ARTICLE IN PRESS

Digital Communications and Networks xxx (xxxx) xxx

Contents lists available at ScienceDirect

Digital Communications and Networks

journal homepage: www.keaipublishing.com/dcan

Sports match prediction model for training and exercise using attention-based LSTM network

Qiyun Zhang [a], Xuyun Zhang [b], Hongsheng Hu [c], Caizhong Li [a], Yinping Lin [d], Rui Ma [e,*]

[a] Weifang Key Laboratory of Blockchain on Agricultural Vegetables, Weifang University of Science and Technology, Shouguang, 262700, China
[b] Department of Computing, Macquarie University, Sydney, Australia
[c] Faculty of Engineering, University of Auckland, New Zealand
[d] No.1 Middle School of Pidu Chengdu, Chengdu, 611730, China
[e] General Education Department, Shandong First Medical University (Shandong Academy of Medical Sciences), Taian, China

ARTICLE INFO

Keywords:
Sports
Prediction
Long short-term memory
Attention
Sliding window

ABSTRACT

Sports matches are very popular all over the world. The prediction of a sports match is helpful to grasp the team's state in time and adjust the strategy in the process of the match. It's a challenging effort to predict the sports match. Therefore, a method is proposed to predict the result of the next match by using teams' historical match data. We combined the long short-term memory (LSTM) model with the attention mechanism and put forward an AS-LSTM model for predicting match results. Furthermore, to ensure the timeliness of the prediction, we add the time sliding window to make the prediction have better timeliness. Taking the football match as an example, we carried out a case study and proposed the feasibility of this method.

1. Introduction

At present, the world economy is in a stage of rapid development, and people's living standards are constantly improving. People pay more attention to material life. With the popularity of the Internet, people can watch various games on TV or live. Among them, football is the largest single event with a huge impact [1]. In Western countries with developed market economies, football has become an important sector of the national economy, even a pillar industry. And in Italy, Britain and other countries, the football industry has become their pillar industry. The football industry can provide tens of thousands of employment opportunities and promote the development of the whole national economy. Therefore, in today's fast-growing business, football and the economy are more and more closely linked. There are many business opportunities in the football industry. There are various types of football leagues, such as La Liga, Premier League, Bundesliga, Serie A, etc. There are also continent-wide competitions, such as the UEFA Champions League and the AFC Champions League. There is also the FIFA World Cup referred to as the “World Cup”. It is a football match that symbolizes the highest honor and has the greatest popularity and influence. Global TV viewers exceed 3.5 billion. At the same time, the tourism economy brought about by football matches is impressive. Winning or losing in sports games often arouses people's great curiosity and promotes the development of gambling. So the prediction of football match results has become a research hotspot.

Prediction technology is also constantly developing and has been applied to many fields, such as point-of-interest prediction [2], stock price prediction [3] and service dynamic planning [4,5]. There are various prediction methods, mainly including linear prediction and nonlinear prediction. Among them, the prediction of sports performance should adopt nonlinear prediction because there are many factors affecting sports competitions. Taking football games as an example, the stadium will affect the mentality of the players, and a home match is more conducive to the players' performance. The players’ physical and psychological quality, the game tactics, weather factors and even some unexpected accidental factors will also affect the result of the game. Therefore, the linear prediction method cannot consider these factors, but the nonlinear method can.

The purpose of our research is to predict the results of sports matches. Taking the prediction of football match results as an example, the result of a football match is caused by various factors, e.g., time, environment. Therefore, capturing the potential characteristics of teams and predicting the performance of teams is a very worthy research topic. In order to predict the results of the game in advance and make the deployment for

* Corresponding author.
E-mail addresses: zhangqiyun@wfust.edu.cn (Q. Zhang), xuyun.zhang@mq.edu.au (X. Zhang), hhu603@aucklanduni.ac.nz (H. Hu), licaizhong@wfust.edu.cn (C. Li), 415077676@qq.com (Y. Lin), 49803937@qq.com (R. Ma).

https://doi.org/10.1016/j.dcan.2021.08.008
Received 23 January 2021; Received in revised form 19 August 2021; Accepted 24 August 2021
Available online xxxx

the game in advance, we choose to use the LSTM model. First, we collect a large number of football match data, then input the football match related information into the LSTM model according to the time sequence for training, and use the trained LSTM model to predict the game results. It also provides available models for other competitions. LSTM can capture the characteristics of teams under time change. Also, we talk about the combination of the attention mechanism and the time sliding window to improve the timeliness of model prediction results. The approaches to our work are summarized below.

- In order to capture the team's potential state and predict the team's game results by using the team's existing data, we use the LSTM model to predict the game results.
- To capture the team's short-term state and better explore the team's potential characteristics, we combine LSTM with attention mechanism and time sliding window and propose an attention-based sports competition performance prediction model (AS-LSTM) to predict the sports competition results.
- We designed a case study to introduce the feasibility of our method.

