

# Machine Learning and Data mining on the innovation of E-sports industry

Yiqing CHEN<sup>1</sup>, Jie MEI<sup>1</sup>, Steward HUANG<sup>2</sup>

<sup>1</sup>Master Student of AIBT program at SKEMA Business School Paris, France

<sup>2</sup>Director of MS of AIBT program at SKEMA Business School Paris, France  
August 2020, Master Thesis of AIBT program at SKEMA Business School  
Paris, France

Received: August 17, 2020. Revised: November 16, 2020. Accepted: November 23, 2020. Published: November 24, 2020.

**Abstract**—AI technology brings many revolutionary innovation opportunities to the e-sports industry. With the help of data mining, we can analyze the advantages and disadvantages of competitors, and predict the trend of the situation in the future. With the help of the agent created by intensive in-depth learning, it can assist players of different levels to carry out routine training, so as to improve the overall activity of the game. With the help of AI's big data advantage, AI can assist E-sports teaching to regard E-sports specialty as an experimental platform for using cutting-edge technology to reform and innovate traditional education and provide forward-looking guidance for future education. This paper uses CNN, LSTM, and LSTM + CNN three model to predict the outcome of the game according to the heroes selected by both teams, and has achieved good prediction results.

**Keywords**—Machine Learning, E-sports, Data mining, data analyst, DOTA2, CNN, RNN, LSTM, MOBA games, Depth Learning.

## I. INTRODUCTION

E-sports is a kind of intellectual antagonism among people, which uses high-tech hardware and software equipment as sports machinery; through sports, it can exercise and improve the thinking ability, reaction ability, coordination ability of mind, eyes and limbs and willpower of participants, and cultivate team spirit. “Standardization refers to the unification of repetitive things and concepts through the formulation, release and implementation of standards in the social practice of economy, technology, science and management, so as to obtain the best order and social benefits” [1].

According to Newzoo, a company that specializes in digital media and market data, the total number of global E-sports audiences will grow to 495 million in 2020 [2]. Among them, core E-sports fans will account for 223 million, with a

year-on-year growth of 25 million and will grow to 295 million in 2023 with an annual compound growth rate of 11.3% (2018-2023). Meanwhile, by 2020, the number of occasional noncore viewers will reach 272 million, up from 245 million in 2019. E-sport's audience attendance is also amazing. In 2017, E-sports officially became the official event of the 2022 Asian Games [3], indicating that the social recognition of E-sports has greatly increased.

With the rapid development of e-sports industry, AI is quietly changing the development mode of this industry [4], AI can not only bring revolutionary opportunities for the development of e-sports, but also E-sports provide excellent opportunities for the development of AI. “Rather, games provide ideal environments that simulate the real world. AI researchers can conduct experiments in games and transfer successful AI ability to the real world.” [5]. Therefore, the application and development of AI in E-sports is not only for games, but also has a long-term strategic significance for real life.

In 2017, AlphaGo defeated world Go champion Ke Jie, which symbolized the end of AI for today's Go world [6]. However, compared with the complexity of RTS (Real Time Strategy) type E-sports games, the difficulty of Go is dwarfed. “There are four major aspects that make RTS games much more difficult compared to GO: Computational complexity; Multi-agent. Playing; Imperfect information; Sparse and delayed rewards.” [4]. Therefore, the field of E-sports is still a blue ocean for AI, and it becomes an excellent place to test AI performance.

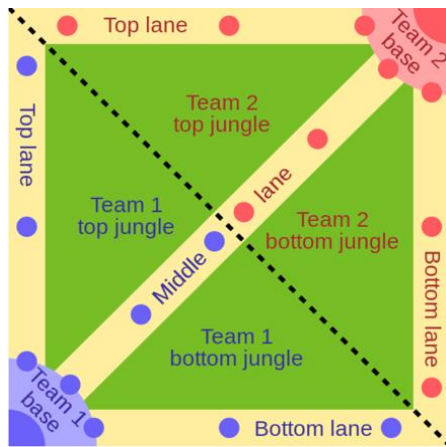


Figure 1: The general layout of a typical MOBA game

Different from the development direction of AI from simple to complex, the development path of E-sports in recent years has always been from complex to simple [7]. The biggest turning point of E-sports development is the comprehensive transformation of RTS games to MOBA (Multiplayer Online Battle Arena) games. MOBA is a 5v5 real-time strategy game. Players control a character in the game, that is, "hero" (take DOTA2 as an example, select from 119 hero pools), and push it to the enemy base for the ultimate purpose of winning. The whole process ends in more than 40 minutes [8]. 30% of online game players in the world are loyal fans of MOBA and especially the growth speed of mobile MOBA is the first [9]. This phenomenon proves that E-sports gradually spread from the core players to the pan entertainment players, and the threshold of Esports is getting lower and lower.

Compared with PC, mobile MOBA game has three advantages: lower equipment requirements, fragmented entertainment mode and stronger social media function [10]. As long as players use smart phones, they can participate anywhere and anytime in E-sports. Undoubtedly, the threshold of E-sports has been lowered to the lowest level. As the leader of mobile MOBA games in recent years, Honor of Kings downloaded more than 800 million times in the App market, with more than 200 million registered users [10]. Unlike DOTA2 and LOL, which have steep learning curve and high complexity, Honor of Kings maximizes the player's need for fragmented entertainment while maintaining the basic principles of competitive play. However, the injection of a large number of noncore players will also cause a large number of loss of novice players, thus affecting the life cycle of the game [11]. Therefore, the emergence of AI as an auxiliary tool will bring long-term sustainable power for the development of E-sports.

To sum up, this paper will focus on the study of the industrial structure changes of e-sports industry in the booming environment of artificial intelligence. At the same time, the significance of E-sports games for AI algorithm evolution and iteration is elaborated in depth.

## II . LITERATURE REVIEW

The purpose of this Literature Review is to benchmark the existing research and explore the innovation of E-sports in the era of AI. First of all, the relevant papers are reviewed, and systematic exploration is carried out to extract similarities, gaps and main research results. The results show that MOBA is the most discussed game type among all the e-sports game categories, and player experience and poisoning behavior are the hot topics of research. Secondly, because MOBA game has a huge group of fans, but the researchers have not carried out innovative exploration on its development path in the AI era, as well as the projection and influence of E-sports on game designers in the AI era, so they put forward suggestions for future research.

### A. Dynamic Difficulty Adjustment system

With the continuous development of the game industry, the revenue generated by this industry has exceeded the music and movie industry. Games are considered to be the most important source of entertainment for contemporary humans. Therefore, more and more resources are invested in this industry. Games The effect has been continuously improved, and the player's experience has been continuously improved. In order to allow players to have a better game experience and continue to maintain interest in the game, dynamic difficulty adjustment technology has emerged. This technology can change the difficulty of the game according to the actual operation level of the player, so that the player remains enthusiastic about the game. We set the player's skills and challenges at about 1: 1, so that players can continue to produce a pleasant game experience; when the game sets the challenge higher than the player's actual skills, the player often feels frustrated; when the game sets When the challenge is lower than the player's actual skills, players tend to get bored because the game is too simple. Therefore, controlling the player's chance of winning is about 50% to better keep the player interested in the game.

However, this theory is generally applicable to most games, not to MOBA games. According to Johnson et al. (2015) [12] and Kwak et al. (2015) [13], MOBA games will have more frustration and more challenge. Andrade, Santana and Jussieu (2005) [14] mentioned three requirements: Rapid adaption to the player's initial level of skill or competency; Track evolution and regression in the player's performance; Remain believable in the elements that are modified.

In order to allow computer opponents to have a similar level of skill as the player or to allow the computer to generate the appropriate level of difficulty, we must first evaluate the player's skill level. In different games, the data that needs to be evaluated for players is different. We need to find data that can accurately reflect the level of authority of the player. It is worth noticing that currently there are no DDA systems which would cover all game genres. We will analyze two different types of game, console game, multiplayer online game.

