



# Performance Index: A New Way To Compare Players

Simon Demediuk, Athanasios Kokkinakis, Moni Sagarika Patra, Justus Robertson, Ben Kirman, Alistair Coates, Alan Chitayat, Jonathan Hook, Isabelle Nolle, Oluseyi Olarewaju, Daniel Slawson, Marian Ursu, Florian Block, Anders Drachen  
University of York, Digital Creativity Labs  
{first}.{last}@york.ac.uk

## 1. Introduction

In team sports, it can be a non-trivial task to compare players and understand how much they contribute and in what way, especially when those players are fulfilling significantly different roles on a team. What metric do you use to compare a Pitcher to a Batter? A Point Guard to a Center? Understanding what degree players are performing and how they compare to one another allows for greater depth in storytelling for the broadcasting team. It also provides audiences with overall performances across team positions.

This problem is forefront in esports. Esports are video games played competitively with an estimated viewership of 495 million, comprising a rapidly growing 1.5bn USD market [1]. The players range from pre-teens to mature adults. Prize pools have reached the multi-million USD level, and esports fans fill Olympic arenas and events attract millions of viewers online [1]. Spanning multiple game genres, esports include digital versions of traditional sports, however the most popular esports games by prize pool or viewership are the Multiplayer Online Battle Area (MOBA), First Person Shooters (FPS), and Real-time Strategy (RTS) games. Other genres are also found in esports but are substantially less popular. MOBAs are team-based magic-themed fantastical games, whereas shooters are graphically realistic and focus on small-team operations and tactics. Contrary to real-time strategy games, there is no unit or building construction in MOBA or FPS games. Therefore, the strategy revolves around team-based tactics and individual character development.

In the past decade, work in esports has given rise to an entire subdomain of Data Science: esports analytics (broadly also referred to as esports data science). Utilizing techniques across game analytics [2] and sports analytics, the focus of data scientists working with esports is to analyze the detailed and voluminous data that these digital games generate for the purpose of locating trends and patterns in the data, towards informing players, teams, coaches, business models as well as content production [3]. Professional teams today employ analysts and data scientists to extract value from the data produced by their own team and those of professional players and competitors worldwide, similar to analysts working in physical sports [4].

Esports broadcasting and coverage follow a similar structure to traditional sports. However, due to their virtual nature and incredible complexity, an extensive and detailed amount of data is available about in-game actions not currently accessible in traditional sport. Jointly, this provides not only an



opportunity and a need to incorporate novel insights about complex aspects of gameplay into the audience experience, enabling more in-depth coverage for experienced viewers, and increased accessibility for newcomers [5]. The latter point is crucial to attracting new players and spectators, as there is a considerable barrier to entry in esports due to their underlying complexity [5, 6].

Using the massively popular esports game Dota 2 [7] as a case study, this paper presents the Performance Index (PI), a novel and unified metric designed to not only fairly compare and understand the performance of each player on a team irrespective of their role, but also provides deep insights into the roles players perform. The PI also extends to other contexts such as for use by professional teams for team and player analysis, swings in performance, and recruitment. Dota 2 is classified as a MOBA, with examples including *League of Legends* [8] and *Smite* [9]. In MOBAs, the concept of role is foremost in the mechanics of the game, with each player developing their virtual characters in specific ways depending on their role and performing widely different actions in the game. This makes evaluating the performance of players notably complex in MOBAs, as compared to FPS games [10-14]. Currently, there is no direct way to compare players in a fair manner, i.e., considering their role as part of the team. This in turn impacts the ability of commentators to quickly compare the performance of players and teams and increases the barrier of entry for novice viewers and players. This forms the primary motivation for the research presented here.

The Performance Index considers the playstyle of each player and operates in real-time, facilitating real-time storytelling and audience engagement. The PI has been successfully deployed at multiple major Dota 2 tournaments across Europe in 2020. Three different platforms have been used to deliver the index, including broadcast overlays, Twitch overlays, and a companion mobile app. While developed for Dota 2, the principles behind the Performance Index is not limited to esports but can be applied in a similar way to other team-based, multi-role sports such as basketball, baseball, and football.

## 2. Background

In this section, we overview Dota 2 and the gameplay for readers unfamiliar with the game. We also provide a brief explanation of the Random Forest Algorithm, which is used as part of the Performance Index calculation. Finally, we present an overview of work performed in the area of esports analytics.

### 2.1. Dota 2

Dota 2 [7] is a MOBA game. In Dota 2, two teams of five players, called Radiant and Dire, each controlling a different hero, battle each other within a closed arena environment in a race to destroy the opponent's ancient, the core structure in each team's home base. The game map is broken down into several distinct areas. At the top right and bottom left corners are the home bases of the two teams. Diagonally in the middle of the map is a river that separates the two teams' home territories. Three roads, called lanes, cross between the two bases. These are designated by the names, Safelane, Offlane, and Midlane (Figure 1). In between the lanes, there are jungle areas that contain buildings with specific properties (shrines, outposts, shops, effigy buildings), as well as neutral enemies referred to as neutral creeps or neutral monsters. The home base consists of various buildings with different functionalities, notably the ancient and two towers that can damage nearby enemy heroes and which must be destroyed before the ancient can be destroyed.



