Proceedings Paper:

Pedrassoli Chitayat, Alan, Block, Florian Oliver orcid.org/0000-0003-0348-6731, Walker, James Alfred orcid.org/0000-0003-2174-7173 et al. (1 more author) (Accepted: 2024) How Could They Win? An Exploration of Win Condition for Esports Narratives in Dota 2. In: Annual Symposium on Computer-Human Interaction in Play (CHI PLAY 2024). ACM (In Press)

https://doi.org/10.1145/3677079

How Could They Win? An Exploration of Win Condition for Esports Narratives in Dota 2

ALAN PEDRASSOLI CHITAYAT, University of York, UK
FLORIAN BLOCK, University of York, UK
JAMES A. WALKER, University of York, UK
ANDERS DRACHEN, University of York, UK

Data analytics is commonly used to enable storytelling and enhance esport coverage. One prominent use of it is win prediction, where machine learning models predict the winner of the game before its conclusion. However, predictions are most commonly results of black-box systems, forcing commentators to produce ad-hoc interpretations. Additionally, broadcasters generally rely other metrics to build narratives, limiting the impact of win prediction models for storytelling. This paper explores an alternative method to win prediction, identifying the needs of broadcasters to guide development of a novel win condition model. By focusing on existing storytelling points, the proposed win condition model can offer greater storytelling opportunities to broadcasters, focusing on the user needs identified from within the esport domain. Rather than utilising game state data to predict the winner, as it is usually done in win prediction, the proposed win condition model uses an exploration of the possible winners to predict the game state needed for each team to win. Lastly, the features identified for win condition are evaluated through a series of machine learning models, which provide a data-driven metric to test and predict win condition in the context of Dota 2, a popular esport title.

CCS Concepts: • **Human-centered computing → HCI theory, concepts and models**; • **Computing methodologies** → *Machine learning*; • **Applied computing** → *Media arts.*

Additional Key Words and Phrases: Esports, User-centred Design, Outcome Prediction, Win Condition, Machine Learning

ACM Reference Format:
Alan Pedrassoli Chitayat, Florian Block, James A. Walker, and Anders Drachen. 2024. How Could They Win? An Exploration of Win Condition for Esports Narratives in Dota 2. *Proc. ACM Hum.-Comput. Interact.* 8, CHI PLAY, Article 314 (October 2024), 22 pages. https://doi.org/10.1145/3677079

1 Introduction

Esport is a form of digital entertainment, in which a video game title is played competitively, usually within a tournament or league setting [12]. The highest level of esports is typically played by highly proficient individuals, and those games can be broadcast to large audiences across the globe, with a live in-person audience sometimes also being present [4]. The large popularity of these titles has led to the fast growth of the esport market as a form of media consumption [6, 22], which has subsequently motivated the development of a new field of academic research - esport analytics - in which Machine Learning (ML) and other Artificial Intelligence (AI) techniques are prominent topics

Authors' Contact Information: Alan Pedrassoli Chitayat, University of York, York, UK, alan.pchitayat@york.ac.uk; Florian Block, University of York, York, UK, florian.block@york.ac.uk; James A. Walker, University of York, York, UK, james.walker@york.ac.uk; Anders Drachen, University of York, York, UK, anders.drachen@york.ac.uk.

ACM 2573-0142/2024/10-ART314
https://doi.org/10.1145/3677079

of research, particularly as esport titles tend to be complex environments, with vasts accessible datasets which creates an ideal study domain for such topics [24].

One of the most common example of ML being used within esport' analytics is win prediction [9], where models have been trained to attempt to predict the winner of a match before its conclusion [2, 13]. In addition to its popularity in academia, win prediction models have been widely implemented in esport titles directly as a form of narrative enhancement for spectators of the game. Figure 1 depicts an example of a win prediction graph that is available for spectators of *Dota 2*[1], a popular esport title.

Fig. 1. An example of the Dota 2's Win Prediction graph as available within the in-game client

As depicted in Figure 1, the win prediction available in the game is presented as a continuous prediction that updates over time, reflecting events that happen in a match. Similar continuous predictions can also be found within academia, in which details of the models used are made available within the literature [13, 28]. Models used for predictions are commonly black-box systems - i.e. it is not usually feasible to ascertain what caused a particular prediction to be generated. As a result, broadcasters and audiences are generally required to use ad-hoc interpretations of the results to build narratives. However, predictions may not always match user expectation, which can lead to a jarring process and generally limit the utilisation of such tools in an ecological context [37].

Furthermore, while more interpretable models exist [37], the competitive and complex nature of the titles typically lead to sharp shifts in output predictions as depicted in Figure 1, providing fewer insights into potential future states, instead providing a description of the current and past states of the match. This can consequently lead to either less trust in the prediction system [33] or, in the case of more knowledgeable viewers, lead to limited insights that were not already known by the users directly [33].

For this reason, this paper investigates a different approach at prediction. By focusing on the narrative and storytelling applications within *Dota 2*, this paper proposes a win condition model. This is achieved by first investigating what broadcasters and esport analysts use as narrative

[1] https://www.dota2.com/home

points to communicate their expectations of the game (particularly at early stages or pre-game phases). By analysing existing broadcast content, features that are perceived to have predictive significance can be identified. As extracted from existing broadcast content, these features are already commonly used in the ecosystem, and thus have significant impact into driving narratives. Their relationships are then explored, in which a model for representing win condition in terms of what is used in narratives is proposed. Then, the proposed model is evaluated through a controlled comparison between machine learning models, in which the causal relationship between variables is investigated and the way in which they can be predicted and utilised is outlined. Lastly, this paper explores some of the ways in which win condition can be incorporated into the broadcast, and how it can be used to enhance existing narratives.

Thus the contribution of this paper is twofold. Firstly, the features which are commonly used in the esport broadcast ecosystem are identified. This is compared to the existing literature as well as empirically evaluated through the use of ML models. Secondly, a narrative focused win condition system is proposed. This system is designed to enhance existing narrative in such a way that it can be more readily applied to the esport broadcast ecosystem when compared to existing win prediction tools and knowledge.

2 Background

This paper aims at investigating aspects of esports that are widely understood by analysts and experts as factors that determine win condition within the domain. To achieve this, *Dota 2* - a popular esport title - is used as a case-study. This game was chosen as a research domain as it contains a wealth of academic research, in the fields such as win prediction [13, 32], feature explorations [10, 32], data visualisation [7] and data-driven storytelling [18]. Additionally, due to the title's popularity and vastly accessible APIs [24, 26, 31], it provides a data-rich environment which allows for ease of data gathering for the training and evaluation of ML models.

Dota 2 is a Multiplayer Online Battle Arena (MOBA) game, in which two teams compete to attack and destroy the enemy base, while defending their own. Teams are composed of 5 players, each playing a unique character (called hero) from a pool of over 100 characters. As players accumulate resources, they are able to customise their characters with abilities and in-game items. This means that players are able to make several decisions on every match that alters the way in which a player can play, adding complexity to the game. This can impact the way in which models can make predictions, as well as the way in which audiences understand the game, as several multifaceted interactions can have significant impact in the game [18, 24].

Therefore, broadcasters of esport content often are required to formulate narratives that, not only entertain audiences, but also inform them on the state of the game [3]. However, the large complexity of the games can make it difficult for broadcasters to fully identify and communicate their insights of the match. Thus, this paper aims at providing a narrative focused, concept to aid in storytelling and enhance the audience experience through data-driven space exploration.

3 Related Work

Since the emergence of esport as a field of research, clear evidence of the benefits of storytelling could be found. Wang and Fan (2022) [34] has outlined that the majority of esport audiences are driven by the excitement of in-game events and the commentary associated with them. To improve in the quality of commentary, the importance of data analysis [3] and data visualisation [7, 36] within in-game metrics for storytelling has been outlined. By investigating data-driven narrative within a live *Dota 2* tournament, Block et al. (2018) [3] designed a framework for interacting with audiences through a highly produced interface of communication. Building from that, Kokkinakis et al. (2020) [18] has later coined the term Data-driven Audience Experience (DAX), which further highlights the importance of using data to derive narratives. As authors highlighted, the complexity typically associated with competitive esport titles can make understanding what is happening in a game a challenging process even for experts. For this reason, it is important to create tools that assist in the understanding of the game, which both help reduce the barrier of entry for new audiences, and improve the overall experience for existing audience. This has been similarly demonstrated by Carlsson and Pelling (2015) [5], who have outlined a framework for engaging with audiences, presenting information interactively to allow them to explore the space and reach their own narrative points. Similarly, Pedrassoli Chitayat et al. (2024) [25] have also demonstrated an active demand for spectators to consume more in-depth content for enhancing storytelling and narratives, outlining population-wide patterns of viewership interactions.