The rest of this paper is organized as follows. Section 2 presents the related works. In Section 3, the preliminary knowledge is presented, including LSTM and attention mechanism. Section 4 details our sports match prediction model. Section 5 introduces our case study, and Section 6 concludes this paper and discusses our future work directions.

2. Related work

The neural network has shown good performance in predicting problems. There are many kinds of neural networks, and the most basic one is the Deep Neural Network (DNN). Rahman [6] proposed a football game prediction framework based on DNN, which shows that deep learning can successfully predict the results of football games or any other sports events. And Park et al. [7] used DNN to deal with sports news, mining news reports related to athletes in news articles and put forward a sports athlete evaluation model. Artificial Neural Network (ANN) is also used to predict sports performance. Montoye et al. [8] collected sports competition data first and then used their design ANN to compare their method with linear regression and linear composite. It proved that the precision of ANN is good.

Convolutional Neural Network (CNN) is one of the representative algorithms of deep learning. It has the ability to represent learning and multi-dimensional data processing. Johnson et al. [9] first used the tactical feature engineering technology to compress space-time and promote the fine-tuning of three pre-trained CNNs. They used the space-time driven CNN model to predict the ground reaction forces and moments of athletes. In addition, Wu et al. [10] proposed a method to identify the semantic time in basketball videos. They used the two-stream 3D CNN framework to predict the group activity recognition of the movement modes, and extracted the appearance features of the athletes through the CNN structure. To judge the success or failure of the team competition, the experiment proved that this method had the most advanced performance. Huang et al. [11] proposed an incremental action tube construction method for spatiotemporal context-aware online action detection and prediction. They used a convolutional recurrent neural network encoder-decoder model to predict the object's action in advance. Although the above methods can deal with multi-dimensional data, it is difficult to handle time-series data. However, the results of sports competitions often change over time. For example, the physical condition of the distance mobilization, the environment of the field and other factors are affected by the state of time.

The recurrent neural network is also one of the representative algorithms of deep learning. At present, recurrent neural networks are also used in various prediction models. Wozniak et al. [12] have used recurrent neural networks to predict the posture of the human body and evaluate the safety of the human body through the collected motion-sensing data. This method helps people at risk. In today's world, sports games provide a lot of statistical information for each player, team, game and season [13]. In other words, there is already a lot of available data on sports. Taking Rugby as an example, Watson et al. [14] considered the continuity of the game, the sequences of the players' actions and the location of the actions, used CNN and RNN to predict the results of the game, achieving good prediction accuracy. Taking football as an example, the football game is a group behavior with a noise background. Strnad et al. [15] used MultiLayer Perceptron (MLP) and RNN to predict the group behavior to judge the attendance of the football game. There are many scenarios where RNN is used for prediction. For example, Zhang et al. [16] provided a multi-task RNN with high-order Markov Random Fields (MRF) to predict the trend of stock price. The multi-task RNN framework can achieve high-quality forecasting results only by extracting information features from single stock original market data without other domain knowledge. In addition, Chen et al. [17] used RNN to mine users' interests on location-based social networks to predict the next point of interest. Oytun et al. used a backpropagation neural network and LSTM to predict the performance of female handball players [18].In addition to athletes' performance prediction, deep learning has also made great progress in other fields, such as image denoising [19] and download delay optimization in edge computing [20].

However, the recurrent neural network also has some disadvantages. For example, it can not capture the long-term dependence between historical data [21]. In the actual training process, some important historical information may be ignored. Therefore, people began to use LSTM, an excellent variant of RNN, to predict time series data. In order to help the basketball team develop tactics, Yu et al. [22] proposed a method combining deep Bi-LSTM with Mixture Density Network (MDN). Through the real-world basketball trajectory data training model, it helps coaches and players decide when and where to shoot so as to win the basketball game. Wei et al. [23] proposed an AE-LSTM model combined with an autoencoder to predict traffic flow, which is helpful to the development of urban planning.

Since many models are not perfect, some optimization strategies provide ideas for improving the model [24–26]. Attention mechanism has become an important concept in neural networks and has been fully studied in different fields [27,28]. After that, researchers began to combine the attention mechanism with the neural network for problem prediction. Wu et al. [29] proposed an attention-based CNN-LSTM-BiLSTM model for short-term load forecasting. Considering the inherent time relationship of speech waveforms, Xie et al. [30] used attention-based LSTM to extract emotional features in human languages and perform voice emotion classification. The application of the attention mechanism improves the performance of the model and the interpretability of the neural network model.

Although there are many prediction methods, it is still difficult to predict sports performance. Taking a football game as an example, the victory of the match is affected by many factors, for example, home-court advantage. A team has an advantage at a home court. In the game, the home team has more fans, and the athletes are more likely to be inspired. Moreover, the referee may have subjective bias. The visiting team is exhausted from the journey. They cannot get the support of many fans, and their physical condition will also be worse compared to the home team. Besides, the victory of a game is closely related to goals and shots of the team, e.g., shots provide opportunities for goals. The more goals, the greater the opportunity to win the game. In other words, the number of goals is related to the victory of the game. However, it does not necessarily mean the team can win with more shots or goals. The possession rate can also affect the result of the game. Even if a team scores a few goals, it can win the game through the players' ball control skills.