#### a. Console game

In a console game, we can calculate the player's clearance time and skill hit rate, kill efficiency, etc. to evaluate the player's skill level, and balance with the player's skills by adjusting the robot's blood volume, movement speed, number, damage, and task difficulty. For instance, Fallout and The Elder Scrolls use the player's level as the representative of the ability and adjust

the strength of the enemy accordingly. when the player's level gets higher, the strength of the enemy will get higher as well, which will make player stay in the flow channel.

Another example is Resident Evil 4, During the game, the system will give players a real-time rating based on the player's performance (a hidden value, there are always 0 ~ 15 total 16 levels). The factors that affect the hidden value are the player's hit rate, the frequency of damage, the size of the injured value, the number of players killed by the enemy, the number of consecutive times, etc. The factors affected by the hidden value are the enemy's physical strength, the enemy Attack power, the enemy's desire to attack, where the enemy appears, the player's attack power, the difficulty of QTE, and so on. That means if the player plays smoothly along the way, the game's level will become very high, and the difficulty will be corrected upwards; if the player is stumbling, then the game's level will be lowered, and the difficulty will be revised down.

The Mario Kart racing game series use "Rubber banding" to balance the challenge and skill, the speed of robot car is relative to the speed of player car, when the player is faster the robot will speed up, when the player is slower the robot will slow down. It is a double-edged sword; it will make the game more competitive when the players don't know "Rubber banding", however, after players knowing the fact, players will lose their interest immediately.

#### b. Multiplayer online game

Comparing with the console game, multiplayer online game is more complex and More diverse. However, some existing games use more direct methods similar to Dynamic Difficulty Adjustment (DDA) in a multiplayer context [15].

Also, in The Mario Kart multiplayer racing game series, with the random item system, the player's final score is mainly determined by the player's skill level, but the item obtained in the game will also have a certain impact on the player's final score. During the game, DDA will be based on the player's ranking. To adjust the probability of player getting different items. When the player ranks closer to the first place, the probability of the player getting better items will decrease a lot, so that the probability of getting worse items will be greater; on the contrary, if the player ranking is relatively behind, then the player will have a greater probability of getting better items. The game designer narrows the gap between players of different levels by changing the drop probability of different items. The benefits of this design are obvious, which makes the game more competitive and enhances the fun of the game to a certain extent. However, this also reduces the fairness of the game to a certain extent. Players will not get various props with the same probability. Players with the top rankings receive relatively unfair treatment, which violates the fairness principle of e-sports.

Compared to the random item system in the Mario Kart multiplayer racing game series, the ranking mode in most multiplayer online games is much fairer. The most attractive part of e-sports is its fairness, so the most popular mode for players in most multiplayer online games is usually the ranked mode. For MOBA games, dynamic difficulty adjustments similar to those used in stand-alone games are difficult to apply in the game, because the attributes and equipment of all heroes

are unified. It is difficult to control the player's winning rate by adjusting the strength of the opponent like a stand-alone game. Therefore, in this type of game, we usually use a matching mechanism to match the player to a similar level of opponent to let the player stay in the "flow channel". Usually we will use Elo and MMR to make the matching system fairer. Elo (Elo rating system) was creating by Arpad Elo, which is originally created for chess to rate chess players. And it is used to calculate the wining rate of each player and calculate the point of each player winning or losing. The formulate of Elo is below:

$$E_A = \frac{1}{1 + 10^{(R_B - R_A)/400}} \quad E_B = \frac{1}{1 + 10^{(R_A - R_B)/400}}$$

$$R'_A = R_A + k(S_A - E_A)$$

$R_A$ : Player A's current points

$R_B$ : Player B's current points

$S_A$ : Actual win/loss value, win=1, tie=0.5, loss=0

$E_A$ : Expected player A's win or loss

$E_B$ : Expected B player's winning or losing value

Because the E value is also estimated, then  $E_A + E_B = 1$

$R'_A$ : Player A's points after the competition

$K$ : constant, In the World Chess Championship,  $K=16$ ; in most game rules,  $K=32$ . Generally, the higher the level of the game, the smaller the  $K$  value. This is to avoid a few games that can change the ranking of high-end top players.

Let's take an example, there are 2 chess players, player A's current points is 1500, and player B's current points is 1600. So, the winning rate of player A is around 36%, the winning rate of player B is around 64%. If player A wins, the player A will get 20 points and player B will lose 20 points. If player B wins, the player B will get 12 points and player A will lose 12 points.

However, Elo is mainly used in 1v1 scenarios, it is very difficult to predict the outcome in team games, because it doesn't know the individual impact of a team member on the overall outcome of the game. So, for the team game like MOBA games, MMR (Match Making rating) is more used than Elo. MMR is a modified Elo system designed to take individual inputs of your personal MMR and make prediction for players based on the overall or the average MMR of each team. So, the more games the players play, the more accurate the MMR is, and finally reflect the true level of players, the winning rate of players will be closer to 50%. The MMR in a lot of games is never rested but they will reset rank in the new season. This is only designed to keep ranks balanced and top ranks "prestigious" and giving players to keep playing rank and improving.

#### B. AI E-sports TEAM

In view of the development history of the competition between AI and human in the target literature, this paper explores the possibility of establishing AI team and putting into formal E-sports. From the beginning of AI's involvement in chess, AI has been playing intelligent games with human beings. From simple 1v1 to today's OpenAI 5v5, let's put forward the assumption of building a team composed of all artificial intelligence agents.

Aiming at MOBA games, we set up an AI team. First of all, according to the literature analysis, we get the feasibility of

technology. DOTA2 is a well-known MOBA game. Compared with the problems solved by DRL (such as Chess and GO), the game complexity of DOTA2 is higher [16], which is mainly reflected in the following aspects:

**Decision sequence length:** the frame rate of DOTA2 is 30/s, while the average duration of a game is 45 minutes. OpenAI five is set to make an action every 4 frames, so the total number of decisions is  $30 * 60 * 45 / 4 = 20250$ , that is to say, more than 20000 moves may be done in a game, while chess is about 80 steps per game, and GO is about 150 steps per game, which are far lower than DOTA2.

**Some states can be seen in the game,** each party can only see the state of their unit and the area near the building (fog of war), so to achieve a high intensity, it needs to infer based on incomplete data, and conduct opponent modeling.

**High dimensional movement and observation space:** DOTA2's game environment is relatively complex, with 10 heroes, dozens of buildings, soldiers, etc. on both sides, which is roughly equivalent to 16000 different variables (including continuous and discrete types) that can be observed in each step to describe the environment; similarly, after discretizing the action space, the available actions in each step are between 8000 and 80000 (depending on the hero); the same, Chess and go are much simpler than this. In chess, about 1000 variables can be observed in each step, and 35 actions can be selected in each step. In go, about 6000 variables can be observed by customers, and 250 actions can be selected in each step.

At present, the AI team is mainly based on the deep neural network learning technology of OpenAI five, but compared with the conventional Dota2 game, OpenAI has two main limitations: first, there are 117 different heroes in the game, only 17 of them are supported in OpenAI five second, some props in the game allow players to control multiple units at the same time, which is disabled in OpenAI five [17]. Under this limitation, OpenAI Five beat the 2018 world champion OG of DOTA 2 and beat 99.4% of the human competitors in the online open experience segment [18].

In fact, if we build a team based on OpenAI, **can we really compete with professional players for a long time?** Next, we will analyze it from the perspective of game strategy.