The home bases also contain a fountain, which acts as a location for reviving heroes that have been eliminated. The arena itself is thus an active component of Dota 2. The Dota 2 arena is under a fog of war mechanic, meaning that only areas the player heroes can see, or where they have placed "wards", items that can be placed to provide vision in an area or in a limited area around friendly buildings and creeps, are fully visible.

Lanes are protected by sequences of towers that will target and damage enemies that approach, and which are controlled by the team in whose half of the map the tower is placed. Every 30 seconds, a group of automated (computer-controlled) creatures (creeps) are spawned in each lane for each team. These groups are referred to as creep waves. A creep wave consists of various melee and distance-based creatures. These creep waves walk down the lanes and engage rival creep waves, enemy heroes, and enemy towers.

There are at the time of writing 119 heroes in Dota 2. Each is different, with different statistics and abilities, making for a complex data space. Prior to a match, each player selects one hero, and in tournaments, it is possible for teams to block some heroes from being picked by the opposing team. Understanding how different heroes synergise with each other, and how they can counter the capabilities of enemy team heroes, is a vital part of securing a good starting position in Dota 2 [12, 15].

Players control their heroes to engage in combat using abilities and items, to kill enemy creeps, heroes, towers, and neutral monsters until the enemy ancient is exposed and can be destroyed. When a hero kills an enemy creep, hero, tower, or neutral monster, they are awarded experience and gold. Experience allows the hero to level-up to gain access to improve their abilities or gain new ones, and gold (an in-game currency) enables the hero to purchase items from the shops in the game. Such items can boost the stats of a hero or even confer entirely new abilities, e.g., healing, teleportation, or similar.

In Dota 2 each player performs a different role as part of the team. Different models for how to break down the roles available in Dota 2 have been proposed, and there are no officially recognized definitions. However, a common definition refers to roles as "positions" and label these one through to five. Positions are generally comparable to the concept of positions in sports such as Basketball, e.g., point guard, center, and forward. Each position focuses on different aspects of the game and has different operational and tactical goals. Positions 1 to 3 are commonly referred to as Carry positions. These are heroes that are initially weak in matches but become very powerful in the later phases of a match. This has given rise to the term "carry", as these heroes can carry a team and secure victory. The carry positions broadly focus on dealing damage to opponents and destroying towers. The specific goals of a carry vary across the three positions. Positions 4 and 5 are referred to as Support positions, and they focus on supporting and helping other characters, for example, via healing friendly heroes or stunning enemy heroes, as well as providing vision on the battlefield. They also try to give carries opportunities to earn experience and gold to unlock more powerful abilities and items. Each position can be played differently in terms of how the core goals are achieved depending on the hero selected (see Section 3.2), and there can even be substantial variation in how the goals are defined depending on the strategy of the team in question. Comparing players in different positions and even in the same position is a non-trivial task. Currently, there is no easy way to compare players in a fair manner, i.e., considering their role as part of the team.