In this situation, how to predict the results of sports competitions is still a challenge. Therefore, we need to find a way to predict the result in advance so that the team can adjust their strategy in advance. And it can facilitate the coach to train the team. We proposed the attention-based LSTM model to predict the results of the game. To improve the timeliness of the prediction results, we will set a time sliding window to ensure the prediction of the next game result.

3. Preliminary knowledge

3.1. Notations and definitions

Definition 1. Football match grade: A team's football match grade is a 7-tuple $F^u_{t_k} = (u, g, l, c, s, h, t_k)$. It indicates that a team u plays a game in place h at time point t_k.

Definition 2. Football match history: A football match history is a set of football match grades of a team. It can be denoted as $FH_u = (u, g_1, l_1, c_1, s_1, h_1, t_1), (u, g_2, l_2, c_2, s_2, h_2, t_2), \ldots, (u, g_k, l_k, c_k, s_k, h_k, t_k)$, Thus, all teams football match histories can be denoted by $AFH = \{FH_{u_1}, FH_{u_2}, \ldots, FH_{u_{|U|}}\}$, where U is the number of all teams.

Definition 3. Football match prediction: Given a team's football match history FH_u, the goal is to predict the team's next possible game results, including results of the competition, the number of goals, ball control rate and the number of goals lost. This is helpful for the coach to find the team status in time and train the team. Table 1 shows the notations used in this paper.

3.2. Long short-term memory

To better mine the historical information of sports training, we choose to use the LSTM network. Though RNN can handle the time series data, the basic RNN is very prone to gradient disappearance and gradient explosion in the training process of the model, and it cannot well capture long-term dependence. LSTM is a good variant of RNN, and it is different from general RNN in updating the state method. The LSTM unit contains three gates, namely: forget gate, input gate and output gate. The structure of an LSTM unit is shown in Fig. 1.

The forget gate determines what information should be discarded or retained. Information from the previously hidden state $h_t - 1$ and from the current input x_t are passed through the sigmoid function. Then, the output value of the forget gate is between 0 and 1. If the value is close to 0, it means that the information will be forgotten. Otherwise, the closer to 1, the more information will be retained. And the calculation formula of forget gate is as follows:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{1}$$

where σ represents the sigmoid activation function, which limits the value between 0 and 1. h_{t-1} is the output of the previous cell, x_t is the input of the current cell, W_f is the weight of the forget gate, and b_f is the bias term. Then, we put the previously hidden state h_{t-1} and the current input x_t into the sigmoid function. This determines which values are updated by converting the values from 0 to 1. Among them, 0 means unimportant, and 1 means important. Then the hidden state h_{t-1} and current input x_t are passed to the tanh function. And the calculation formula of the input gate is as follows:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

where W_i is the weight of the forget gate, and b_i is the bias term. Now we calculate the cell state and update the new value to the cell state.

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{3}$$

$$C_t = f_t^* C_{t-1} + i_t^* \tilde{C}_t \tag{4}$$

where tanh represents the tanh activation function, W_C is the weight of the cell state, and b_C is the bias term.

Finally, the output gate determines what the next hidden state should be. Then the new cell state C_t and the new hidden state h_t are transferred to the next time step. And the calculation formula is as follows:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t^* \tanh(C_t) \tag{6}$$

where W_o represents the weight of the forget state, and b_o is the bias term. Therefore, we can use LSTM to analyze the performance of the sports team. Through appropriate forgetting and memory, we can predict the added results of the team, which will help the coach understand the team's current situation in time and help the team's training.

Table 1
Notations in this paper.

Symbol	Description
u	team
$U = \{u_1, u_2, \ldots, u_M\}$	collection of all teams
$F^u_{t_k}$	collection of all POIs
h_t	hidden vector of the LSTM unit
$P^u_{t_{N+1}, F_k}$	Match result probability of team u at t_{N+1}
f_t, i_t, o_t	forget gate vector, input gate vector and output gate vector of LSTM units
C_t	cell state
$\tilde{C}_t$	candidate state
g	goals
l	goals lost
c	the rate of ball control
s	results of the competition
h	home-court advantage
t_k	the time of the game
FH_u	set of football match grades
AFH	set of all teams football match histories
$P^u_{t_{N+1}, F_k}$	S-LSTM output
$o^u_{t_{N+1}, C_k}$	AS-LSTM output
σ	sigmoid function
$\{w\}$	set of weight matrices for a LSTM model
$\{b\}$	set of bias vectors for a LSTM model

3.3. Attention model

The attention mechanism is essentially derived from the human visual attention mechanism. When human vision perceives things, it generally does not see a scene from the beginning to the end but observes a specific part as needed [31]. The principle of the attention mechanism is similar, and it will focus on the influential part. In the historical training and exercise of a sports team, there are always some more important games and have a greater influence on the subsequent teams. And some sports or training may not attract too much attention and have little effect on subsequent performance prediction. The advantage of the attention mechanism is that it can focus on the influential parts. At present, attention mechanisms are divided into multiple types. Next, we will introduce the attention mechanism we use.