In e-sports, reaction time is an important indicator to measure the ability of a professional E-sports player. If AI is significantly shorter than human in reaction time, then the fairness of AI team is questionable. As mentioned earlier, the frame rate of DOTA2 game itself is 30, OpenAI Five processes one timestep every four frames, which is equivalent to reducing the frame rate to 7.5. At the same time, due to the staggered way of timestep communication between the game and the AI model (the game sends out the observation state in the T timestep, and continues to run to the T + 1 timestep, the AI model returns the corresponding action of the T timestep to the game), and due to the delay mechanism in action design (allow In fact, we can think that the response time of OpenAI five has a delay of 5-8 frames, which is about 167ms to 267ms, while the average response time of reference human is about 250ms, so we basically think that what we do in response time is relatively fair [18].

Under the condition of fair response time, we discussed the tactical diversity of AI team. First, open AI achieved the optimal solution of tactics in a limited range of hero pool (18 specified heroes). The 5v5 regiment and encounter battle of AI were very strong, but in some special cases, it was still unable to deal with [18]. OpenAI is not good at coping with the "4/1 split pushing" tactics [19], and the protection of messenger's lags behind. The deadliest thing is that AI is not good at dealing with stealth units. Once human players enter the stealth state, they are equivalent to invincible. When the win rate of OpenAI itself prediction is lower than 50%, OpenAI team's reaction is different from that of human beings. AI team will passively compete and continuously perform irrational actions with high risk and low income until losing the competition [20].

The reason why AI defeated human beings is that AI can realize "**subtle operation**" based on flexible movement, rapid response, seamless connection skills and precise control of blood volume [18]. However, compared with human beings, it still has a considerable distance in terms of tactical diversity. In the face of the traps or stealth units deliberately set by human beings, it still lacks reasonable reasoning ability. Many common-sense human beings can quickly grasp the palm through more than ten games Grip, but it's a skill AI can't understand.

Therefore, if we want to build an AI team that can really compete with human professional teams, **what are the necessary conditions for the existing E-sport AI to improve?**

The most important condition is to combine the computing power with the algorithm and continue to increase the research investment of AI algorithm efficiency [21]. Although the efficiency improvement of AI algorithm is considerable, it is still not enough compared with the number of new algorithm parameters with exponential growth. According to the efficiency of open model All together this produces a combined factorized action space size of up to  $30 \times 4 \times 189 \times 81 = 1,837,080$  dimensions (30 being the maximum number of primary actions we support)." [18], The improvement of algorithm efficiency enables researchers to conduct more experiments in the same time and reduce the cost of funds.

We continue to develop OpenAI five based on effective computing, which provides technical support for the diversity of AI team at the tactical level.

According to the basic test of OpenAI for model efficiency, the latest rerun version of OpenAI five only needs 1/5 of the computing power to beat the previous generation of algorithm, while the time between the two generations is only three months [21]. OpenAI speculates that the efficiency of the algorithm may exceed the gain brought by Moore's law, because it observes that the number of transistors in integrated circuits roughly doubles every two years. For AI tasks that spend a lot of research time and / or computing power, the improvement of algorithm efficiency may exceed the requirement of hardware efficiency. The concept of "effective computing" put forward by OpenAI researchers combines the improvement of computing power and algorithm, and we can directly see the contribution of various fields to the development of artificial intelligence [21].

Outlook: after the completion of the AI team, it will have a significant impact on the existing E-sports competition pattern.

The official event organizers can excavate the outstanding players and teams hidden in the folk through the AI team, so as to promote the promotion of E-sports events and competition fairness.

### C. Data mining on MOBA games

The data analysis of e-sports industry is a very new field, which is mainly used in LOL, DOTA2, CS: GO, OVERWATCH and other competitive projects. The data sources of e-Competition are mainly manual record data (data given by on-site judges, media, etc.), product side data (game data of the majority of players), historical data collection and game uniform data (data of professional players competing) [7]. At present, the main purpose of data operation is to learn from the technical means and application scenarios of traditional competitive sports:

1) Data content output; many social influences of E-sports are derived from data [1].

2) Players' experience; the main means are data resetting of training match, accuracy of skill release and damage caused in group battle, hot spots of hero's moving position, vision inserting and arranging position, ban-pick order and winning rate, real-time items recommendation system, dynamic winner prediction system. Gradually promote such professional data statistical analysis to the product side and provide it to MOBA game's public players [22].

3) The live broadcast effect; the user's understanding of the game and the extraction of highlights are improved through data visualization [23].

There are two different development directions of data mining in E-sports:

1) In the direction of entertainment development, the third-party data service providers (MAX+, DOTA plus, DOTAbuff) provide visual and easy to understand data services for players through cooperation with the game authorities. At the same time, in the e-game gambling and entertainment, artificial intelligence data mining also has great potential value [24].

2) In the direction of professional development, E-sports coaches carry out daily training for club players through data mining, grade the internal players' abilities, and analyze the weaknesses of competitors through competition data mining. such as: "Player A's training condition is not good in the near future, let player B start for the time being"; "The three heroes the enemy is good at in the Mid are A, B and C"; "The enemy's 'Offlaner' rarely ganks Top.". Data analysis provides a quantifiable basis for coaches, which has existed in the e-sports industry from the beginning to the end.

Because of the universality and practicability of data mining technology, MOBA game data analysis service has become a very important part of optimizing the game experience, whether it is a professional team or an ordinary player, but is data analysis really omnipotent in MOBA games?

A large number of players will generate game behavior data during the game process, which can usually be accessed through the API (Application Programming Interface). These data enable researchers to use different machine learning algorithms to develop game AI, develop optimal strategies, and detect game balance problems [25], in fact, the difficulty of data is not

analysis. For most of the e-sports clubs, there is not much available data to analyze.

Although E-sports is the best sport to get data tanks to inherently digital, the problem is that the data is in the hands of game developers and event organizers. In addition to Valve and Riot Games, there are not many game companies providing open professional event data interface, and the API of these two companies has not been greatly updated for many years [8]. Even if the original data or demo of the event are obtained, most E-sports clubs can only analyze the data and realize the technical reserve of productization through the third-party data analysis service. However, for the protection of their own club information, for the match information with absolute value, the club will not deliver the data analysis service to the third party, especially before the period of world-class match. Therefore, the innovation of algorithm technology is important in the development of E-sports data analysis, but more importantly, how to break the "data barrier" and realize the free interaction of data.

Shall we win the game if we do data analysis? It should be noted that the data never tells you how to win but provides a tool. When all people get the same data, they have a deeper understanding of the game. Teams with a deeper understanding of the data can usually see more phenomena through data analysis [26]. At the end of the game, we usually see a panel of injuries, which shows how many injuries a player caused in the process of the game. This is the most intuitive data experience. On this basis, we can calculate the ratio of output to total output, output rate per minute, and economic conversion output efficiency. However, we can not only use one index or a group of data or the same standard to judge the quality of players, which depends on the perspective of data users to understand the game, so it has a strong subjectivity [24]. Data analysis is only an objective tool, which cannot give the most correct answers to the players. Artificial intelligence can achieve relatively correct recommendation and strategy guidance through clustering and regression algorithm [8]. Thus, the real decision-making is the human players themselves.

Even so, the AI based data mining and recommendation system is still an indispensable and powerful tool for the primary and intermediate players in MOBA games [25], "We believe that such a system could bring about a revolutionary method of self-improvement wherein a player is able to better understand his/her shortcomings and pin-point the particular aspects of the game where he/she might be lacking - whether it be too many deaths, too little champion kills, or too little minion kills." [8].

### D. E-sport Bets

The rapid rise of e-sports and the digitization of the gaming industry have led to the rapid development of the e-sports betting industry, and the e-sports betting market is larger than many people think.

According to statistics, in 2015, 590,000 paid event activists participated in real money fantasy e-sports betting, 1.72 million paid events participated in e-sports betting, and about 3.2 million players conducted in-game items through the informal gaming market Bet. By 2020, an estimated 19.4 million bettors

will bet \$ 23.5 billion in eSports, generating \$ 1.8 billion in revenue for operators.