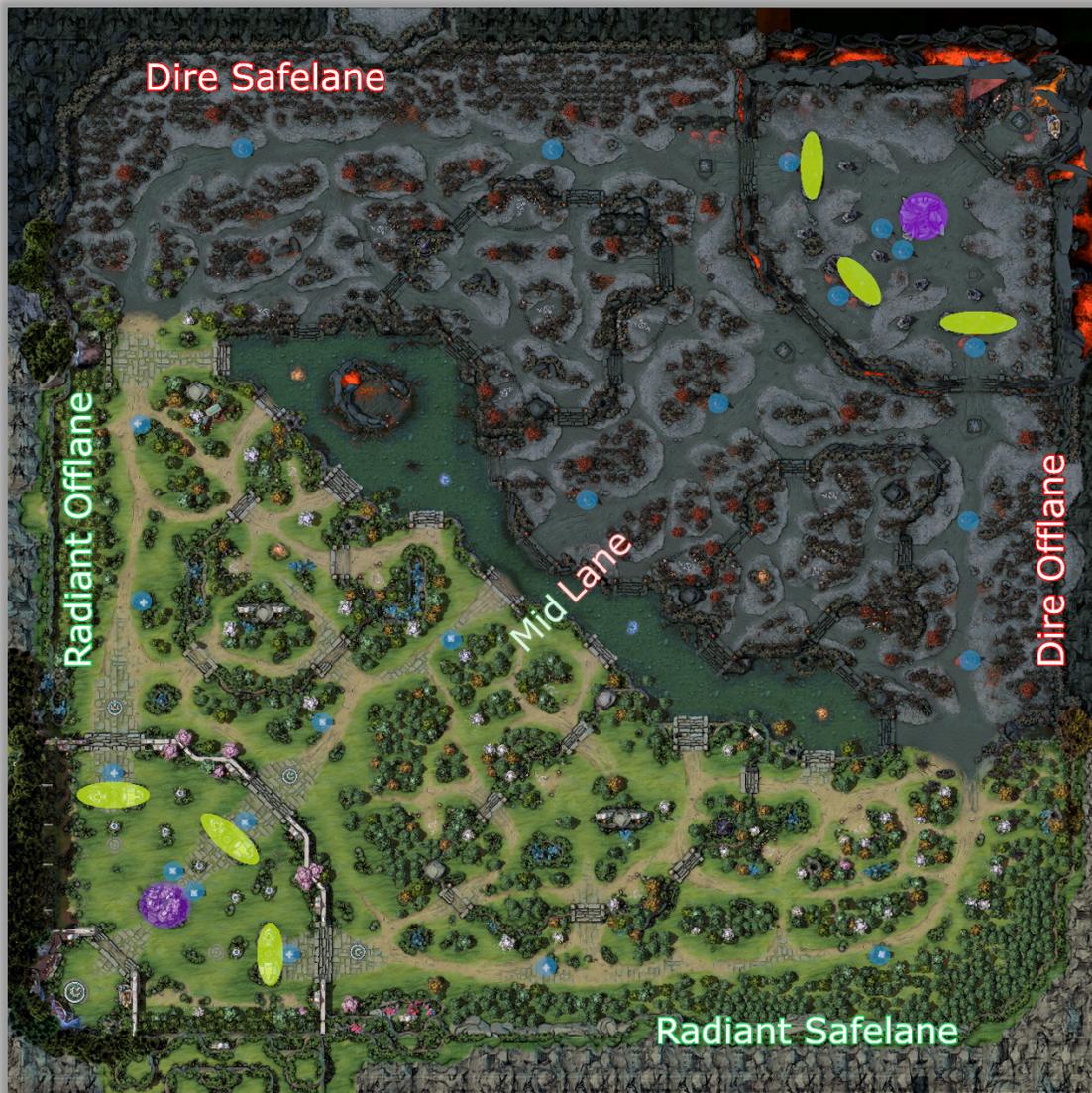


Figure 1: The map of Dota 2. Towers are highlighted in blue, Racks ("barracks") are highlighted in yellow, and the Ancients are highlighted in purple. The three lanes are marked with the names associated with their corresponding teams

## 2.2. Random Forest Algorithm

The Random Forest Algorithm is a supervised learning algorithm used in Machine Learning for classification and regression tasks. In this paper, we will be using it as a classifier. This section is designed to give a basic understanding of the algorithm. The algorithm is used for calculating the Performance Index is explained later in Section 4.

The Random Forest Algorithm was first developed by Breiman [16] and later expanded on by Cutler et al. [17]. The Random Forest Classifier is an ensemble method that builds multiple decision trees. Decision trees are made using binary recursive partitioning [18]. The basic principle is that



each tree within the random forest produces a prediction based on its input data. Based on the "wisdom of crowds", the prediction with the highest number of entries amongst the trees is the prediction for entire random forest. Each tree within the forest is designed to have a low correlation with other trees, as uncorrelated models produce ensemble predictions of higher accuracy than individual predictions.

To build uncorrelated decision trees two methods are employed, Bagging and Feature Randomness, as part of the Binary Recursive Partitioning. Decision trees are sensitive to training data, as small variances within the training data result in different decision trees. Bagging is the process by which each tree in the random forest ensemble will randomly select training data with replacement. For example, if the training data  $T_d$  has a sample size five and is a collection of vectors  $X_n$ , and Classifications,  $Y_n$ ,  $T_d = \{X_1 Y_1, X_2 Y_2, X_3 Y_3, X_4 Y_4, X_5 Y_5\}$ , an individual tree may use a training set of  $T_d = \{X_1 Y_1, X_2 Y_2, X_2 Y_2, X_5 Y_5, X_5 Y_5\}$ , and a different tree may use  $T_d = \{X_1 Y_1, X_1 Y_1, X_3 Y_3, X_3 Y_3, X_5 Y_5\}$ . Feature Randomness is a process in which each decision only trains on a random subset of features rather than all available features. For example, if each  $X_n$  contains the same set of five features,  $F_m$ ,  $X_n = \{F_1, F_2, F_3, F_4, F_5\}$ , one tree may use only use  $X_n = \{F_1, F_2\}$  from within their vector in their training set, and a different tree may use  $X_n = \{F_1, F_4\}$ .

### 2.3. Behavioral Analytics

There is a growing, interdisciplinary body of research surrounding esports [3, 19], covering domains such as consumer research, sports and performance, consumer research, legal research, data visualization, and ethnographic research. One of the most active domains is esports analytics. Esports analytics, as a domain of inquiry, has evolved rapidly over the past decade. Prior to 2010, only a limited number of research publications existed, and the uptake of machine learning and AI in the esports industry was limited. Since then, the domain has expanded rapidly thanks to a combination of freely available data, commercial interest, and a broad scope of data science challenges across analytics, AI, and machine learning [11, 14, 15, 20-25].

Esports analytics operates with a variety of challenges, with the general purpose being to identify patterns in gameplay data, including spatio-temporal pattern analysis [23, 26, 27] and to inform players and teams, very early examples of this include [28, 29]. Machine learning and AI has been employed to generate match predictions [21, 22, 25], event predictions [3, 30] or identification of successful strategies and factors influencing gameplay [13, 31], or coaching systems [20, 32]. Most of the work done in esports analytics is focused on post-match analysis and predictions, with a smaller component focusing on real-time analytics [6, 14, 15]. Despite the growing wealth of esports analytics research, the domain is still in its infancy. Therefore, the challenge of identifying roles and playstyles and the evaluation of performance beyond basic metrics has been minimally addressed at the time of writing, with only a few publications on the topic [10, 12, 14, 33].

