio, and kill-death-assists are the most important predictors. Semenov et al. (2016) compared the performance of different machine learning algorithms to predict the outcome of Dota 2 based on the draft information (picked heroes and the number of hero roles in each team). Their dataset included matches played in three modes and was split based on players' skill levels. Their findings show that factorization machines have the best performance and the model's prediction accuracy depends on the skill level of the players.

Differently from the aforementioned works, we leverage historical statistics about heroes, players, and the combination hero-player (in addition to draft information) to compute features and build prediction models for Dota 2 outcome. We also restrict our dataset to professional matches only as the game is played seriously in competitions and this brings more confidence to our analysis and the drawn conclusions and implications. Furthermore, our aim is not only to provide a good-performing win prediction model, but also explain the prediction by means of feature importance and compare the important factors for the prediction in blowout and regular matches.

Methodology

In this section, we describe the processes to collect and clean the data, compute the features for the machine learning models, build and evaluate the models, and compute feature importance. Figure 1 presents an overview of our methodology.

Collecting data

We collected all the Dota 2 match data from the OpenDota platform,[4] an open source platform developed and maintained by volunteers. All the information available within OpenDota is obtained from the Steam platform, the largest online game distribution platform. Note that only users who have the "Expose public match data" option enabled in the Dota 2 client have their match data collected by OpenDota.

The data collection process is composed of two parts. Initially, we collected the match identifiers of the Dota 2 matches using the SQL query functionality of OpenDota. We then used the identifiers to make requests to the OpenDota API[5] using the *matches* endpoint. We ended up with 86,925 professional matches (JSON format) played between October 26th 2012 and May 6th 2020. These files contain detailed information about the match, such as the start time, the duration, and players' information (experience points and gold earned, number of kills, deaths, assist, among others). They also contain whether the Radiant team won the match or not, which corresponds to the output of our prediction model (the truth label).

Cleaning data

We removed 768 matches which did not use the *Captains mode* or the *All Pick mode*. In fact, the *Captains mode* is the standard mode for competitions and represents 99% of our data. As we investigate win prediction models and prediction's explanatory factors, it is important to consider only matches in which no player abandoned the match to be consistent. We removed 30,870 matches in which at least one player abandoned the match. In the end, we studied the 55,287 remaining matches. These matches were split into 3 groups: time blowout, score difference blowout, and regular. Time blowout refers to the blowout matches with an extremely short duration. To compose this group, we selected the bottom 10% of the matches based on the match duration, which has durations between 6 and 21 minutes. Score difference blowout refers to matches with a large score difference between teams. Although the winning team does not necessarily have a higher score than the opponent, such a large difference might represent a certain "ease" for the winner. To compose this group, we selected the top 10% matches based on the score difference, which corresponds to score differences between 23 and 65. The time and score difference blowout groups ended up with 5,528 matches each. Regular refers to matches with duration and score difference values around the median. We selected matches within the 1st and 3rd quartiles (25% - 75%) for both duration and score difference. We further removed duplicates and overlapping matches between the regular and time/score difference blowout groups The regular group ended up with 36,348 matches. The class distribution is fairly balanced for all groups, with percentages of class 1 (Radiant's victory) of: 50.4%, 59.1%, and 50.7% for regular, time blowout, and score difference blowout groups, respectively. We followed the same steps for all the 3 groups: computed features, built and evaluated the models, and computed feature importance.

Computing features

We extracted a feature set with 457 features that capture relevant information to build the win prediction models. Although we could use post-match features (i.e, features collected for the match that we are predicting the winner for after its end, such as the amount of gold each player had), this would be more trivial and less useful in practice. Therefore, we collected only pre-match features, i.e., all the information which is available right before the match starts, such as attributes of the picked heroes and historical statistics of heroes and players. Table 1 presents an overview of the features.