Several authors have produced work aimed at reducing the complexity of the game through investigating different features. Demediuk et al. (2021) [10] has designed a methodology to evaluate the performance of professional players within a match. This is done by first identifying their role within the team [11], and then comparing their current performance in Key Performance Indicators (KPIs) to data-generated targets. Through this methodology, the author is able to not only inform audiences of how well a player is performing, but also contextualise this insight, by highlighting the relevant KPIs used to determine this measurement. This example highlights how contextual information (i.e. the important KPIs for a given role) can help give audiences more information about the output of models, aiding at reducing the complexity and enabling DAX.

Another prominent example is in the win prediction (or outcome prediction) domain which has seen a wealth of knowledge [13, 32]. In this case, models generally produce an win prediction as a function of game state (i.e. the current knowledge of the state of a particular game is used to predict the outcome). The state of the game is typically represented by a combination of a range of variables or features. This is then used to produce a confidence number that is generally interpreted as the win chance of one of the team, as also depicted in Figure 1. However, details of how this prediction is achieved is typically not available, as they are results of complex machine learning systems with varying weights and biases.

More interpretable forms of win predictions, such as what has been proposed by Yang et al. (2022) [37], can provide audiences and broadcasters with more explanations to why a particular team is favoured to win. Similarly, Makarov et al. (2018) [20] has outlined a prediction system which uses KPIs with the knowledge of professional player roles to produce predictions. The KPIs used to produce the prediction can be highlighted aiding in interpretability of the prediction. However, the ever evolving nature of esport matches can often lead to sharp shifts of powered in the game, which radically changes the advantages of team within the game, and consequently impacting the odds of winning the match. This has a consequence to the way in which win prediction is presented, leading to sharp dips and valleys in graphs as game events have their abrupt impact into the game and the win prediction resulting from ML models adapts to reflect the new data. This phenomenon is observed in the example win prediction graph in Figure 1 and in similar items from the literature [13, 37], where the graph can change from heavily favouring one team to immediately and abruptly favouring another. Thus, the resulting win prediction graphs are generally understood by audiences and broadcasters as a representation of any given moment in a match, rather than a true prediction of the final expected outcome [33]. Additionally, esport audiences are commonly also players, and highly experienced spectators may have different needs, often being able to interpret the game situation without the aid of a win prediction graph [15, 33]. This effect may also impact spectator trust in such systems, especially when the results of the prediction do not align with their expectations. Furthermore, as outlined by Hodge et al. (2021) [13], a win prediction model that is too accurate would have a negative impact into the audience experience, as it may remove the uncertainty and thus the emotional arousal of consuming esport content. The author also suggests that a win prediction model that is too inaccurate would, however, not be considered a trustworthy tool by audiences and would generally be disregarded or ignored. This dilemma has lead to a call for more research into the application of win prediction models [33] and the way in which such insights are communicated to audiences, particularly as it is known that the user experience is directly related to the in-game excitement and present in the livestream [21, 34].

In the other hand, other predictive models have also been proposed beyond that of win prediction, such as team-fight predictions [31], encounter detection [27] and character death prediction [16, 26]. These examples outline the importance of data and how it can be used to accurately predict, not only the outcome, but several aspects and features of the game. In particular, different architectures of Neural Networks have been shown to be an effective way to predict features and events within a the game [8, 16, 26, 31]

In addition to interpretable win prediction models [37], other authors have outlined the importance of interpretability. Ahmad et al. (2019) [1] has outlined different ways of modeling player behaviour with a focus on understandability and interpretability. This work relies on human-labels to ensure the models can be interpreted when visualised. While the authors mention that the human-in-the-loop approach is time-consuming, it brings light to the importance of the way in which models and visualisations are applied in the ecosystem.

Additionally, the way in which esport games are design typically evolves [24, 30], which leads to a change in how players interact with the environment [17]. Thus, the esport playing community typically develops metas, which are community driven conventions to how to play in a particular iterations of the game (i.e. game patches). This outlines how changes to the environment are a key and constant factor within the domain of esports and thus models must be able to adapt to keep up to date with the newest game patch in the same way that players adapt to the changes in the environment [24, 30, 32].

Beyond the context of esports, causal effects relationships have been proven to be an effective way of explaining predictions to users [23], in particular within the outcome exploration domain [14]. In the other hand, games (such as esport titles) are designed to have multiple ways to achieve a goal, thus the condition necessary to achieve a goal may vary [29]. While several features are generally understood as relevant for win prediction [13, 27, 32] and thus are expected to inform win condition, no investigation of the use-case as understood and utilised by specialist or audiences more broadly was identified.

In the context of games, more generally, win condition is a simple concept well understood by game designers [29]. In its simplest form, win condition refers to the main game objective, such as capturing the enemy king in the game of chess. However, just as the game of chess may have several states in which only one outcome can be achieved [35], win condition can be achieved before the main goal of the game is reached. Therefore the win condition system proposed by this paper explores the ways of predicting the state needed for a team to achieve victory, and how this can be used alongside existing narratives that do similar win condition stories within the existing *Dota 2* broadcast ecosystem.

To summarize, a clear need for storytelling and narrative focus tools for esport content can be understood [3, 5, 7, 18, 25]. Esport audiences are driven by the excitement of in-game events and the commentary associated with them [34], and data-driven tools and narratives have been shown to be a reliable way to enhance this excitement [3, 18]. Win prediction has been a well established practice within the domain [32, 37] however questions towards the integration and application of such models are yet to be fully explored [13, 33]. Features explorations have identified several aspects and events within the game that are understood to be imperative towards determining the winner of a match [16, 26, 27, 30, 32]. Win condition is an intuitive and well understood aspect of game design [29], which outlines the state needed to achieve victory in a particular title. While mechanical definitions of win condition simply refers to the main game objective, certain game states can only lead to the same conclusion, i.e. identifying when a game has been won by a team without the game ending is another form of win condition. While game events and the volatile nature of esports can lead to several sharp changes in win prediction, as models update to reflect the new game states, a win condition prediction would be more consistent. Additionally, a model that is specifically designed to enhance existing narratives could aid the way in which data is used to cover the broadcast of esport matches, and improve the DAX. Thus, this paper proposes a new way to use data that aims at identifying and predicting the win condition of esport games. The concept of win condition proposed in this paper is primarily designed through investigating existing narratives and then evaluated through ML models, which have been known and reliable forms of modeling and predicting events and outcomes in the field of esports [10, 16, 26, 31, 32].

4 Defining Win Condition

This paper investigates the ways in which *win condition* can be defined, including the features associated with win condition and their relationships. This section provides a general definition of the term, and how it contrasts to a *win prediction*.

As explored in Section 3, win prediction is a well studied subject within esports, with several machine learning models being proposed in the existing literature [13, 32, 37]. Those models are intended to provide a prediction of who will win a match, given the current state of the game. As explored in this paper, this has several implications to storytelling, including potential detrimental effects including in audience trust and suspense building [13, 33].

By contrast, this paper proposes *Win Condition*, which is designed to more seamlessly integrate with existing esport narratives. This is achieved by first reversing the flow of data from typical win prediction models. Rather than predicting who will win the game, win condition produces a prediction of *how* a team may win the game. Therefore, while the output of a typical win prediction system is a percentage chance for a team to win a game, the output of a win condition system is the game state that is estimated to be needed in order for a team to win a game.

Additionally, win prediction models typically utilise a wide range of features to produce predictions (i.e. to represent the game state as an input). These features are selected with the intent to maximise model performance and accuracy, and may not always correlate to user needs (see Section 3). By contrast, win condition is designed for storytelling, with the intent of being readily integrated into existing narrative. For this reason, the features used as output of the win condition system (i.e. to represent the game state needed to win) were selected from existing narrative, ensuring the system matches the broadcaster's user needs (refer to Section 6.1).

By using an exploration system, similar to other works proposed in the literature [37], this paper provides an interpretable output for win condition. Furthermore, ML has been shown to be a reliable way to investigate and analyses different features of esports [32], while data-driven stories have been known to enhance the general audience experience [3, 18]. Additionally, complex causal relationships between variables can be modeled through structured causal models (SCM) [23], which allow for a greater understanding and representation of complex underlying relationships and correlations. For this reason, the Win Condition system proposed by this paper first investigates common narratives (see Section 6.1), then models these features in a SCM (see Section 6.2) and then evaluates it through the use of machine learning (see Section 7).