The essence of the attention function can be described as a mapping from a query to a series of key-value pairs, as shown in Fig. 2. By comparing the similarity between the key and the value, the attention score was obtained. The higher the score, the more similar they are. And as shown in Fig. 3, there are three steps in calculating attention. The first step is weight calculation. We need to calculate the similarity between the query and each key. There are many functions for calculating the similarity, such as dot product, splicing, perceptron, etc. And we choose

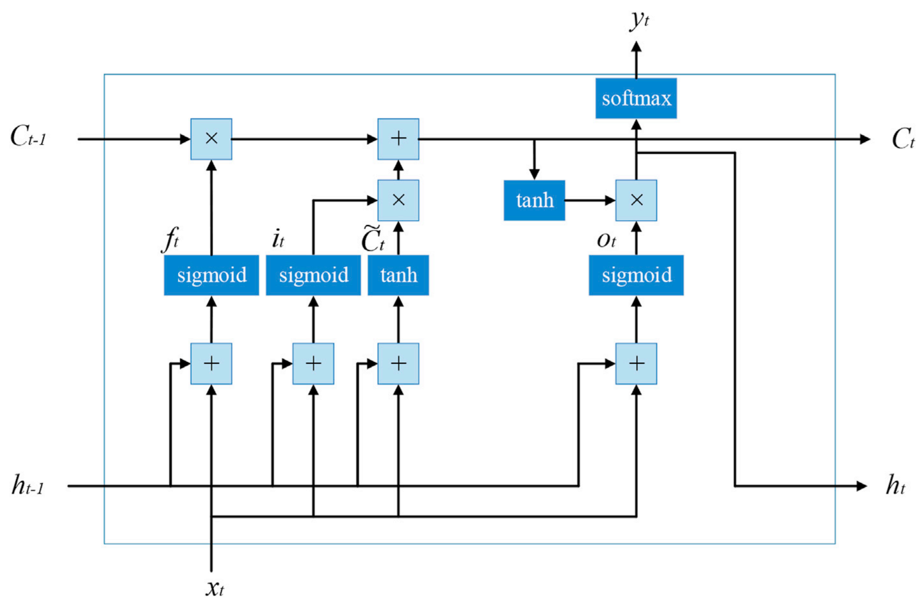

Fig. 1. The structure of LSTM unit.

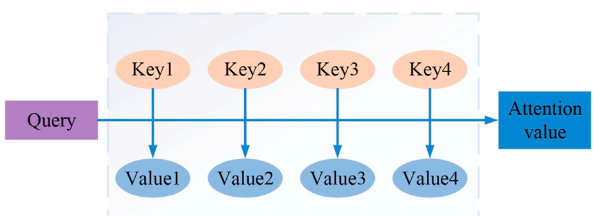

Fig. 2. The essence of attention function.

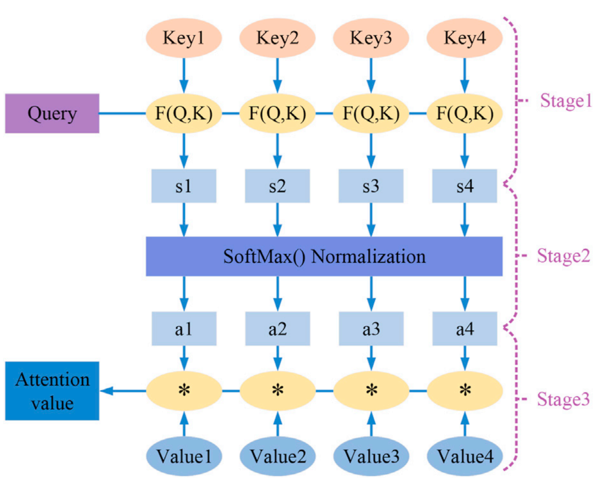

Fig. 3. The step in calculating attention.

to use the dot-product attention because of its efficiency and high speed. The calculation formula is as follows:

$$f_{mul}(Q, K) = QK^T \tag{7}$$

Then we can get the corresponding weights. In order to normalize the weight values, we need to normalize these scores. Softmax function can well realize the normalization process. In the second step, we use a softmax function to normalize the weights. In the third step, we get the final attention. And the calculation formula is as follows:

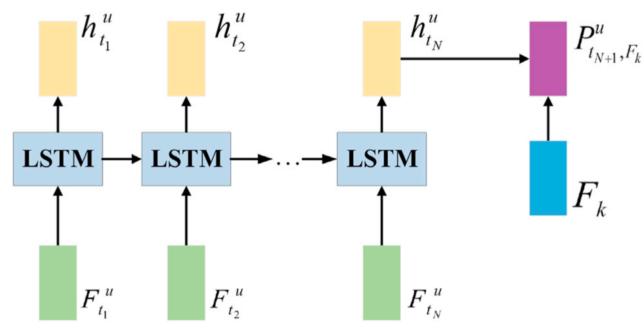

Fig. 4. The architecture of S-LSTM.

$$\text{Attention }(Q, K, V_{val}) = softmax(f(Q, K))V_{val} \tag{8}$$

4. Solution

We present the AS-LSTM model to construct sports match prediction for training and exercise.