Yang et al. [12] modeled combat using graphs and metrics to predict success, integrating simple considerations regarding the role of specific heroes in Dota 2. Lee and Ramler [10] used metadata from the MOBA Heroes of the Storm to identify roles through cluster analysis. Demediuk et al. [14], a precursor work to the study presented here, used ensemble clustering to classify and label individual roles for heroes in Dota 2. Rather than focusing on pre-existing roles defined by expert knowledge, such as used by Yang et al. [12], the authors utilized unsupervised learning to identify roles which each hero can play. This enabled the separation of historical data for each hero.



### 3. Data Collection and Feature Engineering

Before the Performance Index system can be employed, we first must collect and label a dataset of professional Dota 2 matches, as well as define a list of features that will describe the key performance indicators of players within a match.

#### 3.1. Data Collection

As Dota 2 is a video game, every aspect of every match played, both public and professional, is recorded and saved by the developers. The general public may access this repository through a public api. OpenDota [34] provides easy access to match summaries and download locations of full match replays. Once we identify the correct games, we download the full match replays. Using a custom-built parser, we can convert those match replays to extract all the behavioral telemetry data.

Unlike traditional sports, the game of Dota 2 undergoes patches. These patches<sup>1</sup> change different aspects of gameplay, including hero performance, items, buildings, and monsters. Even slight changes to gameplay can have a considerable impact on gameplay and which heroes are selected. Therefore, it is vital to collect and use data from within a single patch to perform any analytics.

In our research, we are interested in professional-level game data. Professional games are considered games that are played at Minor<sup>2</sup> and Major<sup>3</sup> Dota 2 tournaments. We collect professional game data from the current patch and use it to build the performance index models and historical game performance database (see Section 4). As it takes time before there is enough data from a new patch to build effective models, the games from previous patches are used until enough games from the latest patch can be download. While we continually update the models, the results shown in this research are from professional games played during patch 7.26. We used 1000 professional matches from within this patch. At the time when analysis was conducted for the work presented here, patch 7.26 was the current patch.

#### 3.2. Playstyles

As mentioned in Section 2.1, there are five common positions that players perform in Dota 2. However, due to the diversity in the 119 heroes available to choose from at the time of writing, players utilize different playstyles in each position. While it is out of this paper's scope, we have created 10 different playstyles, which are used to refine the positions even further. These playstyles allow us to label each player's performance within the dataset. To achieve this, we employed Archetype Analysis (AA) [35] on the entire dataset, using 7 non-performance metrics, similar to work the role detection work presented in [14]. Using these playstyles, we label each

---

<sup>1</sup> Information about the changes each patch makes can be found at <https://www.dota2.com/patches/>

<sup>2</sup> [https://liquipedia.net/dota2/Dota\\_Minor\\_Championships](https://liquipedia.net/dota2/Dota_Minor_Championships)

<sup>3</sup> [https://liquipedia.net/dota2/Dota\\_Major\\_Championships](https://liquipedia.net/dota2/Dota_Major_Championships)



player's performance in every match in the dataset. The full list of playstyles identified by this work is shown in

Table 1.

It is important to note that this work can be performed without refining the Dota 2 positions into playstyles. However, each player's performance in the dataset would still need to be labeled based on their position. Employing playstyles improves the accuracy of both the Random Forest Classification and Historic Performance Dataset, described in Section 4.

Playstyle Label	Position	Description
Offlane Support	5	Spends most of their time in Offlane and roaming, gaining very little experience last-hits and denies.
Midlane Carry (last-hitter)	2	Spends nearly all their time in the Midlane, focusing on last hits, and very occasionally roaming to other lanes.
Safelane Carry (last-hitter)	1	Spends most of their time in the Safelane and in their jungle, focusing on primarily on last-hits
Midlane Carry (denier)	2	Spends all their time in the Midlane focusing heavily on denies and last-hits. Gaining large amounts of experience.
Offlane Carry (denier)	3	Spends all their time in the Offlane focusing heavily on denies and last-hits.
Safelane Carry (denier)	1	Spends most of their time in the Safelane and in their jungle, focusing heavily on denies and last-hits.
Safelane Support	5	Spends most of their time in Safelane and roaming, gaining very little experience last-hits and denies.
Roaming Carry	3	A rare playstyle, starting in Offlane, these players roaming constantly. Gaining large amounts of experience and last-hits from jungle and lane creeps.
Offlane Carry (last-hitter)	3	Spends most of their time in the Offlane, focusing on primarily on last-hits
Roaming Support	4	Spends their time roaming the map stacking camps in the jungle and helping different lanes when needed.

Table 1: The playstyles identified through Archetype Analysis. These playstyles were used in this research to label each entry in the dataset.