5 Methodology

In order to create a system that is ecologically consistent with the domain, it is first important to understand the main topics of existing narrative, when related to broadcasters making predictions of outcomes. Therefore, a study of existing esport content is performed, in which the coverage of six distinct matches are analysed. The study focused on the draft-phase section of the matches, which is the phase in which characters are selected by both teams. During this phase, the main match has not yet began, therefore the commentary provided is highly speculative when compared to the descriptive commentary that is more predominant during the game-phase. Therefore, broadcasters tend to focus their narrative on their own predictions and understanding of how the game may be played by both teams and what that could mean in relation to the outcome. This can provide an insight into the features and aspects of the game that broadcasters consider important towards winning, and more crucially, what they believe a team must achieve in order to emerge victorious.

Three games from two tournament (for a total of six matches) were selected. Firstly, the lower bracket finals from the 2023 ESL One Berlin[2] was selected. The ESL One Berlin is a major tournament of *Dota 2*, with large coverage both in person and through live broadcast. For brevity, the three 2023 ESL Berlin Lower Bracket Finals games will be referred to as ESL Game 1 through 3 respectively. Secondly, to ensure a diversity of content is explored, the use of three games from a amateur tournament were selected. Despite being played by amateur players, the games were commentated by two professional broadcasters (Ted "Pyrion-Flax" Forsyth[3] and Jake "SirActionSlacks" Kanner[4]) who have also produced content in the 2023 ESL One Berlin and other major *Dota 2* tournaments. The games used were part of 'Pyrion-Flax" weekly in-house tournament, in which members of his community take part in teams in a league format. The three games utilised took place in the 22nd August 2023. For brevity, the three games from this tournament will be referred to as PF Games 1 through 3 respectively.

Audio and video footage for the draft phase of all games were extracted from the Twitch platform, and then transcribed using Descript[5]. Nvivo 14 [6] was used to perform a Content Analysis (CA) [19] in the transcribed data, in order to identify recurring topics in broadcast in relation to win prediction and the conditions for winning.

Once an understanding of the common narrative patterns can be established, it is important to model and formulate the causal relationships between the topics being depicted and the outcome of the game. Structural Causal Model (SCM) have been shown to be an effective tool to determine and express causal relationships [23]. In turn, an exploration of causal relationships can also aid in explainable relationships, particularly when exploring win predictions [14]. Thus this paper utilises SCM to identify the causal relationship between what is perceived by broadcasters to be important features in order to advice on the feature engineering for training and evaluating ML models to storytelling. Such ML models are then utilised in this study to formulate win condition narratives as a data-driven tool for enhanced esport coverage.

6 Features of Win Condition

To identify the features that constitute win condition, this study first performs a content analysis on data collected from real-world tournaments and environments. Then the resulting data is analysed and a SCM is proposed to identify the relationships between variables. Lastly, ML models are trained using the proposed relationships, serving as a data-driven method to evaluate the relationship of perceived features and allow for a win condition system to be established.

[2]https://liquipedia.net/dota2/ESL_One/Berlin_Major/2023
[3]https://liquipedia.net/dota2/PyrionFlax
[4]https://liquipedia.net/dota2/SirActionSlacks
[5]https://www.descript.com/
[6]https://lumivero.com/product/nvivo-14/

6.1 Content Analysis

Through the analysis of the draft phase of the games, it is clear that the narrative presented by broadcasters contains a series of speculations on what will happen in the game and why. Certain codes can be extracted when the narrative of the broadcast is analysed. Commentary tends to focus on similar aspects of the game to guide their predictions and explanations. The general codes have been outlined in Table 1.

Table 1. CA codes and occurrences

Main Code	Sub Code	Professional	Amateur	Total
Performance	Combined	59	21	80
Performance	Kills	31	12	43
Performance	Timings	20	4	24
Performance	Comfort picks	8	5	13
Meta	Combined	37	25	61
Meta	In meta	29	15	44
Meta	Not in meta	8	10	17

6.1.1 Kills.

The code for Kills refers to the instances where broadcasters focus on the predicted score or the ability to secure or deny kills from the enemy team.

For example, the following extracted quote was observed in ESL Game 1:

> My big question is are you gonna get through the wall of tide [referring to Tidehunder, a hero that has been selected and relating it back to a previous comment where they described it as a "wall" because of its survivability] like this hero is really gonna determine the game

In this case we observe the narrative driving the attention to the high survivability of one of the heroes. The narrative implies how the ability of the hero to remain alive is a key aspect of the character, while putting pressure in the enemy team to kill them. In other words, the casters are outlining how a successful outcome for either team will be connected to whether or not the hero is killed during the game.

Another example extracted from the broadcast of PF Game 2:

> How does he get anywhere near an enemy? My boy [referring to one of the players] gets chained up [referring to abilities which stops characters from being able to move temporarily] at any time, he's dead, it's over, there's nothing he can do.

In this example the narrative expresses how a character is particularly vulnerable to being killed if their movement is disabled. This suggests that the commentators identified this character survivability as a relevant factor for a teams success, forming a narrative where they expect the character to be killed repeatedly, which reduces the perceived chance for that team to win the match.

6.1.2 Timings.

This code contains content which refers to match duration and other timing related subjects. In other words, it includes references to time-dependent features, which can generally be encapsulated by the time of the match (i.e. duration).

One example extracted from PF Game 1:

> Monkey King [character name], uh, hard counter to Axe [character name] in the laning phase [refers to the early stages of the game], but towards the end of the game can... Flip on its head.

In this example, the casters are outlining how the team that has picked the character "Axe" has an advantage in the earlier stages of the game, as Axe performs particularly well against "Monkey King", being able to restrain the hero's capabilities at the early stages. However, as the game progresses the advantage shifts and "Monkey King" is no longer at a disadvantage. This puts great perceived importance on the timing of the game, which can determine the outcome.

Another example of casters emphasising time as a key factor in determining the outcome can be observed in this extract from ESL Game 3:

> it's not like they [team 1] can't win late, but I definitely do favor them [team 2] in the later portion of the mid game more so

In this example, the casters are outlining how the team composition from team 2 is stronger the longer the game takes to finish. This means that team 1 would be expected to win the game faster, while team 2 would be perceived as having an advantage in later stages. Thus, if the match has a short duration, team 1 would be expected to win, while otherwise a longer match duration would benefit team 2. Another example extracted from the same game can be seen below, where the casters continue with the early and late advantage narrative:

> Liquid [team 1] definitely, effectively saying, guys, we're gonna be here in 25 minutes, and 9Pandas [team 2] need to weather the storm and buy enough time

6.1.3 Comfort Pick.

Comfort picks refer to instances where the narrative draws attention to a player's experience with a particular character. In this case, casters outlined how playing characters in which players have demonstrated particular abilities can be strongly beneficial to their perceived chance of wining.

The following example was extracted from ESL Game 1:

> They're playing to their teammates strengths. Roger [player name] on this Chen [character name] or Enchantress [character name] has looked... he's the best Chen or Enchantress at the tournament.

In this quote it can be observed how the broadcaster was outlining how Roger had previously demonstrated high degrees of proficiency while playing both Chen and Enchantress (two characters can be played similarly and are able to fulfil similar roles in a team). Additionally, the caster has outlined them as "the best" at both characters in the entire tournament, which demonstrates playing this character would increase their perceived expectation of winning the game. A similar pattern can be observed in the following extract (PF Game 2), where a hero ban is explained to be due to a comfort pick denial:

> Science Whiz Ben [player name], uh, an OD [abbreviation of character name] specialist, which is why OD's been banned,

6.1.4 In Meta.

This extracts refer to casters outlining how changes in game design have disproportionately improved certain characters or items. This can be observed in the extract from PF Game 3:

> Witch Doctor [character name] is probably, uh, in the best support at the moment

Another example extracted from ESL Game 2:

> Monkey King. Uh, this is a hero that has been a bit overlooked, I think, in this meta. I think he's looking insanely strong,

In both examples we can observe how casters perceive certain heroes to be good in given game iteration. It can be noted that the narrative outlined by them puts an expectation of victory if teams play those characters.

6.1.5 *Not In Meta.*

While certain changes to the game may benefit some characters, others may be disproportionately weakened. This code contains narrative in which casters outline aspects of the game they perceive to have a negative impact in the expectation of winning the match due to patch changes. This can be seen in the example from PF Game 1:

> whatever lane Marci [character name] is in is lost. Now that hero is an absolute dog [bad].

Another example of such narratives was extracted from Game 3 of the same tournament:

> Sand King [character name], a hero that used to be good last week, um, now, no longer

In this particular example, the caster used the fact that this particular character had been changed recently to build their narrative.