4.1. Sports-aware LSTM model

We use LSTM to predict sports performance. Inspired by Ref. [32], we can know that LSTM is very good at processing non-linear data and capturing the non-linear dependence between the team's historical achievements. The state of a team will change with the team members' state, the external environment and many other factors. Therefore, it is impossible to predict performance only by a linear model. But our goal is to get the result of the next match by predicting the result of the game. By predicting the result of the match in advance, the coach can adjust the training of the team in time. It can make the team better, and it is conducive to the training and exercise of the team.

For our proposed LSTM model for predicting sports performance, we input the team's historical competition results into the LSTM network in chronological order to obtain the output of each unit. Finally, it predicts the probability of success and failure at the end of the game and the number of shots, goals, possession rate and goals lost in the game. The coach trains the players through the result and state of the game, thereby reducing the probability of failure and increasing the probability of success. Fig. 4 illustrates the architecture of S-LSTM.

We define the 7-tuple $F^u_{t_k}$ as the input at each time point. So we can update each hidden vector $h^u_{t_i}$ by using current input and previous hidden vector $h^u_{t_{i-1}}$. Then, using the output of LSTM unit, we can predict the results of the next team competition. The definition is as follows:

$$h^u_{t_i} = LSTM\left(W_F F^u_{t_i}, h^u_{t_{i-1}}\right) \tag{9}$$

$$P^u_{t_{N+1},F_k} = \left(W_N h^u_{t_N}\right)^T \left(W_F F^u_{t_i}\right) \tag{10}$$

where $W_F \in \mathrm{R}^{d\times d}$ and $W_N \in \mathrm{R}^{d\times d}$ are transition matrices. We will get the probability of victory or defeat. When the probability of winning exceeds 50%, the team has a greater probability of winning. When the probability of winning is equal to 50%, the probability of a draw is higher. At this time, the coach needs to change the team's strategy in time to improve the team's winning probability. When the probability of winning is less than 50%, the probability of team failure is higher. At this time, the coach needs to improve the training level of the team, pay attention to the state of the athletes, change the strategy appropriately, and strive to turn defeat into victory.

4.2. Attention-based sports-aware LSTM model

For the sports match prediction, not all the historical matches have the same important impact on the next game. Some competitions are more important, so the influence will certainly be greater, and the influence on the next competition will be greater. In contrast, some of the usual training games are relatively less influential. In the actual training match, the team may have a strategic reservation during the competition. And it is also difficult for the team members to simulate the psychological formal competition state, which may lead to insufficient preparation [33]. Therefore, for the sports match prediction, we should pay more attention to historical results, which have a greater impact on the prediction of the subsequent results. The attention mechanism accomplishes this job well. The attention mechanism will assign different weights to different historical game records [34]. The history records with more influence will be assigned a higher weight, while the competition records with less influence will be assigned a smaller weight. After performing the normalization operation, we will assign an appropriate weight to each historical record, and the sum of all weights is 1.

It is also worth noting that the prediction of sports competition results must be highly time-sensitive. If the predicted result will appear in a few weeks or even longer, there is no timeliness and no practical value. Therefore, what we need to predict is the match of the team that may appear in the next game. Inspired by Ref. [35], we use a sliding time window to focus on the team's short-term game status. The sparse available game data hinders the training of the model. Different teams have different times and frequencies of competition. Assuming that a specific time is determined as a standard sliding window, the difference of sports match time interval will seriously affect the training of the model. In contrast, it is more stable to determine a sliding window based on the number of matches. The number of matches can better ensure that there is enough continuous data to participate in the training of the model. As for the selection of the size of the sliding window, if the sliding window size is large, it is difficult for users with fewer competitions to participate in the training and testing of the model, and it is not conducive to the short-term prediction of results. A small sliding window is not conducive to capturing the dependence between the data. We have analyzed the conditions when the sliding window is 5,10,15, respectively, and judged that the model with a sliding window size of 5 is suitable. When the sliding window size is 5, it will focus on the user's first ten game states to predict the user's next game result. The size of the sliding window needs to be set according to the actual status of the team.