### 3.3. Feature Selection

As mentioned previously, all telemetry data from within a match of Dota 2 is recorded. We need to develop a concise feature set that would include all of the playstyles' key performance indicators. Through discussions with subject matter experts, we developed a feature set shown in Table 2. This feature set covers the most critical metrics for each of the playstyles. We will use Random Forest Classification in Section 4.1 to find how important (the weights) each feature is to each playstyle.



Data Feature Name	Feature Description
XP	The player's total amount of experience points, from enemy player kills, creeps, natural monsters, and objectives.
LEVEL	The current level of the player's hero. The maximum level is 30
MAX_HEALTH	The current total health pool of the player's hero. As hero items are challenging to model with Random Forest Classification, this metric is used as an indicator of the current player's items.
MAX_MANA	The current total mana pool of the player's hero. Similar to MAX_HEALTH this metric is used as an indicator of the current player's items.
LAST_HITS	The total number of creeps and natural monsters killed by the player.
DENIS	The total number of friend creeps killed by the player, denying the gold and experience from the enemy team.
NET_WORTH	Net Worth is the sum of the gold in the player's bank and the gold value of the player's items in their inventory and on the courier.
KILLS	The total number of killing blows on enemy heroes performed by the player.
DEATHS	The total number of times the player's hero has been killed.
ASSISTS	The total number of enemy hero kills the player has contributed to.
CAMPS_STACKED	The total number of natural enemy camps that have been stacked <sup>4</sup> .
SUPPORT_GOLD_SPENT	The total amount of gold spent to buy support specific items such as wards.
HERO_DAMAGE	The total amount of damage that the player has dealt.

Table 2: The list of features used for the Performance Index with associated feature descriptions.

## 4. Performance Index

The Performance Index is a cross-role comparator, meaning it can fairly compare to all players' performance regardless of their role. In this case study using Dota 2, we use the Performance Index to compare the performance of players across playstyles. To summarize, the Performance Index is calculated using a weighted sum (found using Random Forest Classification) of the current player's performance (compared to historical performances) across each of the features outlined in Section 3.3, all with respect to their current playstyle. We do this calculation at minute intervals after the first five minutes of the match. This results in a Performance Index value for each player on the team and can be used to compare their current performance in the match.

The Performance Index can be used in real-time on live games, providing broadcasters additional storytelling material and giving more significant performance insights for spectators. Conceptually

<sup>4</sup> For more detail on camp stacking please see [https://dota2.gamepedia.com/Creep\\_Stacking](https://dota2.gamepedia.com/Creep_Stacking)

we can break up the Performance Index development into two Phases, an Offline Phase and an Online Phase, with the full system shown in Figure 2.

The Offline Phase requires a collection of historic game replays (see Section 3.1). A full description of the Offline Phase is provided in Section 4.1. The Online Phase requires a live data stream from the game and can also be performed on non-live game replays. A full description of the Online Phase is provided in Section 4.2.

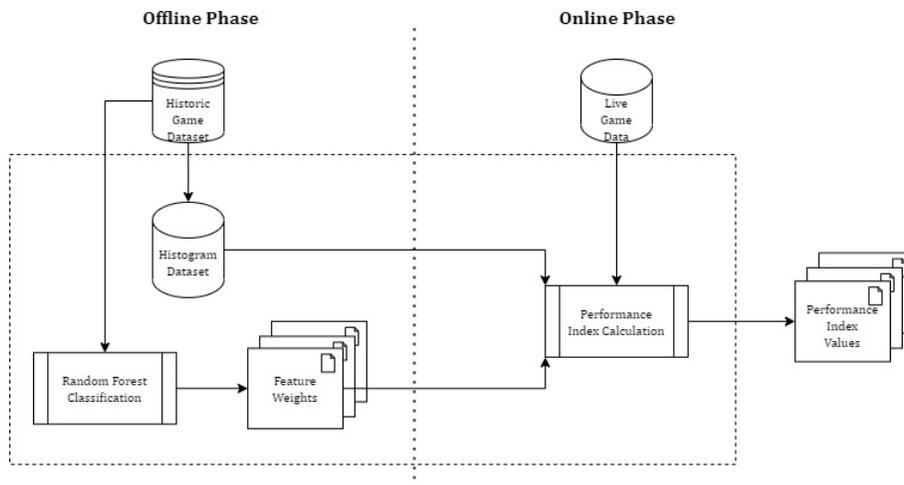


Figure 2: Visual description of the Performance Index system. The Performance Index is calculated through the combinations feature weights and percentile information from historic games. This allows for performance index values in live games to be calculated.

#### 4.1. Offline Phase

The Offline Phase only needs to be performed once, after enough data has been collected for the current patch. There are two critical components to the Offline Phase, building a Histogram Dataset and performing a Random Forest Classification.

The Histogram Dataset is built by first grouping each player's performance in every game in the Historic Game Dataset based on their playstyle,  $S$ . As the values for each feature change throughout it is further broken down into time steps,  $t$ , with each time step being 1 minute. A histogram,  $H_{fSt}$ , is constructed for each feature,  $f$ . These histograms are used as a lookup table for the Performance Index equation, providing a percentile,  $P_{fSt}$ , of the live player's current performance for a given feature and playstyle.