Similar narratives can also been observed in professional games coverage, such as the two extracts from ESL Game 2:

> Wisp [character name] has fallen out of priority a little bit for teams at this tournament.
> ...
> been a pretty loosey hero overall, Nyx [character name].

6.1.6 *Content Analysis Synthesis.*

Overall, the content analysis shows that broadcasters rely on four main features to make predictions.

- Player past experience with a character
- Team composition and how that reflects with the current patch
- The kills potential of a team
- The ideal timings for a team

As outlined in the literature, the features utilised by broadcasters have also been proven to be reliable factors for predicting the outcome, in particular player past performances, match duration and kills differentials [32]. Furthermore, game updates are designed to alter the balance of the game. This can significantly change the viability of a character [17] and the way in which they are played [24]. Therefore, comparing the team composition not only to each other, but also to the current patch can lead to a better understanding of the context in the game [30]. Thus, the narrative points raised by casters generally align with features used to produce, or known to impact, predictions in the ML literature.

6.2 Structural Causal Model

Generally, win prediction models use some form of data to represent the game, (or game state) to predict the winner. This can be represented in a SCM at its most abstract form depicted in Equation 1. However, the proposed model in this study attempts to predict the game state needed to win the game. Thus, a win condition model can be broadly represented in the SCM depicted in Equation 2.

$$Winner := f_{Winner}(State) \tag{1}$$

$$State := f_{State}(Winner) \tag{2}$$

Content analysis of the ecosystem outlined the key factors used by broadcasters to ascertain the state. Therefore the variable "State" can be broken down into player historic performance &

preferences (*Player*), team composition (*Lineup*), *Kills* and match duration (*Duration*). However, those variables have complex causal relationships between them, so in order to produce a model, it is important to identify and attempt to model this structure.

Using these features, it is then possible to break down the win condition causal relationship as follows:

$$
\begin{aligned}
Duration &:= f_{Duration}(Winner, Lineup, Kills) \\
Kills &:= f_{Kills}(Winner, Lineup, Duration) \\
Lineup &:= f_{Lineup}(Player) \\
Player &\in U
\end{aligned}
\tag{3}
$$

Firstly, Kills and Duration are the only variables that change once a game starts. Therefore, for the purpose of describing a condition to be reached in game, those are the main variables that will be used to describe the state goal. However, both Kills and Duration have a direct causal relationship with Player and Lineup, as the specific values for these features may have an impact in the number of kills present in the game and the overall match duration. Furthermore Lineup is assumed to be a function of Player, as players pick their characters and they are presumed to know their experiences and preferences. Lastly, although a relevant feature, Player is identified as an unobserved variable (U). This is a limitation of this study as measuring each individual players is beyond the scope of this paper. Certain aspects of the literature have represented this feature by identifying their past performances with the character [32], however this approach is limited to specific players and the general approach proposed by this study is not player specific.

As outlined by the full breakdown of the space, modeling it into a single SCM graph is infeasible, due to the complex and cyclic relationship observed between Kills and Duration. Thus, a simplification of the space can be proposed, in which 2 models are explored, depicted in Equation 4 as well as in the SCM graph displayed in Figure 2.

$$
\begin{aligned}
Model\ (1):\quad & \mathbb{E}_{Kills}[\mathbb{E}[Duration|Winner, Lineup]] \\
& or \\
Model\ (2):\quad & \mathbb{E}_{Duration}[\mathbb{E}[Kills|Winner, Lineup]]
\end{aligned}
\tag{4}
$$

As both proposed models are simplification of the space, some aspects of the relationship is lost in the abstraction. However, the relationship may display a stronger asymmetry in one of the models, given additional confounding variables which may not be trivially understood. In this case, a ML model may perform more accurately and more consistently given one of the two input features as modeled by the simplification proposed. This is explored by training a set of ML models - in this case as predictive neural networks (NN), as depicted in Equation 5.

$$
\begin{aligned}
Model1 \Rightarrow\quad & NN_1 : (Winner, Lineup) \mapsto Duration; \qquad NN_2 : (Winner, Lineup, Duration) \mapsto Kills \\
Model2 \Rightarrow\quad & NN_3 : (Winner, Lineup) \mapsto Kills; \qquad NN_4 : (Winner, Lineup, Kills) \mapsto Duration
\end{aligned}
\tag{5}
$$

By comparing the averaged performance of both models, it may be possible to identify a stronger asymmetry between variables. This can lead to an exploration of the space, which can be integrated into the broadcast of esport content. In this exploration, rather than attempting to predict the Winner, a depiction of both possible outcomes is displayed, with the predicted necessary State for both teams being identified, such that $\mathbb{E}[State|Winner = 1, Lineup]$ and $\mathbb{E}[State|Winner = 0, Lineup]$. Where State is defined by the resulting features as identified by the best performing series of models (1) or (2) based on the combined output by their respective NNs.

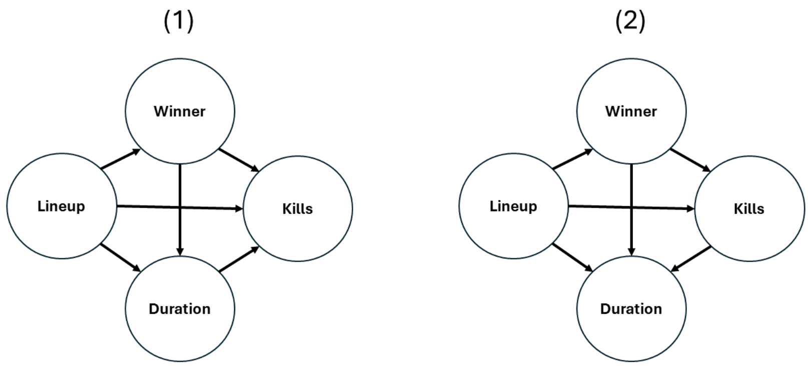


Fig. 2. A SCM representation of the proposed simplification of the space

7 Modeling Win Condition

Once the main narrative features are identified, ML models can be used to represent and evaluate win condition. In order to train such models, appropriate data must be collected.

Firstly, to account for patch changes and patch specific metas, the use of a patch-aware representation of heroes is utilised. As described in the literature, by leveraging game design data through patch notes, it is possible to perform patch agnostic analysis by changing the way in which characters are represented to account for their capabilities, rather than other forms of encoding (such as character IDs). Thus CCR was used, retrieved from the existing literature as provided by the authors previous work [24].

Data from all professional games for Dota 2 for which CCR was available at the time of the study was collected in order to train and evaluate ML models. This was collected through the Open Dota platform, in the form of SQL queries, which can be extracted free of charge using the explorer feature [7]. Data collected for those games included:

- IDs of character selected (later encoded into the CCR equivalent [24])
- Patch in which the match took place (used to translate character IDs into CCR)
- Total number of kills obtained by the Radiant team (i.e. team 1)
- Total number of kills obtained by the Dire team (i.e. team 2)
- Match duration in seconds
- is team 1 winner (represented as a binary variable - note that it is not possible to draw in *Dota 2*).

These features can then be related to the states defined in Section 6.2, where: The Lineup feature is represented by the CCR of the selected characters. Since the CCR of a character is representative of their intrinsic capabilities for a given patch, a machine learning model can leverage the feature to identify relationships and mechanics that are connected to what a player may perceive as the meta. Furthermore, in the same way that $Player \in U$ is assumed, it is also assumed that professional players will generally select character they identify as in meta. Thus, the data being used to train a neural network would be curated by the players themselves to reflect the meta, and only the

[7]https://www.opendota.com/explorer

interactions of capabilities would need to be detected by a ML model. This is facilitate by the way in which CCR is designed, and by the patch-aware values it provides. The remaining variables (Winner, Duration and Kills) are represented by their respective values directly, where Winner=1 when Radian (team 1) wins, and Winner=0 when Dire (team 2) wins, Kills is the total number of kills acquired by each team and Duration is the duration of a match in seconds.

A total of *59,019* games were collected for this study from 6 distinct patches (7.27 to 7.32). Only games which reached an end state were included, to ensure no games with server errors or other technical issues were included.

In order to train and evaluate the models proposed, four neural network models were trained (NN1-4) as described in Section 6.2. NN1 utilised the Lineup (encoded with CCR) and the game outcome (Winner) as input features to predict the final duration of the game. NN2 utilised the same features as well as the match duration in seconds to predict the final score for each team. These two NN combined represent Model 1, which assumes a stronger asymmetry between Duration and Kills (i.e Duration is easier to predict without Kills than Kills are to predict without Duration).