It can be said that AS-LSTM is an improvement of S-LSTM, and the two models can reflect the logical rationality of the paper. The AS-LSTM model we proposed includes three steps in design. First of all, we use the most basic LSTM model to predict competition results. Secondly, in order to pay more attention to the influential historical competition results, we combine LSTM with an attention mechanism. And we can get the attention weight vector and weighted hidden representation $r^u_{t_N}$ to pay more attention to the team's historical competition results. Then, we replace the original LSTM hidden layer state $h^u_{t_N}$ with the calculated attention weight to get the final probability of the match result. The calculation formula is as follows:

$$\mathrm{r}^u_{t_N} = \sum\nolimits_{i=1}^{N} a_i h^u_{t_i} \tag{11}$$

$$o^u_{t_{N+1},C_k} = \left(W_N r^u_{t_N}\right)^T \left(W_F F^u_{t_i}\right) \tag{12}$$

Third, to ensure the timeliness of the prediction results, that is, to ensure that the predicted result is the team's next match performance, we use the sliding window to better extract the team's short-term competition status. To sum up, we proposed an attention-based sports match prediction model AS-LSTM for training and exercise. Fig. 5 illustrates the architecture of AS-LSTM. And our pseudocode is shown in Algorithm 1. Although our model has considered many factors that affect the results of the sports match, there are still some other factors that have not been considered, such as the physical condition of the athletes and the unexpected factors that may appear on the court. Moreover, the addition of a sliding window can not perfectly predict the possible results of each match. Therefore, coaches need to add temporary consideration when they use the model prediction results, and they can not completely rely on our model. On the other hand, it also shows that our model is not perfect and still needs continuous improvement.

Algorithm 1. AS-LSTM

Algorithm 1 AS-LSTM

Require: teams U and their football matches historical records AFH

Ensure: ATCA-GRU model $\{M\}_u$

```
    //building training model
    B=U_u B^u=ϕ;
    for each team u in U do
        for each football match record F^u_{t_k} in FH_U do
            Compute embedding vector E^u_{t_k} of F^u_{t_k}
        end for
        Add a training instance{(E^u_{t_k}), (f'^u_{t_k})} to B^u
    end for
    //model training
    Initialize the parameter set Θ
    while exceed (maximum number if iterations) == false do
        for each team u in U do
            Randomly select a batch of instances B^u_b from B^u
            Find Θ minimizing the objective with B^u_b
        end for
    end while
    return {M}_u
```

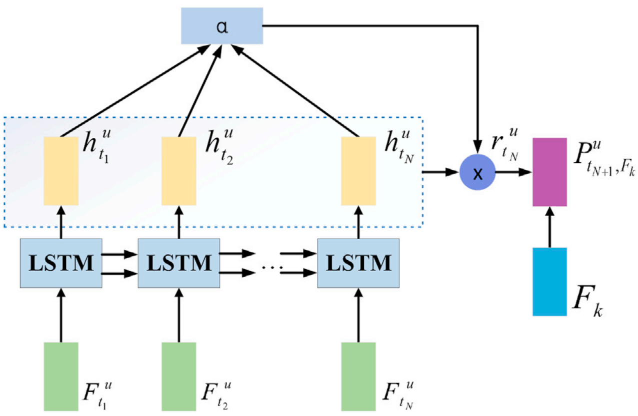

Fig. 5. The architecture of AS-LSTM.

4.3. A case study

In this section, we proposed a case study to prove the feasibility of our AS-LSTM model. We take the football match as an example. The specific steps are as follows.

Table 2
The historical match records of team A and team B.

Team	Goals	Goals lost	Ball control rate	Home court	Time	Results
A	0	2	33%	1	25-07-2019	Lose
	3	2	54%	1	30-07-2019	Win
	1	0	60%	1	05-08-2019	Win
	2	2	51%	1	10-08-2019	Draw
	1	0	66.3%	1	14-08-2019	Win
	2	3	45%	1	19-08-2019	Lose
	0	0	47%	1	26-08-2019	Draw
	1	4	25%	1	30-08-2019	Draw
	0	0	46.7%	1	09-09-2019	Draw
	…	…	…	…	…	…
	2	1	51.7%	0	21-03-2020	Win
Team	Goals	Goals lost	Ball control rate	Home court	Time	Results
B	2	0	53%	0	15-09-2018	Win
	1	3	44%	0	30-07-2018	Lose
	0	4	40%	0	05-08-2018	Lose
	1	0	51%	0	10-08-2018	Win
	1	0	62%	0	14-08-2018	Win
	1	1	45%	0	19-08-2018	Draw
	2	0	47%	0	26-08-2018	Win
	0	1	39%	0	30-08-2018	Lose
	1	1	47%	0	09-09-2018	Draw
	2	1	53%	0	09-09-2018	Win
	…	…	…	…	…	…
	1	0	51.7%	1	21-03-2019	Win

Step 1: Data preparation

We considered the record of the teams in nine months. The data includes goals, goals lost, the ball possession rate, whether the team has the home advantage, the match time and the match results. Among them, if the team has the home-court advantage, it is recorded as 1; otherwise, it is recorded as 0. For the information of each team match, we express it as a vector for subsequent input into LSTM for training. And there are three kinds of results: win, lose and draw. In the process of model training, we use the data of the first seven months as the training set to train our model. The competition data of the last two months is the validation set. Table 2 shows the historical match records of two teams (Team A and Team B), respectively.