We also need to find the importance of each feature, which we will be calling the weights,  $W_{fSt}$ , for the different playstyles at each time step. These weights indicate how important each feature is for a playstyle in relation to winning the game. Specifically, each feature has a different level of importance for each playstyle. To find these weights, we use a Random Forest Classification algorithm outlined previously in Section 2.2.

For each of the 10 playstyles and each time step,  $t$ , between 5 and 30 minutes, we build a Random Forest Classification model for a total of 250 models. At each time step and for each playstyle each datapoint  $X_{iST}$  is vector of values for each feature from the historic dataset at a given time step, with the game outcome  $Y_i$  as the classification, i.e. 0 for a loss and 1 for a win. As there is not the same amount of playstyle performances in the dataset, so not all models are trained with the same number of datapoints.

To build each model, we split the dataset into an 80% train, 20% test set. We use 500 estimators (decision trees in the Random Forest) and train the model. We then test the model's accuracy against the test set and record the accuracy of the model. We repeat this 100 times and use the best model found as the final model for the time step and playstyle. We then repeat the entire process for each time step. For example, Figure 3 **Error! Reference source not found.** shows the best models' accuracy at each time step for the playstyle, Roaming Support. Early in the match, the models have a hard time predicting the winner, this is understandable as there is a lot of uncertainty. As time progresses, however, the models become more accurate. The average game of Dota 2 at the professional level is approximately 30 minutes, as there are limited games that progress after this time, we use 30 minutes as a cut-off point. This is because, with limited data, the accuracy of the Random Forest Classifiers falls off rapidly. The goal of this work is not to improve the accuracy of these models, instead, we use the accuracy as a measure of how confident we are in the level of importance of each of the features.

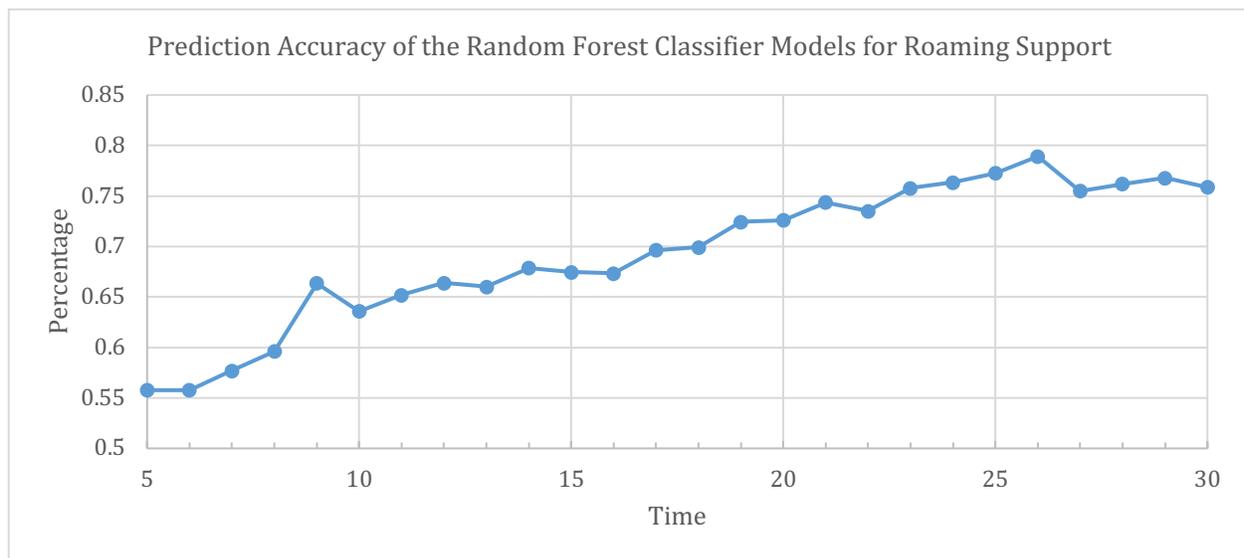


Figure 3: The prediction accuracy of the best Random Forest models for the playstyle, Roaming Support. Each point represents the prediction accuracy of the best Random Forest model found at a given time step.

Using these models, we can now find the importance (weights) of each feature for each playstyle utilizing a permutation importance method. To do this, we record a baseline accuracy by passing a validation set through the Random Forest model. We permute the column values of a single predictor feature, pass all samples back through the Random Forest model, and recompute the accuracy. The importance of the feature is recorded as the difference between the baseline and the drop in accuracy. While this is more computationally expensive than using the mean decrease in



impurity, the results produced are more reliable[36]. We now have a collection of weights of the features for each playstyle at each time step. Now that we know the weights of the features for each playstyle and have a histogram database, we can now calculate the Performance Index of players in a game.

## 4.2. Online Phase

The Performance Index's primary use case is to give a single metric that can compare all players in a live game, regardless of their playstyle. To do this, we utilize a live feed of the game, which will return each feature's current value for each player. We also use our playstyle Detection algorithm at the five-minute mark of the game to determine which playstyle each player is employing. While the primary use case is to utilize this in live matches, the Performance Index calculation can also be performed on non-live historical games.