The development of artificial intelligence has made this e-sports betting industry even more powerful. Artificial intelligence can more accurately predict the results of the game, and many people cannot make predictions. And artificial intelligence can make different degrees of prediction for different games.

For example, for CS: GO, artificial intelligence can not only predict the winner of the game, but also predict the exact score of the game. The winner of each map, each the precise score of the map, whether the winner of an individual round, will produce overtime and the total number of rounds, etc. For DOTA2 and LOL, artificial intelligence can accurately predict the winner of the game, the score of the game, the winning team in each round, the game duration of a single round, and the winner of a blood. In addition to predicting the outcome of the game, artificial intelligence can also change the odds of the game in real time.

Of course, while facing the explosive growth of e-sports betting, we should also be careful of the problems and risks that may arise. The first thing to note is the legal issue. As early as 2016, there was a case of gaming esports. In November 2016, the famous gaming company William Hill obtained a license from the Nevada Gaming Administration Committee to list the League of Legends game at IEM Auckland Station as a Gambling items. The well-known international betting sites including William Hill and Pinnacle Sports have opened gambling projects such as "DOTA 2" and "League of Legends". Of course, the gaming industry is not legal in all countries. For example, in China gambling is illegal, however, if we use vouchers to place bets or rewards in the form of skins, then this will become a legal behavior. In addition, we should also consider minors. Since minors are loyal users of many games and have relatively weak self-control capabilities, there is also a great possibility of participating in gaming. At this time, it is necessary to set up participants by law Minimum age to protect minors.

Another big problem relative to betting industry is match fixing, this behavior is similar to kick-fixing in a football game, and it usually results from several situations: Players are bribed by gambling companies; Players or teams place a lot of bets on the opponent team to win; Due to the rules of the game, in order to avoid confrontation with some stronger teams, and adopted a strategic failure.

Usually the first two will bring huge negative effects, and the people involved in the incident will also be punished. Although the latter is a strategic failure, it is contrary to the spirit of e-sports promotion and will also cause a certain degree of negative impact.

There are many examples of the first two cases:

In 2020, team Newbee match fixing in the dota2 game, resulting in five players being suspended for life, and team Newbee was removed from the Chinese DOTA2 Professional Association

In 2015, in the League of Legends IG VS LGD match, the winners of the two participating teams will be against the strong team EDG in the next round. Both teams want to avoid the

strong team by losing strategically. Although the two teams did not receive a penalty, the rules of the game have also been changed due to the negative matches of the two teams.

In order to prevent similar situations from happening again, for the first two situations, the relevant departments should strengthen supervision and increase penalties. Use the law to deter players and teams to reduce their chances of taking risks.

For the third case, the game organizer should set more reasonable competition rules and impose certain penalties on similar behaviors.

### III . METHODOLOGY

#### A . Research preparation

As we all know, the player is the core of e-sports, so the player's game experience has naturally become the core of E-sports innovation. In the complex MOBA game environment, taking DOTA 2 as an example, the possibility that every five heroes can form different formations is 140,364, 532, therefore, each participant's decision-making will have an impact on the final victory. Before the popularity of AI technology, the player's judgment of the game's victory or defeat was often based on the past game experience, but in reality, with the application of data mining system, players will have a complete understanding of the trend of the game.

This paper selects DOTA2 as a representative sample of E-sports prediction analysis. DOTA2 is a game owned by valve, which has the open source data of dota2 competition and steam. These data sources are reliable and accurate because they are retrieved from steam's official website. The open data website provides DOTA2 data API, and the data is relatively complete, which meets the needs of modeling and analysis. This experiment is to establish a model to predict the winning rate through the selection of heroes of both sides. Therefore, the following data are needed to be obtained, [Radiant hero list], [Dire hero list], [which party wins]. We have started to estimate the quality of a series of machine learning algorithms to predict the outcome of a given hero draft for each team. Under the condition of no ties, then  $p(\text{Radiant winning rate}) = 1 - p(\text{Dire winning rate})$ .

#### B . Dataset

In order to ensure the validity and value of the data, we will adopt the following restrictions:

1) MMR > 4000. MMR is a score standard to measure players' skill level. It has a set of unique algorithms. Players win + 30 and fail - 30 points when playing a rankable game. MMR above 4000 points is in line with the game level of most players.

2)The game time is more than 15 minutes. This condition is to rule out the abnormal ending of the game, mainly due to the problem of network disconnection or giving up the game due to negative emotions [27].

3) Only the rankable game data is used. Because players in the sky ladder competition will be more carefully scrutinizing the lineup and will not be like the ordinary game those random selection of heroes.

$max\_match\_ID$ , that is, search the match data before the game, and the other two variables  $target\_match\_Num$  and

*lowest\_MMR* respectively represents the number of match data to be recorded and the lowest matching score. The workflow is shown in the figure below, in which the blue box indicates the condition:

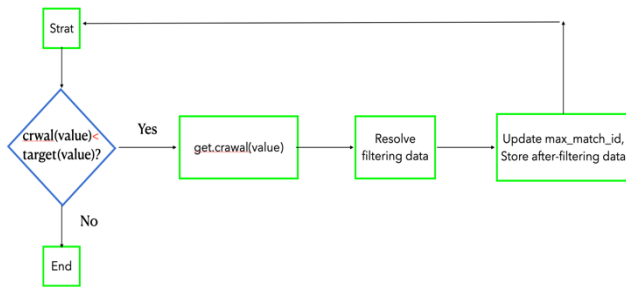


Figure 2: Data workflow schema

We crawled 100000 game data through the python crawler. The 100000 pieces of data include all the high segment match data from 2:00 on October 16 to 13:00 on October 29. Each piece of data has five attributes, they are Game ID, date, Radiant hero, Dire hero, Radiant victory (type Boolean). A training (test) sample is divided into two parts: input and output.

The input part is composed of a two-dimensional matrix with the shape of  $[2*129]$ , where 2 represents two vectors of Radiant and Dire, each vector has 129 bits. When there is a hero in Radiant, the corresponding position of the hero ID is set to 1, and the rest positions are 0. Therefore, the input of a sample is a two-dimensional matrix composed of two vectors. (the value range of hero ID is 1~129, but actually there are only 117 heroes. Some values do not correspond to any hero. In order to facilitate sample production, the vector length is set to 129.) The output part is a scalar of shaping, which represents whether Radiant square wins or not. We use 1 instead of true in the data file and 0 for false.

Thus, the final input shape of 100000 samples is  $[100000*2*129]$ , and the output shape is  $[100000*1]$ . After that, we divide the samples, according to the ratio of 8:1:1, 80000 samples are used as the training set, the training set is used for training the model, to help the model adjust the parameters and let the model adapt to these data. 10000 samples as the validation set. The function of validation set is to observe the effect of the model on the validation set after training a round on the training set. If the prediction accuracy of the model on the validation set does not improve, the training will be stopped to prevent the model from over fitting the training set. The rest 10000 samples is test data, test data will be used after validation data, mainly used to evaluate the generalization ability of the final model.

### C. Model

The mechanism of DOTA2 game is complex, the types of heroes are various, and the playing methods are rich and changeable. According to the existing literature research, scholars who solve this kind of prediction problem mostly use naive Bayes, random forest, support vector machine and other methods to make a lot of attempts. However, due to the complex relationship between heroes' collocation and restraint,

Therefore, it is difficult to obtain good prediction results. For example, Conley & Perry pointed out in his study that although logistic was used The combination of regression and KNN can achieve 69.8% accuracy rate in training set, but this method can not only be used under harsh conditions, but also very slow: for the case of  $k = 5$ , cross validation takes 12 hours [25], which is obviously unrealistic for players who need to predict the results in real time in the game. Therefore, we decided to break through the mindset and adopt Depth Learning algorithm to improve the existing prediction level.