We have 5 main categories of features: match feature, team feature, player features, hero features, and hero-player

[4]https://www.opendota.com/

[5]https://docs.opendota.com/

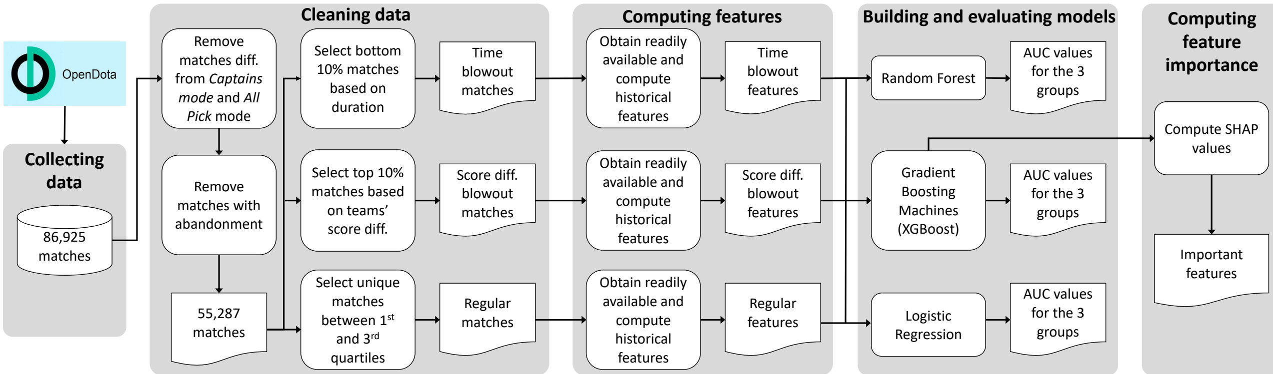


Figure 1: Study methodology.

combined features. Hero-player features capture the experience of the player with the picked hero based on information collected from previous matches where the player picked the same hero as in the current match. All the player and hero-player features use up-to-date data, as well as the *hero_winrate_** feature, which means data covering professional matches upto the date of the match being evaluated. For up-to-date features, we computed the average of the feature values from past matches.

Note that hero attributes (e.g., *strength_**, *agility_**, and *intellig_**) change over time as new Dota 2 versions are released. Therefore, we collected the changelog for each hero[6] and we used the proper attribute value according to the version of the match being evaluated (we have access to the match version from the collected JSON files from OpenDota). For the categorical feature *heroes_**, we adopted the one-hot encoding method since the algorithms require all inputs to be numerical. This feature represents a large portion of the feature space due to the large number of heroes, being a 119-dimensional binary array (one array for the Radiant team and another for the Dire team). In addition, *roles_** is a numeric array of length 9, since there are 9 different roles: *carry*, *nuker*, *initiator*, *disabler*, *durable*, *escape*, *support*, *pusher*, and *jungler*. There is one *roles_** feature array for Radiant and one for Dire. Throughout the text, we refer to them with the specific role name properly appended. For all features except *heroes_**, *roles_**, and *first_pick* there are five values, each corresponding to one team's player/hero. In addition, note that all features with the symbol "*" have a Radiant and a Dire version, indicated throughout the text with the suffixes "r" and "d" when appropriate.

Building and evaluating models

We model the winner prediction as a binary classification, where the Radiant team either wins (class 1) or loses (class 0). Note that, when the Radiant team loses, the opponent team (Dire) wins, as there are no ties in Dota 2. We chose three learning algorithms for the classification task: gradient boosting machines (Friedman 2002), random forest (Breiman 2001), and logistic regression (Cox 1958). We performed 10-fold cross-validation to evaluate the models using the Area Under the Receiver Operating Characteristic Curve (AUC) metric and a grid search to find the optimal values for XGBoost and Random Forest hyper-parameters. The AUC measures the classifier's capability of distinguishing between a positive class (Radiant win) and a negative class (Radiant defeat) and ranges from 0.5 (random guessing) to 1 (best classification performance). We also made sure to avoid data leakage by not having overlap between the training and testing sets. We implemented our models using the scikit-learn package.