For a given player in a live match, at each time step beginning at 5 minutes, we calculate the performance index as follows:

- We use the live feed of the game to return a set of values for each feature,  $\{f_1, f_2, \dots, f_{13}\}_{St}$  with the players corresponding playstyle  $S$ , at time step,  $t$ .
- Using the histograms,  $H_{fSt}$ , from Section 4.1 we lookup what percentile the player's performance is for each feature, returning  $\{P_1, P_2, \dots, P_{13}\}_{St}$ .
- We retrieve the set of weights,  $\{W_1, W_2, \dots, W_{13}\}_{St}$ , from Section 4.1 for the players corresponding playstyle  $S$ , at time step,  $t$ .
- Using the following equation  $PI_t = \sum_{f=1}^{13} (P_{fSt} * \frac{W_{fSt}}{\sum_{f=1}^{13} W_{fSt}})$  we then calculate the player's Performance Index for that time step,  $t$ .
- We then repeat this process for each player in the game. Passing the result to various tools to be utilized by broadcasters, and spectators, demonstrated in Section 5.
- For each minute, from 5 – 30 minutes repeating this process. For each minute after 30 minutes the  $H_{fS30}$  and  $W_{fS30}$  are used to calculate the Performance Index.

## 5. Application

The Performance Index considers the playstyles of each player in the game and provides a single metric that can compare their performances fairly. Providing broadcasters with a rich source of information in which to enhance their storytelling. It also reduces player performance's complexity, so spectators can quickly and easily understand how their team is performing.

As part of the Weavr Project<sup>5</sup> the Performance Index has been used at multiple Dota 2 tournaments across Europe. Currently, there are three methods for delivery of the Performance Index. The first method is through the use of a broadcast stream overlay tool. This tool is controlled by the broadcasters and can be toggled off and on to enhance their storytelling. Figure 4 shows an image of the Performance Index at the ESL Premiership Winter Skirmish 2019 tournament. Each player's current Performance Index is below their character frame at the top of the screen. As this is a broadcaster overlay, it is viewed by all spectators.

---

<sup>5</sup> Weavr homepage, <https://weavr.tv/>

The Performance Index has been incorporated into a Twitch<sup>6</sup> overlay tool. Twitch is a popular streaming service utilized by nearly all esports broadcasters as the primary location for spectators to watch games. The spectator controls the Twitch overlay tool, and they can view the Performance Index at any time they like. The Performance Index is displayed as a number as well as a color-coded radial bar. Below the Performance Index, is the top three most important features and their value for each player. Enabling spectators to better understand why a player has a particular Performance Index value. Figure 5 shows the open Twitch overlay tool at ESL Germany 2020 tournament.



Figure 4: The Performance Index broadcast overlay fielded at the ESL Premiership Winter Skirmish 2019 tournament. The Performance Index value for each player is displayed below their character frame. This overlay can be toggled on and off by the commentators.

<sup>6</sup> Twitch homepage, <https://www.twitch.tv/>



Figure 5: The Performance Index displayed by the Weavr Twitch overlay tool, at the ESL Germany 2020 tournament. The tool is accessed through a control panel on the edge of the screen and displays a window over the current game stream.



Figure 6: The Performance Index displayed in the Weavr Companion App at the ESL Thailand 2020 tournament.

Weavr has also developed a companion app<sup>7</sup> for smartphones aimed at spectators for use in stadiums and at home. The companion app provides numerous features for spectators, adding a wealth of insights. As part of the app, there is a feature wheel that shows the player's performance across all features outline in Section 3.3 as well as the Performance index. Spectators can click on a hero's portrait on the left and right side, which brings up a head-to-head comparison of the Performance Index. This allows them to compare any two players from opposing teams directly. If the spectator clicks on the Performance Index bar, it will open a window that displays the top 3 most important features for that player's playstyle and their current percentile performance in that feature. Figure 6 shows a match in the Weavr Companion App at the ESL Thailand 2020 tournament, the image on the left is the feature wheel, with the image on the right displayed once a spectator clicks on the Performance Index bar. During ESL Birmingham 2020, ESL Germany 2020 and ESL Thailand 2020 tournaments, more than 200,000 people directly interacted with the Performance Index in the Twitch version and the mobile version.

<sup>7</sup> Weavr Companion Application, <https://weavr.tv/dota2/>



## 6. Conclusion

The paper presents a novel metric called the Performance Index, which allows all players in the game to be compared irrespective of their role. Using a labeled historical dataset of player performances and combining that with each feature's importance, based on their playstyle, the Performance Index can be calculated for every player in a live match. The Performance Index also extends to other contexts such as professional teams for, team and player analysis, and swings in performance and recruitment. While the case study used in this paper uses Dota 2, the principle behind the Performance Index is not limited to esports but can be applied similarly to other team-based, multi-role sports such as basketball, baseball, and football. The playstyles utilized in this work can be replaced by position or role on the team. The only limitation to the Performance Index's use in live matches is access to a live feed and delivery methods.

## 7. Acknowledgements

This work has been created as part of the Weavr project (weavr.tv) and was funded within the Audience of the Future program by UK Research and Innovation through the Industrial Strategy Challenge Fund (grant no.104775) and supported by the Digital Creativity Labs (digitalcreativity.ac.uk), a jointly funded project by EPSRC/AHRC/Innovate UK under grant no. EP/M023265/1.