Considering that the input of a sample is a two-dimensional matrix composed of two sparse vectors, we build three models: CNN model, LSTM model and CNN + LSTM model [21].

#### a. CNN model

The mainstream of neural network used in pattern recognition is directed learning network, and unsupervised learning network is more used for clustering analysis. For guided pattern recognition, because the category of any sample is known, the spatial distribution of samples is no longer based on its natural distribution tendency. Instead, it is necessary to find an appropriate spatial division method according to the spatial distribution of similar samples and the degree of separation between different types of samples, or to find a classification boundary, so that different types of samples are located in different regions Domain. This requires a long and complex learning process, constantly adjust the position of the classification boundary used to divide the sample space, so that as few samples as possible are divided into non similar regions.

Convolution network is essentially a kind of input-output mapping. It can learn a large number of mapping relations between input and output, without any precise mathematical expression between input and output. As long as the convolution network is trained with known patterns, the network has the mapping ability between input and output pairs. Convolution network performs supervised training, so its sample set is composed of vector pairs of the form: (input vector, ideal output vector).

Because the feature detection layer of CNN learns from the training data, the explicit feature extraction is avoided when using CNN, and the neural network can learn from the training data implicitly. Moreover, because the weights of the neurons on the same feature mapping surface are similar, the network can learn in parallel, which is a major advantage of convolution network compared with the neural network connected with each other.

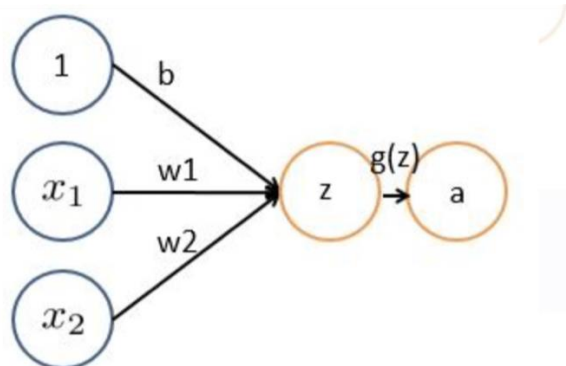


Figure 3: Neuron unit model

In the basic neurons of  $wx + b$ ,  $x_1$  and  $x_2$  represent inputs,  $w_1$  and  $w_2$  are weights. Each input is given a weight,  $b$  is bias,  $g(z)$  is activation function, and  $a$  is output. Here we use the sigmoid function, the expression of sigmoid is as follows:

$$g(z) = \frac{1}{1 + e^{-z}}$$

Where  $z$  is a linear combination, for example,  $z$  can be equal to:  $z = b + x_1 * w_1 + x_2 * w_2$ . By substituting a large positive number or a small negative number into the  $g(z)$  function, we can see that the result tends to 0 or 1.

Therefore, the graphical representation of the sigmoid function  $g(z)$  is as follows (the horizontal axis represents the definition domain  $Z$ , and the vertical axis represents the range  $g(z)$ ):

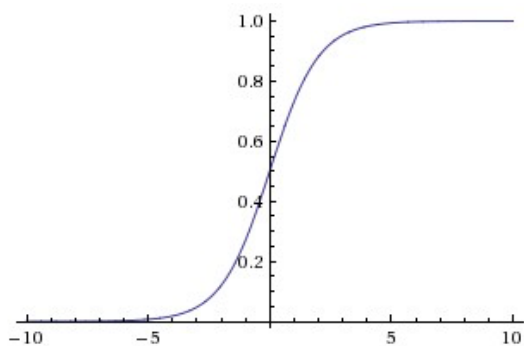


Figure 4: Sigmoid function image

In other words, the function of sigmoid function is to compress a real number from 0 to 1. When  $z$  is a very large positive number,  $g(z)$  will approach 1, while  $Z$  is a very small negative number, then  $g(z)$  will approach 0. In other words, if the sum of  $g(z) \rightarrow 0$  and  $g(z)$  is positive, then the sample is negative.

After the basic function model is selected, the input is convoluted by one-dimensional convolution check whose dimension is 3. After pooling and two full link operations, the dimension is changed to  $[1*1]$ . Finally, the sigmoid activation function is used to limit the output between  $[0,1]$ , that is, the winning probability of the corresponding sample. The following figure shows the dimension changes of matrix and vector.

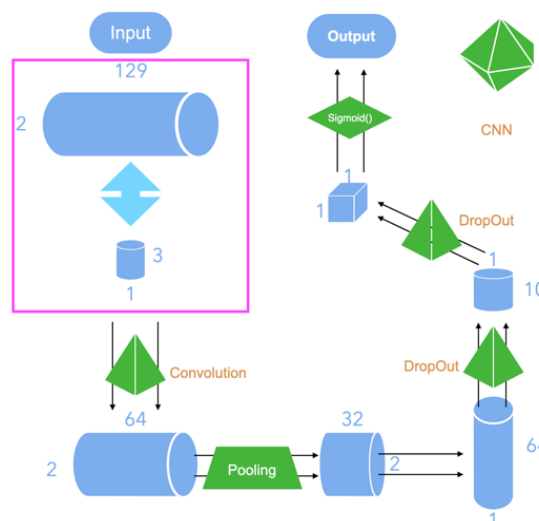


Figure 3: CNN model of winner prediction

Considering that the two-dimensional convolution will cross the vector, that is, the heroes of Radiant and Dire are convoluted together, which may not be helpful to the prediction results. Here, Conv1d is used to convolute the input in one dimension. After convolution, the matrix with dimension  $[2,64]$  is obtained, and then the matching maxpooling1d() function is used to add pooling layer. Next, use the reshape() function to adjust it to a one-dimensional vector, plus two dropout and density layers to convert the output to a scalar.

```

1 model = Sequential()
2
3 model.add(Conv1D(cnn_output_dim,kernel_size,padding='same',activation='relu',input_shape=(team_num,hero_id_max)))
4 model.add(MaxPooling1D(pool_size=pool_size,data_format='channels_first'))
5
6 model.add(Reshape((int(team_num*cnn_output_dim/pool_size),), input_shape=(team_num,int(cnn_output_dim/pool_size))))
7 model.add(Dropout(0.2))
8 model.add(Dense(10),input_shape=(team_num,cnn_output_dim/pool_size))
9 model.add(Dropout(0.2))
10 model.add(Dense(1))
11 model.add(Activation('sigmoid'))
12 model.compile(loss='mse',optimizer='adam',metrics=['accuracy'])
    
```

Figure 11: CNN model implementation code

In the actual parameter adjustment process, the convolution kernel length, convolution output vector dimension, dropout ratio and other parameters are not fixed and can be flexibly adjusted according to the training effect of the model.

*b. LSTM model*

Recurrent neural network (RNN) is a kind of neural network used to process sequence data. Compared with the general neural network, it can deal with the data of sequence change. For example, the performance of the same hero in different lineups is quite different. RNN can solve this problem well.