Conversely, NN3 utilised the outcome and Lineup to predict the total number of kills for both teams, while NN4 utilises both variables as well as the number of kills to predict the duration. This constitutes Model 2, which assumes a stronger asymmetry between Kills and Duration. Comparing the performance of the four NNs allows for an investigation of the two proposed simplifications for the structure causal models.

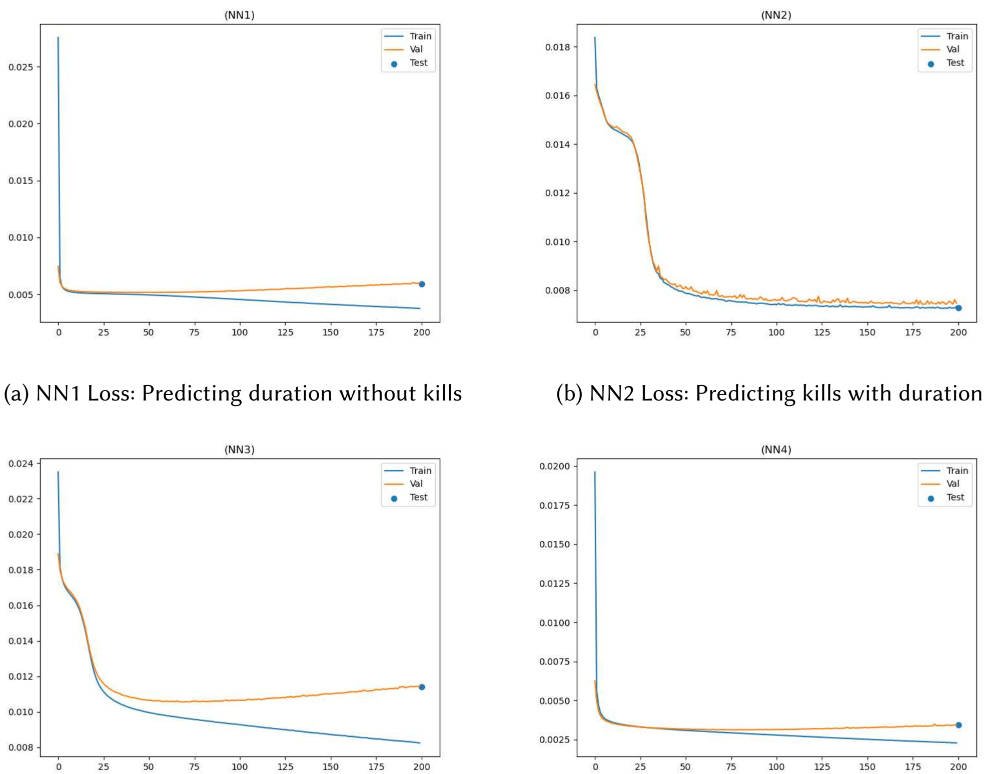
All four models were trained using the same architecture, consisting of 4 hidden layers of [128, 64, 32, 8] neurons each. ML models were trained for 200 epochs with a batch size of 512 using an Adam optimizer. The dataset was divided using a 60/20/20% split between training, test and validation datasets respectively, with the loss graphs depicting the training process available in Figure 3.

As demonstrated in Figure 3, the training and validation graphs for NN3 indicate that the model started to overfit to the data rapidly, with a noticeable deviation between the training and validation graphs by the 50th epoch. Similarly, graphs for NN1 and 4 both demonstrate that the networks were affected by overfitting with a noticeable deviation by the 100th epoch. While Figure 3b does not demonstrate any apparent overfitting, the overall performance of the network had plateaued rapidly by the 50th epoch.

Furthermore, it can be noted that both models that predict the duration (NN1 and 4) outperformed the models attempting to predict the number of kills. The test loss of NN1 can be observed as approximately 6e-3 and NN4's at 4e-3. In contrast, the loss observed in the test dataset for NN3 is approximately 1e-2, where even when provided the duration of a match, NN2's test loss was observed at approximately 7e-3. While the differences in performance are likely due to the properties of the features, it does suggest that a stronger asymmetry between Duration and Kills is observed (albeit as a consequence of the prediction itself). This is also supported by the values of the training losses themselves when compared between all four NN, as demonstrated in Figure 4, where all loss for training were overlaid for ease of parsing. Furthermore, Table 2 depicts the training, validation and test results observed by the 200th training epoch.

8 Discussion

This paper proposes the use of narrative focused machine learning models to produce a win condition system that is designed to be applied to existing broadcast content present in the esport ecosystem. The aim of the win condition model is to provide context to how each team could win the match, rather than attempting to predict who is most likely to win. Win condition as a conceptual narrative feature has been shown to be present in existing coverage of esport as shown in the analysis of existing esport content (see Section 6.1). When analysing the content of esport



(a) NN1 Loss: Predicting duration without kills

(b) NN2 Loss: Predicting kills with duration

(c) NN3 Loss: Predicting kills without duration

(d) NN4 Loss: Predicting duration with kills

Fig. 3. The graph demonstrating the training, test and validation losses for all 4 NNs individually

Table 2. Train, Validation and Test loss values (to 4 decimal values) of all 4 NN as of the 200$^{\text{th}}$ training epoch.

NN	Training	Validation	Test
NN1	0.0038	0.0059	0.0059
NN2	0.0073	0.0074	0.0073
NN3	0.0083	0.0115	0.0114
NN4	0.0023	0.0036	0.0035

commentary in the ecosystem, speculations of game states and outcomes have emerged as common patterns by professional commentators. By extracting and condensing the features that are already in common place when producing commentary, the ML approach of defining and predicting win condition can be more readily applied to the broadcast framework seen in the domain.

Furthermore, the observed features present in the coverage of esport content is in agreement to what is observed in the literature as reliable features for predicting the outcome. Firstly, the current game iterations and changes to the meta are consistent topics of narrative. Similarly it has been shown that the meta have a significant impact into the win prediction models, as well as relevant

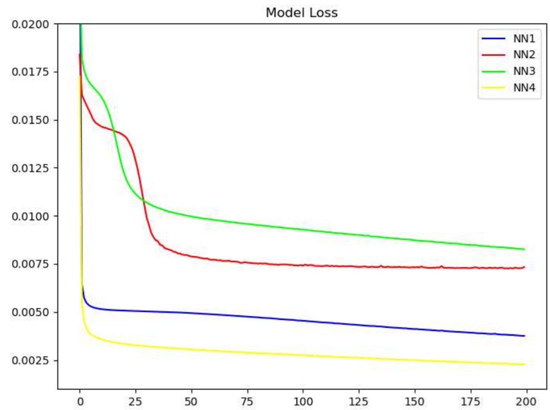


Fig. 4. Loss graphs observed while training all 4 NN combined into one graph

features to understand and represent the domain [17, 24, 32]. Secondly, broadcasters consistently refer to the different stages of the game, stating their expected temporal advantages to different teams. This is understood as a reliable feature for predicting the winner, where the outcome of matches that are exceedingly short or long can generally be predicted more reliably [32].

Lastly, the most predominant feature used in commentary to build a win condition narrative is that of kills. This also reflects the existing literature, where confrontations (and the results of those confrontations) as well as overall score have been shown to be a key metric to produce win predictions [13, 27]. Furthermore, the outcome of blowout score matches have also been proven to be more reliably predicted by ML models in the existing literature [32].

Therefore, the features of importance for winning a game, as perceived by broadcasters, were identified in the SCM in Section 6.2. As observed in Figure 4, match duration as an output generally produces more accurate results than the total number of kills in a match. This is demonstrated as both NN1 and 4 produce more accurate predictions than NN2 and 3. This was observed despite NN2 utilising an additional feature (match duration), which includes knowledge that is only available at the end of the match. In contrast, NN1 does not utilise the knowledge of kills as an input, which contributes to demonstrating this phenomenon. Therefore, as demonstrated by the experiment, Duration seems to be more reliant on Lineup, given the Winner, than Kills. Conversely, Kills is more dependent on Duration, and cannot be as easily ascertained by ML models using only the Winner and the Lineup. This demonstrates an asymmetry in the causal relationships between the variables, where Duration is understood to impact the value of Kills more heavily than the inverse.

In the other hand. Duration as an input feature of Kills prediction seem to have a noticeable impact in the performance of the model, as observed when comparing Figures 3b and 3c. As shown in the differences in training and validation graphs, NN3 is observed to overfit rapidly, with a negative gradient trend line for the training graph and a positive gradient trend in the validation graph by the 50$^{\text{th}}$ epoch. Conversely, NN2 displayed a relatively stable graph by the end of the training period, with no significant deviation between training and validation trends. This suggests that match duration as an additional feature helps prevent the effect of overfitting, thus allowing the NN to more easily generalise to unseen data.