Gradient boosting machines is an ensemble learning algorithm, which combines "weaker" models into a stronger final model. At each iteration, one model is built on the errors of the previous model. Ultimately, the contribution of each base model to the final one is found by minimizing the overall error of the final model. This algorithm is robust to features on different scales (therefore, we do not need to normalize or standardize the feature values) and can model non-linear relationships between the features and the output. In this work, we adopt a scalable and high-performing implementation of gradient boosting machines, known as XGBoost (Chen and Guestrin 2016), which can automatically handle missing values in the data. Random forest is also an ensemble learning algorithm which is based on several decision trees. This algorithm is capable of avoiding overfitting (which is common in simple decision trees). Regarding missing values, we used data imputation, replacing the missing values with the median of the feature. Logistic regression is a classification algorithm that models the relationship between the features and provides the probability of a specific output. For this algorithm, we also adopted data imputation with the median value for missing values. Furthermore, we normalized the data so it is scaled to a fixed range.

Computing feature importance

Feature importance is a common method to explain machine learning models' predictions, which is based on computing the impact that each feature has on the outcome (Holzinger 2018). However, since features interact with each other in many complex ways, it might get very difficult to compute feature importance. Shapley values (Shapley 1953) is

[6]https://dota2.gamepedia.com/Heroes

Feature	High-level category	Type	Description
first_pick	Match feature	Boolean	Indicates if Radiant was the first to pick in the draft phase
heroes_*	Team feature	Categorical	Heroes picked by the Radiant or Dire team for the match
player_winrate_*	Player feature	Numeric	Up-to-date win rate of each player
roles_*	Hero feature	Numeric array	Number of heroes of each role
strength_*	Hero feature	Numeric	Strength the hero has at the start of the match
agility_*	Hero feature	Numeric	Agility the hero has at the start of the match
intellig_*	Hero feature	Numeric	Intelligence the hero has at the start of the match
strength_gain_*	Hero feature	Numeric	The strength factor that the hero gains per level
agility_gain_*	Hero feature	Numeric	The agility factor that the hero gains per level
intellig_gain_*	Hero feature	Numeric	The intelligence factor that the hero gains per level
health_*	Hero feature	Numeric	Amount of health of the hero at the start of the match
health_regeneration_*	Hero feature	Numeric	Amount of health the hero regains each second
move_speed_*	Hero feature	Numeric	The speed at which the hero can move over a second
hero_winrate_*	Hero feature	Numeric	Up-to-date win rate of the hero
hp_winrate_*	Hero-player feature	Numeric	Up-to-date win rate of the pair hero-player
hp_xp_min_*	Hero-player feature	Numeric	Up-to-date earned experience/min of the pair hero-player
hp_gold_min_*	Hero-player feature	Numeric	Up-to-date earned gold/min of the pair hero-player
hp_death_min_*	Hero-player feature	Numeric	Up-to-date deaths/min of the pair hero-player
hp_taken_damage_min_*	Hero-player feature	Numeric	Up-to-date damage/min received by the pair hero-player
hp_kill_min_*	Hero-player feature	Numeric	Up-to-date number of kills/min of the pair hero-player
hp_assist_min_*	Hero-player feature	Numeric	Up-to-date number of assists/min of the pair hero-player
hp_caused_damage_min_*	Hero-player feature	Numeric	Up-to-date damage/min caused by the pair hero-player
hp_heal_min_*	Hero-player feature	Numeric	Up-to-date healing/min of the pair hero-player

Table 1: Features used in our models. *r or d: throughout the paper we use both suffixes to refer to the Radiant team or the Dire team and we use the * symbol to refer to the feature for both teams. All numeric features are computed for each hero or player in a team, so there are five values for each of those.

an explanation method based on game theory and it provides theoretically optimal values and a fair division scheme of the overall importance distribution among features. In this paper, we adopted SHAP (SHapley Additive exPlanations) (Lundberg and Lee 2017) to compute feature importance. SHAP is a state-of-the-art technique for interpretability and explaining predictions (Wang et al. 2019) and it provides an approximation of the theoretically optimal Shapley Values.