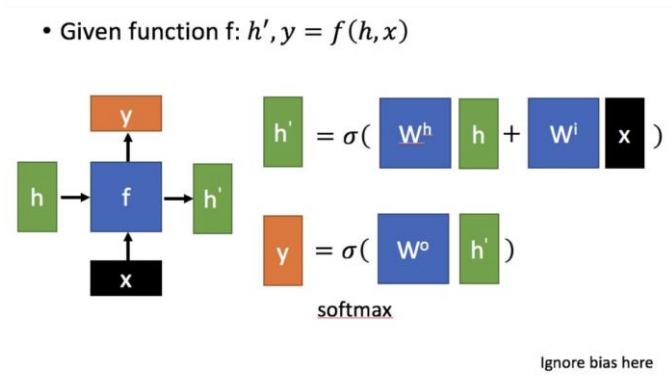


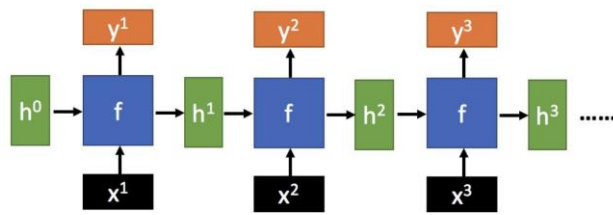
Figure 5: Naïve RNN model

Here,  $x$  is the input of data in the current state, and  $h$  is the input of the previous node received.  $y$  is the output in the current node state, and  $h'$  is the output passed to the next node. From the formula in the figure above, we can see that the output  $h'$  is related to the values of  $x$  and  $h$ .

However,  $y$  often uses  $h'$  to put into a linear layer (mainly for dimension mapping), and then uses softmax to classify and get the required data. How  $y$  is calculated by  $h'$  depends on the usage of the specific model.

Through the input of sequence form, we can get the following form of RNN.

• Given function  $f: h', y = f(h, x)$  h and h' are vectors with the same dimension



No matter how long the input/output sequence is, we only need one function  $f$

Figure 6 : Recurrent Neural Network model

Long short-term memory, LSTM is a kind of special RNN [28], which is mainly used to solve the problems of gradient disappearance and gradient explosion in the process of long-term training. Compared with ordinary RNN, LSTM can perform better in longer DOTA2 data sequences.

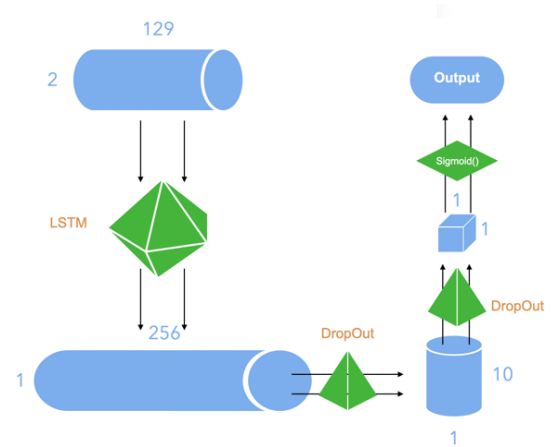


Figure 4: LSTM prediction model

Below is the code for building the LSTM model, where hidden\_ The size parameter is the length of the output eigenvector, and it is also an adjustable variable.

```

1 model = Sequential()
2 model.add(LSTM(hidden_size, input_shape=(team_num, hero_id_max),
3 return_sequences=False))
3 model.add(Dropout(0.2))
4 model.add(Dense(10))
5 model.add(Dropout(0.2))
6 model.add(Dense(1))
7 model.add(Activation('sigmoid'))
8 model.compile(loss='mse', optimizer='adam', metrics=['accuracy'])
    
```

Figure 7 : LSTM implementation code

*c. CNN+LSTM model*

The schematic diagram of the model construction is as follows. It is very similar to the CNN model. The only difference is that the reshape operation is replaced by the LSTM layer to generate a feature vector with a length of 256. The schematic diagram of the model construction is as figure 5.

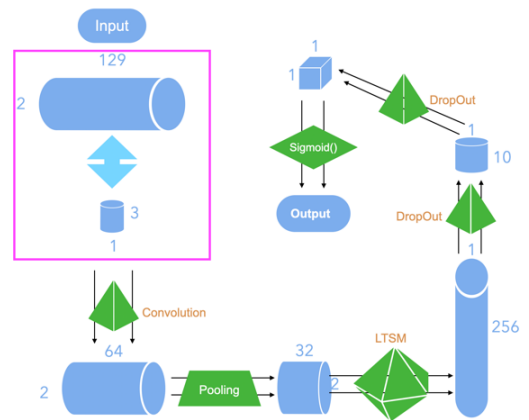


Figure 8: CNN+LSTM prediction model

The code of CNN + LSTM model is as follows, which is similar to the former two models:

```

1 model = Sequential()
2 model.add(Conv2D(cnn_output_dim, kernel_size, padding='same', input_shape=
(team_num, hero_id_max))
3 model.add(MaxPooling2D(pool_size=pool_size, data_format='channels_first'))

4 model.add(LSTM(hidden_size, input_shape=(team_num, cnn_output_dim/pool_size),
return_sequences=False))
5 model.add(Dropout(0.2))
6 model.add(Dense(10))
7 model.add(Dropout(0.2))
8 model.add(Dense(1))
9 model.add(Activation('sigmoid'))
10 model.compile(loss='mse', optimizer='adam', metrics=['accuracy'])
    
```

Figure 9: LSTM+CNN model implementation code

d. Callback functions

The callback function is to check the effect of the model on the verification set after each round of training. If the prediction effect of the model verification set is worse than that of the previous round after this round of training, the callback function can adjust the learning rate or stop the training.

```

1 callbacks = [keras.callbacks.EarlyStopping(monitor='val_loss', patience=2,
verbose=0, mode='min'),\
2   keras.callbacks.ReduceLROnPlateau(monitor='val_loss', factor=0.5, patience=1,
verbose=0, mode='min',\
3   epsilon=0.0001, cooldown=0, min_lr=0)]
4 hist = model.fit(tx, ty, batch_size=batch_size, epochs=epochs, shuffle=True, \
5   validation_data=(validate_x, validate_y), callbacks=callbacks)
6 model.save(model_saved_path+model_name+'.h5')
    
```

Figure 10: Callback function implementation code

First, set the callback function, where monitor='val\_loss' is to monitor the loss of the validation set, if the loss does not get smaller after one round of training, then the callback will be performed; patience is to wait for the rounds, in the above code, if one round of training 'val\_loss' does not get smaller, then the learning rate will be adjusted (ReduceLROnPlateau), if two rounds of training 'val\_loss' does not get smaller, then the training will be terminated (EarlyStopping). Finally, we use the model.fit() function to start training. tx, ty is the input and output of the training set, in the validation\_data parameter we need to pass our validation set, and in the callbacks parameter, we need to pass our set callback function.

IV . RESULTS

A. Experimental results

Model will predict the winning rate of each team, if the predict winning rate is over 50%, and the final winner team is the same as model predicted, then it is a correct prediction. The accuracy is equal to the number of correct predictions divide by the number of all predictions.

We will train the model to predict the test set, after this day of repeated tuning, we got several models with good prediction effect, the highest model prediction accuracy can reach 58%. In other words, for the task of "Guessing the winner by looking at the lineup", the model can reach 58% accuracy. In order to have a more comprehensive understanding of the model's predictive effect, we calculate the predictive accuracy of the model in the test set, training set, and validation set, and calculate the predictive accuracy of the model in the case of higher scores. Take the following model as an example:

```

Accuracy of test set: 0.5779
Accuracy of training set: 0.6196375
Accuracy of validation set: 0.5807
The accuracy of test set prediction success rate above 0.6: 0.6300398759614131 (2844/4514)
The accuracy of test set prediction success rate is above 0.75: 0.7433862439862434 (281/378)
The accuracy of test set prediction success rate is above 0.8: 0.7422680412371134 (72/97)
The accuracy of test set prediction success rate is above 0.85: 0.875 (7/8)
The accuracy of test set prediction success rate is above 0.9: 0.0 (0/0)
The accuracy of test set prediction success rate is above 0.95: 0.0 (0/0)
    
```

Figure 11: Simple slice of prediction results

It can be seen that the prediction effect of the model on the training set is slightly better than 61%, while the prediction accuracy on the training set and verification set is around 58%, and there is no obvious over fitting phenomenon.