It is also noteworthy that the models utilised in this study were designed using the same architecture. No hyper-parameterization was performed to fine-tune the training process or reduce the effects of overfitting or improve the performance. Therefore, the final obtained models could be improved with continued work in the design and training of the models. However, the results observed by the four NN can be interpreted as a controlled experiment, in which the SCM models proposed in Section 6.2 can be compared and contrasted.

When evaluating Model (1) and (2) it is clear that the State needed for a team to win can be more reliably represented in terms of Model (1). While Model (2)'s NN4 has been seen to produce the most reliable predictions, NN3 is also seen to produce the least reliable results of the four NNs, which creates a large disparity between performance of the predictions of the features. The effects of this disparity is worsen, when considering the use-case would require the output of the least reliable prediction (NN3) being used as an input feature to run NN4, which would impact the reliability of the resulting prediction, since the input data itself may not be reliable. Additionally, despite no attempts to mitigate the effects of overfitting, it is clear that NN2 has been affected the least by this phenomenon and thus can more easily generalise to unseen data. Furthermore, the differences in performance between NN1 and NN4 are amplified by the overfitting observed in NN1. This is expected, as no attempts were made to reduce the effects, therefore, given fine-tuning it may be possible to produce more generalisable predictions of the match duration with results that more closely aligns with what is observed in NN4. While the same argument can be made for NN3, the overall performance of the kills prediction still outlines a stronger asymmetry towards predicting duration, which suggests that Model (1) may be more reliable at predicting and depicting the State for data-driven narratives.

Thus, this paper proposes a model for driving the prediction of win condition which aligns with Model (1) as suggested in Section 6.2. The win condition proposed has been depicted in Figure 5 where the use of NN1 and NN2 have been introduced to a dataflow diagram depicting how the State (as defined by Kills and Duration) could be generated to enhance existing storytelling narratives. In this figure, Winner* represents one of $Winner = 1$ or $Winner = 0$, where in a real world use-case, both values are utilised to produce two distinct results, allowing broadcasters to compare their narratives with the predicted State depicted by Model (1).

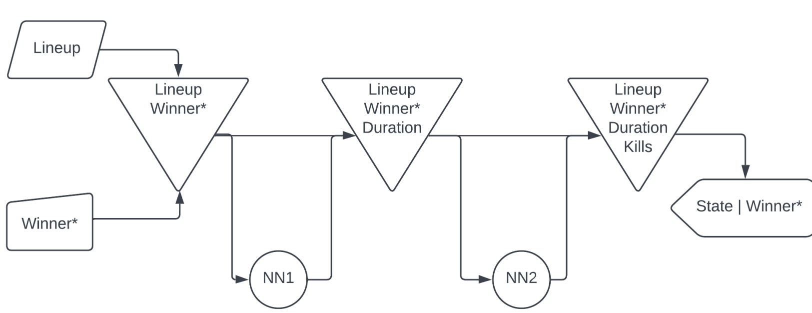


Fig. 5. A dataflow diagram depicting how a prediction that depicts the win condition can be generated in a live scenario.

As the proposed model does not give insights into the most likely outcome, but rather depicts a State to be achieved for a team, it is purely a depiction of goals. This can provide greater narrative opportunities, and particularly as the feature engineering of the model was performed through the features that are already utilised in the existing ecosystem. By leveraging common talking points, and producing results that match commentaries, this model aims to enhance the experience without suffering from the dilemma present in general win prediction models, which can detract from the excitements and user needs of audiences and broadcasters.

8.1 Utilising Win Condition

In order to aid the understanding of how Win Condition, as defined by this paper, can be integrated into a *Dota 2* esport broadcast, some hypothetical example use-cases can be discussed. It is important to highlight, however, that the proposed integration methods are hypothetical, and more work needs to be done to assess the impact of win condition to an actual broadcast, as well as more formally designing and implementing any visualisation tools. However, this section provides some possible ways in which integration could occur, to aid in future development of a Win Condition system and discusses some possible implications.

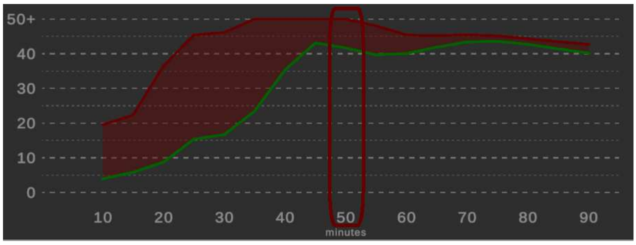
There are many ways in which Win Condition could be integrated into broadcast. One simple and unintrusive way is to use the system defined in Figure 5 to generate textual information and present them to analysts and broadcasters. In this scenario NN1 would produce two predictions for the duration (one for each team as a winner), which is then used as input for NN2, which would also generate two predictions (one for each team as a winner). This could then be formatted in sentences - either manually with a human operator, or through an automated system - and presented to broadcasters. As an example, data from ESL Game 3 was used in the proposed Win Condition system. For comparison, the actual final result of this game concluded with a Radiant [team 1] victory at 21 minutes, 36 seconds with a kill score of 23 to 7 in favour of Radiant. However, in order to increase human readable, all Duration predictions (i.e. the output for NN1) were round to the nearest 5 minutes. This step will undoubtedly have an impact in the kills prediction system (i.e. NN2), which is a trade off between performance and user-needs. The generated outputs were manually formatted, and are presented as follows:

- The Radiant team [team 1] requires more kills than Dire [team 2], predicted to win at approximately 20 minutes, with 25 kills vs. 9 kills to Dire.
- The Dire team [team 2] requires fewer kills than Radiant [team 1], predicted to win at approximately 50 minutes, with 53 kills vs. 42 kills to the Radiant.

This use-case would require minimal amount of training and integration with existing broadcasters, especially if an automated way of generating the sentences is produced. Analysts would simply be presented with textual information and they may choose to integrate it with their existing narrative. However, this use-case provides minimal exploration. One possible alternative that may require some training by broadcasters is to integrate the results of the networks with a visualisation system. Figure 6 depicts how two graphs could be generated following an exploration. In this case, NN2 is run multiple times to predict the kills condition needed for each team at a wide range of match durations. The duration condition (as predicted by NN1) is also drawn, to showcase the primary win condition identified by the system.

In this proposed visualisation, each duration contains two lines connected together by a semi-transparent shading. The green line indicates the predicted amount of kills for Radiant, while the red line indicates the predicted amount of kills for the Dire (output of NN2). Graph (a) and (b) are read independent of each other and refer to a particular team's win condition - i.e. winner=1 for (a) and winner=0 for (b). In this graph, a wide difference between the two predicted lines (such as what is observed in Figure 6b at 30 minutes) indicates that the difference in kills between the two teams must be large for a team to achieve their win condition. Conversely, a shallow difference between the two predicted lines (Figure 6b at 80 minutes) suggests that a team does not require a significant amount of kills more than the enemy team.

(a) Kills predicted per duration (NN2) for Radiant [team 1] as a winner.



(b) Kills predicted per duration (NN2) for Dire [team 2] as a winner.

Fig. 6. A kills prediction graph, outlining the win condition needed (as predicted by NN2) for each team to win at any given time. Lines in green outline the Radiant [team 1] predicted score, while lines in red outline the Dire [team 2] predicted score. The expected duration predicted by (NN1) has also been highlighted with a green/red box for each team.

Lastly, another way in which a win condition system can be integrated is alongside a typical win prediction system, to enhance storytelling. In this case, the win prediction used would require the line-up, match duration and score as inputs, and produces a confidence value for weather Radiant [team 1] will win. A full exploration system can then be performed, and can be used in a live game to enhance the narratives. Following the same exploration methodology depicted in the previous example, a wide range of scores can be used, in addition to duration to identify a team's win condition and measure weather that has been met or not. In order to visualise this, a heatmap per duration can be generated. Figure 7 provides an example use-case for ESL Game 3 for three timestamps (10, 15 and 20 minutes). As the heatmaps display, the model predicted a Radiant victory at all timestamps, however the distance of the dot from the white region at 10 minutes (Figure 7a) is significantly shorter when compared to minute 15 (Figure 7b) and then further increased by 20 minutes (Figure 7c). This exploration enables the user to visualise that the Radiant team is more likely to win, and how far ahead they are compared to where they could be given a different set of score/time (for the given lineup).