## References

1. Newzoo. *Newzoo Global Esports Market Report 2020*, <https://newzoo.com/insights/trend-reports/newzoo-global-esports-market-report-2020-light-version/>.
2. Drachen, A., M.S. El-Nasr, and A. Canossa, *Game Analytics: Maximizing the Value of Player Data*. 2013: Springer.
3. Schubert, M., A. Drachen, and T. Mahlmann. *Esports analytics through encounter detection*. in *Proceedings of the MIT Sloan Sports Analytics Conference*. 2016. MIT Sloan Boston, MA.
4. Alamar, B. and V. Mehrotra, *Beyond Moneyball: The Future of sports analytics*. *Analytics Magazine*. 2012.
5. Kriglstein, S., et al. *Be Part Of It: Spectator Experience in Gaming and Esports*. in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020.
6. Block, F., et al. *Narrative bytes: Data-driven content production in esports*. in *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video*. 2018.
7. Valve Corporation. *Dota 2*. <https://www.dota2.com/play>.
8. Riot Games. *League of Legends*. <https://na.leagueoflegends.com/en-us/>.
9. Hi-Rez Studios. *Smite*, <https://www.smitegame.com/>.

10. Lee, C.-S. and I. Ramler, *A Data Science Approach to Exploring Hero Roles in Multiplayer Online Battle Arena Games*. Data Analytics Applications in Gaming and Entertainment, 2019: p. 49.
11. Semenov, A., et al. *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*. 2016. Springer.
12. Yang, P., B.E. Harrison, and D.L. Roberts. *Identifying patterns in combat that are predictive of success in MOBA games*. in *FDG*. 2014.
13. Eggert, C., et al. *Classification of player roles in the team-based multi-player game dota 2*. in *International Conference on Entertainment Computing*. 2015.
14. Demediuk, S., et al. *Role Identification for Accurate Analysis in Dota 2*. in *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*. 2019.
15. Kokkinakis, A., et al. *DAX: Data-Driven Audience Experiences in Esports*. in *Proceedings of ACM International Conference on Interactive Media Experience (IMX) 2020*. 2020. Association for Computing Machinery (ACM).
16. Breiman, L., *Random forests*. Machine learning, 2001. **45**(1): p. 5-32.
17. Cutler, A., D.R. Cutler, and J.R. Stevens, *Random forests*, in *Ensemble machine learning*. 2012, Springer. p. 157-175.
18. Breiman, L., et al., *Classification and regression trees*. 1984: CRC press.
19. Reitman, J.G., et al., *Esports research: A literature review*. Games and Culture, 2020. **15**(1): p. 32-50.
20. Agarwala, A. and M. Pearce, *Learning Dota 2 team compositions*. Sl: sn, 2014.
21. Johansson, F. and J. Wikström, *Result prediction by mining replays in dota 2*. 2015.
22. Makarov, I., et al. *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*. 2017. Springer.
23. Rioult, F., et al., *Mining tracks of competitive video games*. AASRI procedia, 2014. **8**: p. 82-87.
24. Demediuk, S., et al. *Player Retention In League of Legends: A Study Using Survival Analysis*. in *Proc. of ACM ACSW Interactive Entertainment*. 2018.
25. Hodge, V.J., et al., *Win Prediction in Multi-Player Esports: Live Professional Match Prediction*. IEEE Transactions on Games, 2019.
26. Drachen, A. and M. Schubert, *Spatial game analytics*, in *Game Analytics*. 2013, Springer. p. 365-402.
27. Drachen, A., et al. *Skill-based differences in spatio-temporal team behaviour in defence of the ancients 2 (dota 2)*. in *2014 IEEE Games Media Entertainment*. 2014.
28. Hoobler, N., Humphreys G. and M. Agrawala. *Visualizing competitive behaviors in multi-user virtual environments*. in *Proceedings of IEEE Visualization*. 2004.
29. Miller, J.L. and J. Crowcroft, *Group movement in world of warcraft battlegrounds*. International Journal of Advanced Media and Communication, 2010. **4**(4): p. 387-404.

30. Katona, A., et al. *Time to Die: Death Prediction in Dota 2 using Deep Learning*. in 2019 *IEEE Conference on Games (CoG)*. 2019.
31. Pobiedina, N., et al. *Ranking factors of team success*. in *Proceedings of the 22nd International Conference on World Wide Web*. 2013.
32. Gupta, V., et al., *A Team Based Player Versus Player Recommender Systems Framework For Player Improvement*. 2019.
33. Gao, L., et al., *Classifying dota 2 hero characters based on play style and performance*. Univ. of Utah Course on ML, 2013.
34. Cui, A., H. Chung, and N. Hanson-Holtry. *OpenDota*, <https://www.opendota.com/>.
35. Cutler, A. and L. Breiman, *Archetypal analysis*. *Technometrics*, 1994. **36**(4): p. 338-347.
36. Strobl, C., et al., *Conditional variable importance for random forests*. *BMC bioinformatics*, 2008. **9**(1): p. 307.