In addition, for 10000 samples in the test, 4514 were judged to have a winning rate of more than 60% by the model, of which 2844 were correctly predicted, with an accuracy rate of 63%. 378 matches were judged by the model as having more than 75% of the winning rate, of which 281 matches were correctly predicted with an accuracy rate of 74.3%.

97 matches were judged by the model as having more than 80% of the winning rate (the winning rate was determined as 80% when the lineup selection was completed), of which 72 matches were correctly predicted with an accuracy rate of 74.2%.

There were also eight games identified by the model as having close to a 90% chance of winning, and seven were predicted correctly with an accuracy of 87.5%.

Therefore, we can draw a conclusion that the prediction results given by the model have certain reference value.

In order to have some intuitive feeling of the prediction effect, we focus on the lineup with the predicted winning rate greater than 0.85.

```

radiant: ['Drow Ranger', 'Pudge', 'Vengeful Spirit', 'Clinkz', 'Invoker']
dire: ['Lina', 'Witch Doctor', 'Enchantress', 'Lone Druid', 'Abaddon'] prediction: 0.86839567 reality: 0.0

radiant: ['Pudge', 'Sand King', 'Slark', 'Zeus', 'Abaddon']
dire: ['Tinker', 'Batrider', 'Spectre', 'Rubick', 'Bristleback'] prediction: 0.8681598 reality: 1.0

radiant: ['Axe', 'Mirana', 'Templar Assassin', 'Keeper of the Light', 'Legion Commander']
dire: ['Nature's Prophet', 'Ursa', 'Rubick', 'Medusa', 'Bristleback'] prediction: 0.8787791 reality: 1.0

radiant: ['Axe', 'Mirana', 'Meepo', 'Slark', 'Abaddon']
dire: ['EarthShaker', 'Lina', 'Shadow Demon', 'Troll Warlord', 'Ember spirit'] prediction: 0.85831786 reality: 1.0

radiant: ['Pudge', 'Zeus', 'Faceless Void', 'Broodmother', 'Spirit Breaker']
dire: ['Lina', 'Sniper', 'Bristleback', 'Legion Commander', 'Techies'] prediction: 0.86522377 reality: 1.0

radiant: ['Pudge', 'Zeus', 'Night Stalker', 'Silencer', 'Slark']
dire: ['EarthShaker', 'Lina', 'Lifestealer', 'Bristleback'] prediction: 0.85895973 reality: 1.0

radiant: ['Bane', 'Pudge', 'Viper', 'Lifestealer', 'OmniKnight']
dire: ['Tiny', 'Kunkka', 'Lina', 'Faceless Void', 'Centaur Warrunner'] prediction: 0.87895906 reality: 1.0

radiant: ['Queen of Pain', 'Dazzle', 'Night Stalker', 'Silencer', 'Naga Siren']
dire: ['Faceless Void', 'Outworld Devourer', 'Brewmaster', 'Rubick', 'Grimsstroke'] prediction: 0.8581611 reality: 1.0
    
```

Figure 12: Eight-game prediction results

The model predicted the winning rate of Radiant, ranging from 85% to 87%. If ordinary players look at the lineup and predict the final outcome, it is obvious that the success rate of human prediction is much lower than that of the model. At the same time, in these 8 games, the hero "Pudge" played a lot of times, up to 5 times.

Therefore, we have doubts whether this hero is more powerful in the current game version, so that the chosen party will get a greater chance of winning. so, we query Pudge's hero restraint index (the hero restraint index reflects the hero in the situation of occupying a natural advantage or disadvantage, where the positive percentage is regular is the advantage, otherwise it is the disadvantage) on DOTA 2 data mining website "MAX +".

From the MAX+ figure, Tinker, Techies, Specter, Sniper, Enchantress and other heroes all appeared as enemies of Pudge. This also shows that the prediction results of our model are

consistent with the conclusions displayed on the statistical level. therefore, we believe that the CNN + LSTM model based on keras framework has the best fitting degree.

### *B. Prospect of application scenarios*

Scenario 1: Both players have completed the Ban & Pick session.

DOTA 2 has a 2-minute hero picking phase before the game starts, during this time each player has to strategize on BAN and PICK from the point of view of hero coordination or team, once the player finishes picking, they can input their lineups into the model for prediction, If the model gives a predicted win rate of 20% or 10%, the player may consider leaving the game early before entering the game, thus avoiding wasting time and preserving a good competitive position.

Scenario 2: The enemy has completed the Ban & Pick session, and we have only one hero left to choose.

The DOTA2 system provides each player with 30 seconds to pick, at which point the player can enter pre-picked heroes into the model to make predictions, thus selecting the hero with the highest odds of winning given the model.

Scenario 3: Using prediction model for E-sport bets, place bets based on predictions results.

In regular rankable games, the teammates that players match are all strangers, it is difficult to have a high degree of tacit understanding, so personal ability will have a greater impact. However, pro-team pays more attention to the ability of collaboration within the team. So there are large differences between regular rankable games and pro-team matches, and the training data for our model comes from regular players in rankable games, the prediction is more risky on a practical level, but this idea is still valuable and informative if we train the model using a large amount of data from pro-team matches.

### *C. The limitations of the model*

1) The CNN + LSTM model based on keras framework is not mature enough. A lot of computing resources are needed in training, and it will take about 1 minute to use the model, which makes it very inconvenient for players to manipulate.

2) Although lineup selection is a very important factor in MOBA games, it is not a decisive factor. The winning or losing trend of the game is not only affected by the objective factors of hero skill collocation and restraint relationship between heroes, but also by subjective factors such as the players' proficiency in using heroes, the players' selection of heroic equipment, the mentality of players in the game, and the communication between players and their teammates. Therefore, the winning rate given by the model is only a reference for players to make decisions Even if 90% of the winning rate is given by the model, the player is still more likely to lose the game due to these factors. On the contrary, even if the model is given a 10% winning rate, the player still has the hope of winning the game through the correct strategy operation in the game.

3) Due to the defect of training set data, it may lead to over fitting. The training samples are not complete enough to include all the match possibilities. After calculation, it is known that if the 117 heroes in dota2 are randomly arranged and combined as

a group, there are  $2 \times 10^{16}$  possible lines without repetition. Therefore, we only study the possibility of  $5 \times 10^{-12}$ , which will have a great impact on the accuracy of the model and produce the phenomenon of over fitting.

4) DOTA2 There is a huge problem in the internal balance of the game, which will produce a lot of interference for the model prediction. Among the 100000 game records in this data set, there are 54814 radiant victories and 45186 dire victories. If we compare the victory of radial and dire to the probability of positive and negative coin toss, the probability of difference between them reaches 9%, indicating that the current version of the map is flat There are huge problems in balance.

## V . CONCLUSION

This paper present 4 aspects about e-sports, discussing some AI innovations in E-sports. And deeply explain the prediction system of MOBA.

Firstly, we talk about the DDA in different types of game and the principle of DDA in different games. This makes the game more competitive and interesting, allowing players to stay in the flow channel and making the game more attractive to players. In a multiplayer battle game, dynamic difficulty adjustment is accomplished through a matching mechanism, using elo and mmr to achieve this, which allows players to always play against players of similar level.

Secondly, we mention the AI team and the way that Open AI build the AI team, AI team once became the best team in the world. Even the professional teams can't beat them. What's more, we put forward some benefits that AI team can bring us.

Thirdly, we discuss the data mining system in MOBA games, which helps us know the fact behind the data.

Lastly, we focus on the AI in e-sports betting industry. AI is used in the prediction of winning rate in competitions and calculating the odds for each competition. And the we should not ignore many of the problems arising from the e-sports betting industry. We should formulate reasonable laws and regulations to protect minors. The relevant departments should strengthen supervision and strengthen penalties to prevent counterfeiting events.