All of the example visualisations and insights provided have been generated from the models proposed (NN1 and 2) using ESL Game 3's lineup. Therefore, it is possible to discuss how they could be implemented alongside the narratives of the game. As explored in Section 6.1, the following two quotes were extracted from the game's draft phase:

> it's not like they [team 1] can't win late, but I definitely do favor them [team 2] in the later portion of the mid game more so
>
> ...

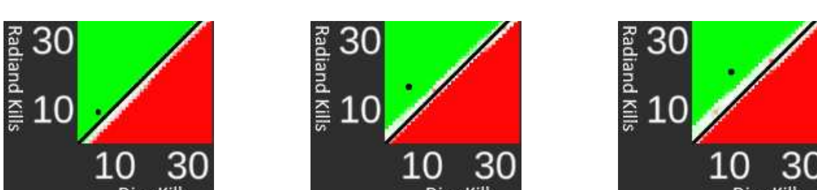
> Liquid [team 1] definitely, effectively saying, guys, we're gonna be here in 25 minutes, and 9Pandas [team 2] need to weather the storm and buy enough time

In this case, the most simplistic textual insights could enhance both narrative, as it affirms the broadcaster assessment that Radiant [team 1] is more favorable to win at earlier stages. The graphs depicted on Figure 6a could also depict how Dire [team 2] would need to perform well at the earlier stages of the game to prevent Radiant [team 1] from reaching their win condition, being able to comment on the amount of kills needed by the Dire to "weather the storm and buy enough time". Similarly, utilising the heatmap presented in Figure 7 in a live setting (i.e. during game) would highlight how the Dire [team 2] was not performing sufficiently well by 10 minutes to last long enough to turn the tide in their favour.

8.2 Limitations & Future Work

While the win condition proposed in this paper has been designed to enhance the esport narrative of broadcast, continued work in the subject can improve the impact of this research. This paper has presented some possible first steps into the ways Win Condition can be integrated into broadcast, however continued work must be carried to fully investigate this. Particularly in evaluating the impact of win condition into the broadcast and ultimately the audience experience, as well as any ethics implications in the use of this system, and design considerations. Therefore, a full investigation of visualisation and integration methods, followed by user studies is an imperative continuation step towards win condition in esports research, to enhance audience experience.

Additionally, the proposed win condition model was defined through the use of a SCM formulated from investigating the user needs of a selection of broadcasters, by studying existing narratives. As this was the first step in this topic, this paper limited the evaluation to the draft-phase, as it provides the most speculative narrative, as the game has not yet started. This allowed for an investigation of what broadcasters perceive as important features that define condition. However, other user needs may arise from continued work looking at in and post game analysis including commentary and posy-game panels as well as utilising a different set of broadcasters. Furthermore, SCM graphs provide a good way to explore causal relationships, however this is done through a series of abstractions. While this has enabled an in-depth investigation of the selected features, continued work in feature refining can lead to better forms of win condition, following the same patterns observed in similar topics, such as win prediction.



(a) Minute 10 (b) Minute 15 (c) Minute 20

Fig. 7. A hypothetical exploration example combining a traditional Win Prediction model with the Win Condition System proposed in this paper. Heatmap outlining the chance of Radiant win given any set of scores from 0 to 30 kills per team. Predictions in favour of Radiant (high heat) are coloured in green, prediction in favour of Dire (low heat) are coloured in red and predictions with no significant favour (mid heat) are coloured in white. Heatmap generated for each timestamp is given, and the actual score observed at those periods are plot as a black dot.

Furthermore, the steps outlined om this paper focus on *Dota 2* narratives. While other similar titles (particularly ones in the same MOBA genre) may have similar features that would allow for an application of the findings, future work should investigate the concept of win condition in similar terms, but in relation to other titles, to allow for a more broad understanding of win condition as a narrative tool across a wider esport spectrum.

9 Conclusion

In conclusion, it is clear that both esport expert broadcasters (see Section 6.2 and the academic knowledge available in the literature places key importance on score, timings and meta as core factors that impact the outcome of a match [16, 24, 27, 32]. Despite this, some evidence that the insights obtained from win prediction models are not fit to the user-needs in relation to aiding stories or commentaries [13, 33, 37]. Conversely, this paper outlined how win condition as a narrative topic is a prominent part of the commentary coverage of esport content (see Section 6.1). Thus, this paper draws from existing win prediction literature and the user-needs identified to design a win condition system. By reverting the flow of data that produces win predictions, the proposed win condition model uses an exploration of all possible outcomes to predict and outline the game state needed for each team to win, based on features that are already widely utilised by experts to describe the game state. Thus, while previous work in win prediction purely gives a confidence-score to the most likely winner, the proposed model leverages narrative explorations to enhance storytelling while focusing on the needs of broadcasters and audiences, aiming to enhance the experience of consuming esport content.

Acknowledgments

This work was supported by the EPSRC Centre for Doctoral Training in Intelligent Games & Games Intelligence (IGGI) EP/S022325/1.

References

[1] Sabbir Ahmad, Andy Bryant, Erica Kleinman, Zhaoqing Teng, Truong-Huy D. Nguyen, and Magy Seif El-Nasr. 2019. Modeling Individual and Team Behavior through Spatio-temporal Analysis. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Barcelona, Spain) *(CHI PLAY '19)*. Association for Computing Machinery, New York, NY, USA, 601–612. https://doi.org/10.1145/3311350.3347188

[2] R. Ani, Vishnu Harikumar, Arjun K. Devan, and O.S. Deepa. 2019. Victory prediction in League of Legends using Feature Selection and Ensemble methods. In *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*. 74–77. https://doi.org/10.1109/ICCS45141.2019.9065758

[3] Florian Block, Victoria Hodge, Stephen Hobson, Nick Sephton, Sam Devlin, Marian F. Ursu, Anders Drachen, and Peter I. Cowling. 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* (SEOUL, Republic of Korea) *(TVX '18)*. Association for Computing Machinery, New York, NY, USA, 29–41. https://doi.org/10.1145/3210825.3210833

[4] Kabir Bubna, Michael G. Trotter, Remco Polman, and Dylan R. Poulus. 2023. Terminology matters: defining the esports athlete. *Frontiers in Sports and Active Living* 5 (2023). https://doi.org/10.3389/fspor.2023.1232028

[5] Christian Carlsson and Alex Pelling. 2015. Designing spectator interfaces for competitive video games. Master's thesis, Report No. 2015:129, Chalmer University of Technology, Gothenburg, Sweden.

[6] Marcus Carter and Ben Egliston. 2021. The work of watching Twitch: Audience labour in livestreaming and esports. *Journal of Gaming & Virtual Worlds* 13, 1 (2021), 3–20. https://doi.org/10.1386/jgvw_00025_1

[7] Sven Charleer, Kathrin Gerling, Francisco Gutiérrez, Hans Cauwenbergh, Bram Luycx, and Katrien Verbert. 2018. Real-Time Dashboards to Support eSports Spectating. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play* (Melbourne, VIC, Australia) *(CHI PLAY '18)*. Association for Computing Machinery, New York, NY, USA, 59–71. https://doi.org/10.1145/3242671.3242680

[8] Alan Pedrassoli Chitayat, Athanasios Kokkinakis, Sagarika Patra, Simon Demediuk, Justus Robertson, Oluseji Olarewaju, Marian Ursu, Ben Kirmann, Jonathan Hook, Florian Block, et al. 2020. WARDS: Modelling the Worth of Vision in MOBA's. In *Intelligent Computing: Proceedings of the 2020 Computing Conference, Volume 2*. Springer, 63–81.

[9] Lincoln Magalhães Costa, Anders Drachen, Francisco Carlos Monteiro Souza, and Geraldo Xexéo. 2023. Artificial Intelligence in MOBA Games: A Multivocal Literature Mapping. *IEEE Transactions on Games* (2023), 1–23. https://doi.org/10.1109/TG.2023.3282157

[10] Simon Demediuk, Athanasios Kokkinakis, Moni Sagarika Patra, Justus Robertson, Ben Kirmann, Alistair Coates, Alan Chitayat, Jonathan Hook, Isabelle Nölle, Oluseyi Olarewaju, Daniel Slawson, Marian F. Ursu, Florian Block, and Anders Drachen. 2021. Performance Index: A New Way to Compare Players. In *Proceedings of the 2021 MIT Sloan Sports Analytics Conference*. IEEE, United States. 2021 MIT Sloan Sports Analytics Conference.