For the categorical *heroes_** feature, we followed the recommendations[7] of the SHAP developers: we summed the SHAP values for each one-hot encoded feature so that we can provide the importance for the high-level feature, i.e., to what extent the hero itself impacts the prediction. Note that a more detailed analysis of the importance of each hero individually can be performed, but we opted to present the results of the high-level feature only due to space restrictions. We also followed developers' recommendation[8] regarding cross-validation: we performed a sensitivity analysis and, since there was almost no variation in the important features for each fold, we report the SHAP values of only one fold.

RQ1: How well can we predict victory in blowout and regular Dota 2 matches?

We are able to predict the winning team with a high performance for score and time blowout matches. Table 2 presents the AUC values obtained for each model and for each group. As we can see, XGBoost presented the best performance for all the three groups, with an AUC varying from 0.65 (regular matches) to 0.86 (time blowout matches). Logistic regression presented better performance than random forest for all groups, with AUC values of 0.61 against 0.60 (regular), 0.79 against 0.74 (time blowout), and 0.77 against 0.70 (score difference blowout).

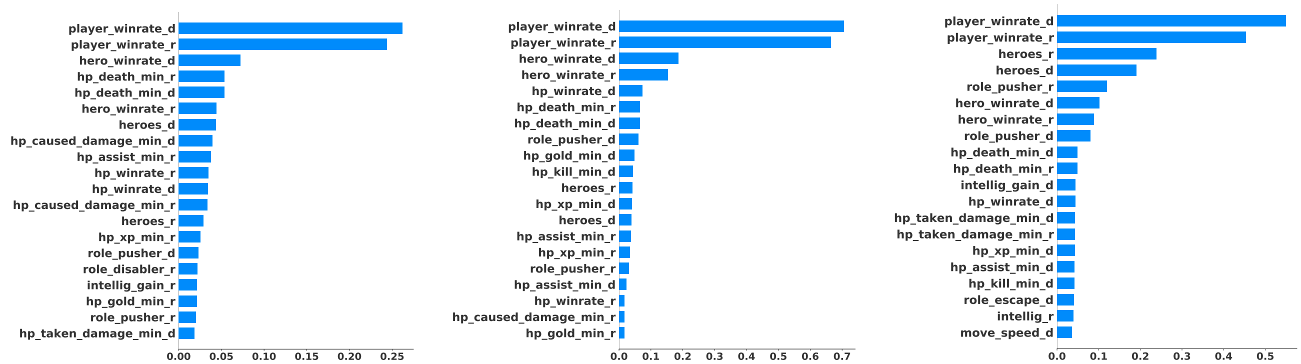
Model	Regular	Time blowout	Score blowout
XGBoost	**0.65**	**0.86**	**0.78**
Random Forest	0.60	0.74	0.70
Logistic Regr.	0.61	0.79	0.77

Table 2: AUC of win prediction models.

The lower performance of the models for regular matches is understandable given the difficulty of the prediction task using only pre-match information. We performed some experiments with a different set of features aggregated from the existing ones (e.g., using the mean or the median win rate for players in a team rather than using each player's win rate). However, we obtained marginally worse AUCs, therefore we kept the current feature set. One possible reason for the better performances obtained for the time and score difference blowout groups might be that the match information is a more clear/stronger indicators of the winning team. Furthermore, most of our hero-player features reflect the team scores, therefore, a larger feature set may be necessary to capture other aspects of the match.

[7]https://github.com/slundberg/shap/issues/397

[8]https://github.com/slundberg/shap/issues/337



(a) Impact on Radiant's victory (regular)

(b) Impact on Radiant's victory (time blowout)

(c) Impact on Radiant's victory (score blowout)

Figure 2: Feature importance (by means of SHAP values) for all groups of matches.

RQ2: Which factors are correlated with victory in blowout and regular Dota 2 matches?

In this section, we discuss the most important factors for a team's victory using SHAP to compute the feature importance. We selected the best performing model (XGBoost) to analyze the important features. SHAP assigns an importance value to each feature in a single prediction. By computing the average importance of each feature across all predictions (e.g., testing set), we are able to obtain the overall importance of each feature on the victory prediction. Figure 2 presents the feature importance plots provided by SHAP.