Step 2: Add sliding window

We set the size of the sliding window to 5. That is to say, we use the results of the team's recent five games to predict the results of future matches. After adding the sliding window, we need to process the data of the team. Take team A as an example, and the processed data is shown in Fig. 6. The sliding window contains the team's five-match data at a time. Each time, a game record will be added backwards. In this way, we can

Fig. 6. The process of prediction.

Table 3
Comparison between actual results and predicted results.

Team	Actual results	Predicted results	Accuracy (%)
A	draw, lose, draw, draw, win	win, lose, lose, draw, win	60
B	win, win, draw, lose, draw	win, win, draw, lose, lose	80
…	…	…	…
N	draw, win, win, lose, lose	win, win, win, lose, lose	80

better train the model to predict the team's next match performance. Therefore, the coach can also give team guidance and special training in advance based on the team's possible match conditions.

Step 3: AS-LSTM model training

We input the data after adding the sliding window into the attention-based LSTM model for training. First, the team's match data is processed into a vector. After sorting according to the match time, a sliding window is introduced. For each team, the first sliding window contains the data of the first ten matches, and the second sliding window moves backwards once, including the data from the second match to the eleventh match, thus pressing this operation to get all the data input. By training the data of the first seven months, we can get a better prediction model and then verify the data of the next two months. In the verification, we will get the prediction results of the model and save the prediction results from judging the performance of the model.

Step 4: Performance evaluation of AS-LSTM model prediction

We can compare the predicted results with the real value to judge the performance of the model. In addition, we can choose to predict the results of the next few games (e.g., top-1, top-3, top-5). Although the model can predict the performance of the match, the most important thing is the result of the match. Therefore, we choose to show the predicted results of the next five matches and compare them with the real values. The comparison process is shown in Table 3. Therefore, it can be proved that our AS-LSTM model is true and effective.

5. Conclusion

In this paper, we propose a prediction method for sports games to help team training and exercise.To capture the dependence between the team's historical competition results, we use the LSTM model. In order to pay more attention to important competition results, we combine the attention mechanism with LSTM. Finally, in order to ensure the timeliness of sports game prediction, we add a time sliding window and propose the LSTM model of attention-based motion perception. Finally, the feasibility of the method is verified by an example. In our future research, more real experiments and applications are underway.

For future work, we will search more appropriate sliding time window, a more effective word embedding method, more accurate attention to athletes' psychological and physical states, and constantly improve the accuracy of the model prediction.

Declaration of competing interest

No conflicts of interest.