[11] Simon Demediuk, Peter York, Anders Drachen, James Alfred Walker, and Florian Block. 2019. Role Identification for Accurate Analysis in Dota 2. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE '19)* 15, 1, 130–138. https://doi.org/10.1609/aiide.v15i1.5235

[12] Jessica Formosa, Nicholas O'Donnell, Ella M. Horton, Selen Türkay, Regan L. Mandryk, Michael Hawks, and Daniel Johnson. 2022. Definitions of Esports: A Systematic Review and Thematic Analysis. *Proc. ACM Hum.-Comput. Interact.* 6, CHI PLAY, Article 227 (oct 2022), 45 pages. https://doi.org/10.1145/3549490

[13] Victoria J. Hodge, Sam Devlin, Nick Sephton, Florian Block, Peter I. Cowling, and Anders Drachen. 2021. Win Prediction in Multiplayer Esports: Live Professional Match Prediction. *IEEE Transactions on Games* 13, 4 (2021), 368–379.

[14] Md Naimul Hoque and Klaus Mueller. 2022. Outcome-Explorer: A Causality Guided Interactive Visual Interface for Interpretable Algorithmic Decision Making. *IEEE Transactions on Visualization and Computer Graphics* 28, 12 (2022), 4728–4740. https://doi.org/10.1109/TVCG.2021.3102051

[15] Wooyoung (William) Jang, Kevin K. Byon, and Hyunseok Song. 2021. Effect of Prior Gameplay Experience on the Relationships between Esports Gameplay Intention and Live Esports Streaming Content. *Sustainability* 13, 14 (2021). https://doi.org/10.3390/su13148019

[16] Adam Katona, Ryan Spick, Victoria J. Hodge, Simon Demediuk, Florian Block, Anders Drachen, and James Alfred Walker. 2019. Time to Die: Death Prediction in Dota 2 using Deep Learning. In *2019 IEEE Conference on Games (CoG '19)*. 1–8. https://doi.org/10.1109/CIG.2019.8847997

[17] Athanasios Kokkinakis, Peter York, Moni Sagarika Patra, Justus Robertson, Ben Kirman, Alistair Coates, Alan Pedrassoli Pedrassoli Chitayat, Simon Demediuk, Anders Drachen, Jonathan Hook, et al. 2021. Metagaming and metagames in esports. *International Journal of Esports* 1, 1 (2021).

[18] Athanasios Vasileios Kokkinakis, Simon Demediuk, Isabelle Nölle, Oluseyi Olarewaju, Sagarika Patra, Justus Robertson, Peter York, Alan Pedrassoli Pedrassoli Chitayat, Alistair Coates, Daniel Slawson, Peter Hughes, Nicolas Hardie, Ben Kirman, Jonathan Hook, Anders Drachen, Marian F Ursu, and Florian Block. 2020. DAX: Data-Driven Audience Experiences in Esports *(IMX '20)*. Association for Computing Machinery, New York, NY, USA, 94–105. https://doi.org/10.1145/3391614.3393659

[19] Klaus Krippendorff. 2018. *Content analysis: An introduction to its methodology*. Sage publications.

[20] Ilya Makarov, Dmitry Savostyanov, Boris Litvyakov, and Dmitry I. Ignatov. 2018. Predicting Winning Team and Probabilistic Ratings in "Dota 2" and "Counter-Strike: Global Offensive" Video Games. In *Analysis of Images, Social Networks and Texts*. Springer International Publishing, Cham, 183–196. https://doi.org/10.1007/978-3-319-73013-4_17

[21] Yue Meng-Lewis, Gavin Lewis, Zhi Lin, and Yupei Zhao. 2022. Examination of Esports Fans' Live Streaming Experiences Using an Extended Expectation-Confirmation Model: A Case Study of the King Pro League. *International Journal of Human–Computer Interaction* 0, 0 (2022), 1–16. https://doi.org/10.1080/10447318.2022.2141008

[22] Newzoo. 2022. *Newzoo's Global Esports & Live Streaming Market Report 2022 | Free Version*. https://newzoo.com/resources/trend-reports/newzoo-global-esports-live-streaming-market-report-2022-free-version

[23] Judea Pearl et al. 2000. Models, reasoning and inference. *Cambridge, UK: CambridgeUniversityPress* 19, 4 (2000), 675–685. https://doi.org/10.1017/S0266466603004109

[24] Alan Pedrassoli Chitayat, Florian Block, James Walker, and Anders Drachen. 2023. Beyond the Meta: Leveraging Game Design Parameters for Patch-Agnostic Esport Analitics. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE '23)*. 116–125. https://doi.org/10.1609/aiide.v19i1.27507

[25] Alan Pedrassoli Chitayat, Alistair Coates, Florian Block, Anders Drachen, James Alfred Walker, James Dean, Mark McConachie, and Peter York. 2024. From Passive Viewer to Active Fan: Towards the Design and Large-Scale Evaluation of Interactive Audience Experiences in Esports and Beyond. In *Proceedings of the 2024 ACM International Conference on Interactive Media Experiences*. 94–107.

[26] Charles Ringer, Sondess Missaoui, Victoria J Hodge, Alan Pedrassoli Chitayat, Athanasios Kokkinakis, Sagarika Patra, Simon Demediuk, Alvaro Caceres Munoz, Oluseji Olarewaju, Marian Ursu, et al. 2023. Time to die 2: Improved in-game death prediction in dota 2. *Machine Learning with Applications* 12 (2023), 100466.

[27] Matthias Schubert, Anders Drachen, and Tobias Mahlmann. 2016. Esports analytics through encounter detection. In *Proceedings of the MIT Sloan Sports Analytics Conference*, Vol. 1. MIT Sloan Boston, MA, 2016.

[28] Antonio Luis Cardoso Silva, Gisele Lobo Pappa, and Luiz Chaimowicz. 2018. Continuous outcome prediction of league of legends competitive matches using recurrent neural networks. *SBC-Proceedings of SBCGames* (2018), 2179–2259.

[29] Katie Salen Tekinbas and Eric Zimmerman. 2003. *Rules of play: Game design fundamentals.* MIT press.

[30] Sunny Thaicharoen, Jeremy Gow, and Anders Drachen. 2023. An Ecosystem Framework for the Meta in Esport Games. *Journal of Electronic Gaming and Esports* 1, 1 (2023), jege.2022–0045. https://doi.org/10.1123/jege.2022-0045

[31] Marko Tot, Michelangelo Conserva, Alan Pedrassoli Chitayat, Athanasios Kokkinakis, Sagarika Patra, Simon Demediuk, Alvaro Caceres Munoz, Oluseji Olarewaju, Marian Ursu, Ben Kirmann, et al. 2021. What are you looking at? Team fight prediction through player camera. In *2021 IEEE Conference on Games (CoG '21)*. IEEE, 1–8. https://doi.org/10.1109/CoG52621.2021.9619038

[32] Markos Viggiato and Cor-Paul Bezemer. 2020. Trouncing in dota 2: An investigation of blowout matches. In *Proceedings of the AAAI conference on artificial intelligence and interactive digital entertainment (AIIDE '20)*, Vol. 16. 294–300. https://doi.org/10.1609/aiide.v16i1.7444

[33] Minglei Wang. 2023. The Communication Effectiveness of AI Win Prediction Applied in Esports Live Streaming: A Pilot Study. In *Computer-Human Interaction Research and Applications*. Springer Nature Switzerland, Cham, 315–325. https://doi.org/10.1007/978-3-031-49368-3_19

[34] Wenche Wang and Jiajia Fan. 2022. Topic mining of real-time discussions: what catches the attention of live-streaming esports viewers? *European Sport Management Quarterly* (2022), 1–22. https://doi.org/10.1080/16184742.2022.2117397

[35] Mark HM Winands, Yngvi Björnsson, and Jahn-Takeshi Saito. 2008. Monte-Carlo tree search solver. In *Computers and Games: 6th International Conference, (CG '08)* (Beijing, China). Springer, 25–36. https://doi.org/10.1007/978-3-540-87608-3_3

[36] Peter Xenopoulos, João Rulff, and Claudio Silva. 2022. ggViz: Accelerating Large-Scale Esports Game Analysis. 6, CHI PLAY, Article 238 (oct 2022), 22 pages. https://doi.org/10.1145/3549501

[37] Zelong Yang, Zhufeng Pan, Yan Wang, Deng Cai, Shuming Shi, Shao-Lun Huang, Wei Bi, and Xiaojiang Liu. 2022. Interpretable Real-Time Win Prediction for Honor of Kings—A Popular Mobile MOBA Esport. *IEEE Transactions on Games* 14, 4 (2022), 589–597. https://doi.org/10.1109/TG.2022.3149044

Received February 2024; revised June 2024; accepted July 2024