The up-to-date win rate of the players of both teams is the most important factor for all groups. As we can see in Figure 2, the top-2 most important features is the up-to-date win rate of players (*player_winrate_**). The up-to-date win rate of the heroes (*hero_winrate_**) is also important for all groups. However, hero win rate plays a more important role in regular and time blowout matches rather than in score blowout matches. Furthermore, we found that win rate of players and heroes of the opponent team (the Dire team according to our modelling) is slightly more important than hero and player win rates of the team we are predicting the victory for (Radiant, in our case). This shows that the statistics of the opponent team matter and the teams should take the opponent team's information into account when planning their strategy. We further found out that Dire features are negatively correlated with the prediction output (Radiant's victory), as expected. For instance, we observed that a higher Dire's hero/player win rate decreases the likelihood of a Radiant's victory and a lower Dire's hero/player win rate increases the likelihood of a Radiant's victory.

Hero and team features are more important for score difference blowout matches compared to regular and time blowout matches. Figure 2(c) shows that the choice of the hero itself is a key feature in score blowout matches, as shown by the *heroes_** feature in the top-4. In addition, the pusher hero role (*role_pusher_**) is a key factor for the Radiant's victory in score blowout matches together with other hero attributes, such as the intelligence gain of Dire heroes. Note that those features are much less important for regular and time blowout matches, as shown by Figures 2(a) and (b).

Hero-player combined features are more important in regular and time blowout matches. Features that capture the player experience with the picked hero (i.e., hero-player features) are important for regular matches and somewhat important for time blowout matches. However, they do not play a major role in the prediction for score difference blowout matches. If we look at the top-10 features for each group, regular matches have 5 hero-player features (*hp_death_min_**, *hp_caused_damage_min_d*, *hp_assist_min_r*, and *hp_winrate_r*) and time blowout matches also have 5 hero-player features (*hp_winrate_d*, *hp_death_min_**, *hp_gold_min_d*, and *hp_kill_min_d*). On the other hand, score blowout matches have only 2 hero-player features among the top-10 (*hp_death_min_**).

Implications of Our Study

Although many factors are involved with the ideal composition of the team and winning a Dota 2 match, some factors are more strongly associated with the victory depending on the type of the match. First, players should focus on choosing heroes with high up-to-date win rates, since hero win rates by themselves seem to be a key factor and are associated with quick victories. Furthermore, our results indicate that the experience of the team's players with the hero to be picked is an important factor. Players should focus on heroes with which they achieved a good performance in previous matches, such as a low number of deaths per minute, to increase their chances of winning. Second, hero attributes (e.g., intelligence) and the use of heroes with the pusher role seem to be associated with a victory with a large score difference. Finally, our work provides the foundations for new recommendation systems based on the hero choice. For instance, in score difference blowout matches, heroes' roles and attributes are key factors for the victory. Such information can be used to increase the accuracy of recommendation systems for heroes in Dota 2.

Conclusion

We studied 55,287 past Dota 2 matches to compare blowout and regular matches regarding win prediction performance and important features for the prediction. We found out that XGBoost performs best, with a maximum AUC of 0.86 for time blowout matches. We also identified that the top-4 features used for prediction are very similar, but hero-player statistics are more important for regular and time blowout matches, and heroes' roles and attributes are more important for score difference blowout matches. Our findings are a first important step towards better prediction models for Dota 2 matches using only pre-match information, which is a difficult task. These models have a wide range of applications, such as the prediction of the winner team during the draft phase and the improvement of recommendation systems for heroes. Additional information from the matches (e.g., more statistics of players) can increase the prediction performance and provide more insights about the important factors.

References

Block, F.; Hodge, V.; Hobson, S.; Sephton, N.; Devlin, S.; Ursu, M. F.; Drachen, A.; and Cowling, P. I. 2018. Narrative bytes: Data-driven content production in esports. In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video*, 29–41.

Breiman, L. 2001. Random forests. *Machine learning* 45(1):5–32.

Chen, T., and Guestrin, C. 2016. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785–794.

Cox, D. R. 1958. The regression analysis of binary sequences. *Journal of the Royal Statistical Society: Series B (Methodological)* 20(2):215–232.