References

[1] H. Li, Analysis on the construction of sports match prediction model using neural network, Soft Computing (2020) 1–11.
[2] H. Xu, P. Wu, J. Wei, Z. Yang, J. Wang, A meta-path-based recurrent model for next poi prediction with spatial and temporal contexts, in: Asia-Pacific Web (APWeb) and Web-Age Information Management (WAIM) Joint International Conference on Web and Big Data, Springer, 2019, pp. 219–235.
[3] S. Jeon, B. Hong, V. Chang, Pattern graph tracking-based stock price prediction using big data, Future Generat. Comput. Syst. 80 (2018) 171–187.
[4] J. Luo, J. Li, L. Jiao, J. Cai, On the effective parallelization and near-optimal deployment of service function chains, IEEE Trans. Parallel Distr. Syst. 32 (5) (2021) 1238–1255.
[5] J. Cai, Z. Huang, L. Liao, J. Luo, W.-X. Liu, Appm: adaptive parallel processing mechanism for service function chains, IEEE Transactions on Network and Service Management 18 (2) (2021) 1540–1555.
[6] M.A. Rahman, A deep learning framework for football match prediction, SN Applied Sciences 2 (2) (2020) 165.
[7] Y.J. Park, H.S. Kim, D. Kim, H. Lee, S.B. Kim, P. Kang, A deep learning-based sports player evaluation model based on game statistics and news articles, Knowl. Base Syst. 138 (2017) 15–26.
[8] A.H. Montoye, M. Begum, Z. Henning, K.A. Pfeiffer, Comparison of linear and non-linear models for predicting energy expenditure from raw accelerometer data, Physiol. Meas. 38 (2) (2017) 343.
[9] W.R. Johnson, J. Alderson, D. Lloyd, A. Mian, Predicting athlete ground reaction forces and moments from spatio-temporal driven cnn models, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 66 (3) (2018) 689–694.
[10] L. Wu, Z. Yang, Q. Wang, M. Jian, B. Zhao, J. Yan, C.W. Chen, Fusing motion patterns and key visual information for semantic event recognition in basketball videos, Neurocomputing 413 (2020) 217–229.
[11] J. Huang, N. Li, T. Li, S. Liu, G. Li, Spatial-temporal context-aware online action detection and prediction, IEEE Transactions on Circuits and Systems for Video Technology.
[12] M. Woźniak, M. Wieczorek, J. Siłka, D. Połap, Body pose prediction based on motion sensor data and recurrent neural network, IEEE Transactions on Industrial Informatics 17 (3) (2020) 2101–2111.
[13] M. Keshtkar Langaroudi, M. Yamaghani, Sports result prediction based on machine learning and computational intelligence approaches: a survey, Journal of Advances in Computer Engineering and Technology 5 (1) (2019) 27–36.
[14] N. Watson, S. Hendricks, T. Stewart, I. Durbach, Integrating machine learning and decision support in tactical decision-making in rugby union, J. Oper. Res. Soc. (2020) 1–12.
[15] D. Strnad, A. Nerat, Š. Kohek, Neural network models for group behavior prediction: a case of soccer match attendance, Neural Comput. Appl. 28 (2) (2017) 287–300.
[16] C. Li, D. Song, D. Tao, Multi-task recurrent neural networks and higher-order markov random fields for stock price movement prediction: multi-task rnn and higer-order mrfs for stock price classification, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019, pp. 1141–1151.
[17] M. Chen, W.-Z. Li, L. Qian, S.-L. Lu, D.-X. Chen, Next poi recommendation based on location interest mining with recurrent neural networks, J. Comput. Sci. Technol. 35 (2020) 603–616.
[18] M. Oytun, C. Tinazci, B. Sekeroglu, C. Acikada, H.U. Yavuz, Performance prediction and evaluation in female handball players using machine learning models, IEEE Access 8 (2020) 116321–116335.
[19] C. Tian, L. Fei, W. Zheng, Y. Xu, W. Zuo, C.-W. Lin, Deep learning on image denoising: an overview, Neural Network. 131 (2020) 251–275.
[20] Y. Li, H. Ma, L. Wang, S. Mao, G. Wang, Optimized content caching and user association for edge computing in densely deployed heterogeneous networks, IEEE Trans. Mobile Comput..doi:10.1109/TMC.2020.3033563.
[21] W. Zaremba, I. Sutskever, O. Vinyals, Recurrent Neural Network Regularization, arXiv preprint arXiv:1409.2329.
[22] Y. Zhao, R. Yang, G. Chevalier, R.C. Shah, R. Romijnders, Applying deep bidirectional lstm and mixture density network for basketball trajectory prediction, Optik 158 (2018) 266–272.
[23] W. Wei, H. Wu, H. Ma, An autoencoder and lstm-based traffic flow prediction method, Sensors 19 (13) (2019) 2946.
[24] Y. Li, C. Liao, Y. Wang, C. Wang, Energy-efficient optimal relay selection in cooperative cellular networks based on double auction, IEEE Trans. Wireless Commun. 14 (8) (2015) 4093–4104.
[25] Y. Yan, J. Ren, G. Sun, H. Zhao, J. Han, X. Li, S. Marshall, J. Zhan, Unsupervised image saliency detection with gestalt-laws guided optimization and visual attention based refinement, Pattern Recogn. 79 (2018) 65–78.
[26] Y. Li, S. Xia, B. Cao, Q. Liu, et al., Lyapunov optimization based trade-off policy for mobile cloud offloading in heterogeneous wireless networks, IEEE Transactions on Cloud Computing.doi:10.1109/TCC.2019.2938504.
[27] M. Zhang, V. Palade, Y. Wang, Z. Ji, Attention-based word embeddings using artificial bee colony algorithm for aspect-level sentiment classification, Inf. Sci. 545 (2021) 713–738.
[28] J. Guo, X. Nie, Y. Ma, K. Shaheed, I. Ullah, Y. Yin, Attention based consistent semantic learning for micro-video scene recognition, Inf. Sci. 543 (2021) 504–516.
[29] K. Wu, J. Wu, L. Feng, B. Yang, R. Liang, S. Yang, R. Zhao, An attention-based cnn-lstm-bilstm model for short-term electric load forecasting in integrated energy system, International Transactions on Electrical Energy Systems (2020), e12637.
[30] Y. Xie, R. Liang, Z. Liang, C. Huang, C. Zou, B. Schuller, Speech emotion classification using attention-based lstm, IEEE/ACM Transactions on Audio, Speech, and Language Processing 27 (11) (2019) 1675–1685.

[31] S. Zhang, G. Wu, J. Gu, J. Han, Pruning convolutional neural networks with an attention mechanism for remote sensing image classification, Electronics 9 (8) (2020) 1209.

[32] M. Sangiorgio, F. Dercole, Robustness of lstm neural networks for multi-step forecasting of chaotic time series, Chaos, Solitons & Fractals 139 (2020) 110045.

[33] J. Stübinger, B. Mangold, J. Knoll, Machine learning in football betting: prediction of match results based on player characteristics, Appl. Sci. 10 (1) (2020) 46.

[34] X.-m. Yu, W.-z. Feng, H. Wang, Q. Chu, Q. Chen, An attention mechanism and multi-granularity-based bi-lstm model for Chinese q&a system, Soft Computing 24 (8) (2020) 5831–5845.

[35] Z. Zhu, Y. Wang, X. Zhou, L. Yang, G. Meng, Z. Zhang, Swav: a web-based visualization browser for sliding window analysis, Sci. Rep. 10 (1) (2020) 1–4.