We discussed in depth the use of ai in predicting the results of the competition. After the preparation phase, we first select the appropriate data by screening the data, and then through the three models of CNN, LSMT, CNN+LSMT, we can get higher The accuracy rate of the match result prediction rate. Then we discussed three scenarios that can be used, namely:

Both players have completed the Ban & Pick session.

The enemy has completed the Ban & Pick session, and we have only one hero left to choose

Using prediction model for E-sport bets, place bets based on predictions results.

Finally, we discussed the limitations of the model.

References

- [1] Xia Yuhui. 2018. "Analysis on the Influencing Factors of the Standardization Development of E-Sports in China." *Advances in Social Sciences* 07(02):118–22.
- [2] Newzoo 2020, Global Esports Market Report: January 2020. Retrieved from [http://www.\(URL\)](http://www.(URL))
- [3] Graham, B.A (2017, April). eSports to be a medal event at 2022 Asian Games. *The Guardian*. Retrieved online from [http://www.\(URL\)](http://www.(URL))
- [4] Qi, Shen, Chen Chun-Chih, and Wu Shu-Ming. 2019. "Research of the Attractiveness Factors of MOBA Mobile Games Based on Evaluation Grid Method." *Journal of Physics: Conference Series* 1325:012004.
- [5] Wu, Bin. 2019. "Hierarchical Macro Strategy Model for MOBA Game AI." *Proceedings of the AAAI Conference on Artificial Intelligence* 33(01):1206–13.
- [6] Byford, Sam. 2017. "Google's AlphaGo AI Defeats World Go Number One Ke Jie." *The Verge*.
- [7] LIU Wei, WANG Mu-xuan. Rethinking on the Command and Control System[J]. *Science and Society*, 2016, 6(3): 86-103
- [8] Bhattacharya, Rohit, and Azwad Sabik. n.d. 2016. "Data-Driven Recommendation Systems for Multiplayer Online Battle Arenas." 8.
- [9] Super Data. 2018. Worldwide digital games market: February 2018. Retrieved from [http://www.\(URL\)](http://www.(URL))
- [10] Qi, Shen, Chen Chun-Chih, and Wu Shu-Ming. 2019. "Research of the Attractiveness Factors of MOBA Mobile Games Based on Evaluation Grid Method." *Journal of Physics: Conference Series* 1325:012004.
- [11] Mora-Cantallops, Marçal, and Miguel-Ángel Sicilia. 2018. "MOBA Games: A Literature Review." *Entertainment Computing* 26:128–38.
- [12] Johnson, D., Nacke, L. E., Wyeth, P., 2015. All about that base: differing player experiences in video game genres and the unique case of moba games. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, pp. 2265–2274.
- [13] Kwak, H., Blackburn, J., Han, S., 2015. Exploring cyberbullying and other toxic behavior in team competition online games. *Social Dynamics* 22 (28), 47.
- [14] Andrade, G., Santana, H., & Jussieu, B. P. (2005). Challenge-Sensitive Action Selection: an Application to Game Balancing Vincent Corruble. In *Proceedings of the 2005 IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT'05)*.
- [15] Baldwin, Alexander. 2017. "Balancing Act: The Effect of Dynamic Difficulty Adjustment in Competitive Multiplayer Video Games." Phd, Queensland University of Technology.
- [16] Raiman Jonathan, Zhang Susan, and Wolski Filip. 2019. "Long-Term Planning and Situational Awareness in OpenAI Five."
- [17] White, Alfonso, and Daniela M. Romano. 2020. "Scalable Psychological Momentum Forecasting in Esports." *ArXiv:2001.11274 [Cs]*.
- [18] OpenAI, Christopher Berner, Greg Brockman, Brooke Chan, Vicki Cheung, Przemysław Dębiak, Christy Dennison, David Farhi, Quirin Fischer, Shariq Hashme, Chris Hesse, Rafal Józefowicz, Scott Gray, Catherine Olsson, Jakub Pachocki, Michael Petrov, Henrique Pondé de Oliveira Pinto, Jonathan Raiman, Tim Salimans, Jeremy Schlatter, Jonas Schneider, Szymon Sidor, Ilya Sutskever, Jie Tang, Filip Wolski, and Susan Zhang. 2019b. "Dota 2 with Large Scale Deep Reinforcement Learning." *ArXiv:1912.06680 [Cs, Stat]*.
- [19] Wang, Tian. n.d. 2018. "PREDICTIVE ANALYSIS ON ESPORTS GAMES – A CASE STUDY ON LEAGUE OF LEGENDS (LOL) ESPORTS TOURNAMENTS." 46
- [20] Bansal, Gagan, Besmira Nushi, Ece Kamar, Walter Lasecki, Daniel S. Weld, and Eric Horvitz. n.d. 2019. "Beyond Accuracy: The Role of Mental Models in Human-AI Team Performance." 10.
- [21] Hernandez, Danny, and Tom B. Brown. 2020. "Measuring the Algorithmic Efficiency of Neural Networks." *ArXiv:2005.04305 [Cs, Stat]*.
- [22] Araujo, Vladimir, Felipe Rios, and Denis Parra. 2019. "Data Mining for Item Recommendation in MOBA Games." Pp. 393–97 in *Proceedings of the 13th ACM Conference on Recommender Systems*. Copenhagen Denmark: ACM.
- [23] Barman, Nabajeet, Saman Zadtootaghaj, Steven Schmidt, Maria G. Martini, and Sebastian Moller. 2018. "GamingVideoSET: A Dataset for Gaming Video Streaming Applications." Pp. 1–6 in 2018 16th Annual Workshop on Network and Systems Support for Games (NetGames). Amsterdam: IEEE.
- [24] Yang, Pu, Brent Harrison, and David L. Roberts. n.d. 2016. "Identifying Patterns in Combat That Are Predictive of Success in MOBA Games." 8.
- [25] Semenov, Aleksandr, Peter Romov, Sergey Korolev, Daniil Yashkov, and Kirill Neklyudov. 2017. "Performance of Machine Learning Algorithms in Predicting Game Outcome from Drafts in Dota 2." Pp. 26–37 in *Analysis of Images, Social Networks and Texts*. Vol. 661, Communications in Computer and Information Science, edited by D. I. Ignatov, M. Yu. Khachay, V. G. Labunets, N. Loukachevitch, S. I. Nikolenko, A. Panchenko, A. V. Savchenko, and K. Vorontsov. Cham: Springer International Publishing.
- [26] Li, Quan, Peng Xu, Yeuk Yin Chan, Yun Wang, Zhipeng Wang, Huamin Qu, and Xiaojuan Ma. 2017. "A Visual Analytics Approach for Understanding Reasons behind Snowballing and Comeback in MOBA Games." *IEEE Transactions on Visualization and Computer Graphics* 23(1):211–20.
- [27] Shim, Jae Youn, Tae Hyun Kim, and Seong Whan Kim. 2014. "Decision Support of Bad Player Identification in MOBA Games Using PageRank Based Evidence Accumulation and Normal Distribution Based Confidence Interval." *International Journal of Multimedia and Ubiquitous Engineering* 9(8):13–24.
- [28] Chung, Junyoung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. 2014. "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling." *ArXiv:1412.3555 [Cs]*.
- [29] Yiqing Chen, Jie Mei and Steward Huang, August 2020. "Research on innovation of E-sports industry in the area of Artificial Intelligence" Pp.4-56 Master Thesis of AIBT program at SKEMA Business School Paris.

**Creative Commons Attribution License 4.0  
(Attribution 4.0 International, CC BY 4.0)**

This article is published under the terms of the Creative Commons Attribution License 4.0

[https://creativecommons.org/licenses/by/4.0/deed.en\\_US](https://creativecommons.org/licenses/by/4.0/deed.en_US)