Demediuk, S.; York, P.; Drachen, A.; Walker, J. A.; and Block, F. 2019. Role identification for accurate analysis in dota 2. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 130–138.

Drachen, A.; Yancey, M.; Maguire, J.; Chu, D.; Wang, I. Y.; Mahlmann, T.; Schubert, M.; and Klabajan, D. 2014. Skill-based differences in spatio-temporal team behaviour in defence of the ancients 2 (DotA 2). In *2014 IEEE Games Media Entertainment*, 1–8. IEEE.

Eggert, C.; Herrlich, M.; Smeddinck, J.; and Malaka, R. 2015. Classification of player roles in the team-based multi-player game dota 2. In *International Conference on Entertainment Computing*, 112–125. Springer.

El-Nasr, M. S.; Drachen, A.; and Canossa, A. 2016. *Game analytics*. Springer.

Friedman, J. H. 2002. Stochastic gradient boosting. *Computational statistics & data analysis* 38(4):367–378.

Gopya. 2020. Why the dire might be better than the radiant after dota 2 patch 7.24 (last visited: April 30^{th}, 2020).

Hanke, L., and Chaimowicz, L. 2017. A recommender system for hero line-ups in moba games. In *Thirteenth Artificial Intelligence and Interactive Digital Entertainment Conference*.

Holzinger, A. 2018. From machine learning to explainable ai. In *2018 World Symposium on Digital Intelligence for Systems and Machines (DISA)*, 55–66. IEEE.

Katona, A.; Spick, R.; Hodge, V. J.; Demediuk, S.; Block, F.; Drachen, A.; and Walker, J. A. 2019. Time to die: Death prediction in dota 2 using deep learning. In *2019 IEEE Conference on Games (CoG)*, 1–8. IEEE.

Looi, W.; Dhaliwal, M.; Alhajj, R.; and Rokne, J. 2018. Recommender system for items in dota 2. *IEEE Transactions on Games*.

Lundberg, S. M., and Lee, S.-I. 2017. A unified approach to interpreting model predictions. In *Advances in neural information processing systems*, 4765–4774.

Luo, Z.; Guzdial, M.; and Riedl, M. 2019. Making cnns for video parsing accessible: event extraction from dota2 gameplay video using transfer, zero-shot, and network pruning. In *Proceedings of the 14th International Conference on the Foundations of Digital Games*, 16. ACM.

Makarov, I.; Savostyanov, D.; Litvyakov, B.; and Ignatov, D. I. 2017. Predicting winning team and probabilistic ratings in "dota 2" and "counter-strike: Global offensive" video games. In *International Conference on Analysis of Images, Social Networks and Texts*, 183–196. Springer.

Ravari, Y. N.; Bakkes, S.; and Spronck, P. 2016. Starcraft winner prediction. In *Twelfth Artificial Intelligence and Interactive Digital Entertainment Conference*.

Ravari, Y. N.; Spronck, P.; Sifa, R.; and Drachen, A. 2017. Predicting victory in a hybrid online competitive game: The case of destiny. In *Thirteenth Artificial Intelligence and Interactive Digital Entertainment Conference*.

Rondina, S. 2018. What is dota 2? (last visited: April 30^{th}, 2020).

Schubert, M.; Drachen, A.; and Mahlmann, T. 2016. Esports analytics through encounter detection. In *Proceedings of the MIT Sloan Sports Analytics Conference*, volume 1, 2016. MIT Sloan Boston, MA.

Semenov, A.; Romov, P.; Korolev, S.; Yashkov, D.; and Neklyudov, K. 2016. Performance of machine learning algorithms in predicting game outcome from drafts in dota 2. In *International Conference on Analysis of Images, Social Networks and Texts*, 26–37. Springer.

Shapley, L. S. 1953. A value for n-person games. *Contributions to the Theory of Games* 2(28):307–317.

Wang, D.; Yang, Q.; Abdul, A.; and Lim, B. Y. 2019. Designing theory-driven user-centric explainable ai. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15.

Webb, K. 2019. The $120 billion gaming industry is going through more change than it ever has before, and everyone is trying to cash in (last visited: May 31^{st}, 2020).