Computers, Materials & Continua

Tech Science Press

DOI: 10.32604/cmc.2023.037293
Article

Regional Economic Development Trend Prediction Method Based on Digital Twins and Time Series Network

Runguo Xu*, Xuehan Yu and Xiaoxue Zhao

School of International Relations, Yonsei University, Seoul, 03722, Korea
*Corresponding Author: Runguo Xu. Email: xrg96@yonsei.ac.kr
Received: 29 October 2022; Accepted: 19 May 2023; Published: 30 August 2023

Abstract: At present, the interpretation of regional economic development (RED) has changed from a simple evaluation of economic growth to a focus on economic growth and the optimization of economic structure, the improvement of economic relations, and the change of institutional innovation. This article uses the RED trend as the research object and constructs the RED index to conduct the theoretical analysis. Then this paper uses the attention mechanism based on digital twins and the time series network model to verify the actual data. Finally, the regional economy is predicted according to the theoretical model. The specific research work mainly includes the following aspects: 1) This paper introduced the development status of research on time series networks and economic forecasting at home and abroad. 2) This paper introduces the basic principles and structures of long and short-term memory (LSTM) and convolutional neural network (CNN), constructs an improved CNN-LSTM model combined with the attention mechanism, and then constructs a regional economic prediction index system. 3) The best parameters of the model are selected through experiments, and the trained model is used for simulation experiment prediction. The results show that the CNN-LSTM model based on the attention mechanism proposed in this paper has high accuracy in predicting regional economies.

Keywords: Regional economic development; attention mechanism; digital twins; time series network

1 Introduction

In the economic society, the government's correct macro-control is based on the accurate analysis of various economic information. Through accurate economic forecasts, the government can take corresponding economic adjustment measures to properly control economic activities so that economic activities can achieve the desired effect and improve their operation quality [1]. Accurate economic forecasting will greatly enhance the predictability of economic policies and reduce mistakes in economic decision-making and regulation. Therefore, economic forecasting has become an important basis for countries and regions to formulate economic policies, make investment decisions, and formulate development plans [2]. With the development of China's economy and the profound reform of its economic system, macroeconomic regulation is playing an increasingly important role. Economic forecasting is more and more widely used in various economic fields such as planning, statistics, operation, and management, and its significance is increasingly recognized and concerned by people [3]. From the scope involved, economic forecasting can be divided into three categories: macroeconomic forecasting, microeconomic forecasting, and intermediate-level economic forecasting. Macroeconomic prediction mainly studies the prediction of national income, the development trend of industrial and agricultural output value, the change of economic structure, and the effect of economic system reform and policy adjustment [4]. Microeconomic forecast refers to various forecasts made within the scope of enterprises with independent economic accounting. It takes the economic activities carried out by an enterprise as the research object and predicts the development trend of the enterprise's product output, sales volume, market share, etc. [5]. The research object of this paper is regional economic forecast, which belongs to the category of middle-level economic forecast and is between the conceptual analysis in the macro scope and the specific analysis in a micro-scale. Regional economic forecasting is mainly to conduct quantitative analysis, processing, and forecasting of economic data such as the development scale and speed, production and sales, resource development, and the regional balance of the various regional sectors. It is a necessary condition and fundamental basis for correctly planning regional development, developing strengths and avoiding weaknesses, and giving play to regional advantages [6]. Economic forecasting has three characteristics: scientific, approximate, and limited. First, economic forecasting is based on historical data, recent data information, and certain economic theoretical background, using certain methods and models. It analyzes the correlation between the prediction object and the relevant factors, thereby revealing and summarizing the characteristics and changing rules of the economic prediction object. Therefore, the prediction is scientific [7]. Secondly, prediction is the estimation and speculation of the state of economic events before they occur, and the development of things is not a simple repetition. This is always affected by various factors, so there are always some deviations between the prediction in advance and the actual results in the future, which can only be approximated. Therefore, economic forecasting has a strong approximation, and the statistical forecasting method is an approximation in probability [8]. Finally, the forecasters' understanding of the research object is always limited by their knowledge, experience, and prediction art ability. Sometimes, the inaccuracy and incompleteness of economic data will lead to the accuracy and one-sidedness of forecast results. Therefore, the economic forecast results have certain limitations [9]. A correct understanding and analysis of these characteristics of economic forecasting can avoid hindering the research and application of forecasting due to incorrect views. Doubt and denial of forecast results without analysis will lead to an insufficient basis for future planning and decision-making [10]. If the government fully believes in the predicted results, the work will lack flexibility. It is not objective and realistic to demand the accuracy of prediction too much [11]. In traditional economic forecasting theory, scholars generally use regression analysis and other methods to fit, analyze and forecast relevant economic indicators based on mathematical modeling [12]. The artificial neural network (ANN) model is an intelligent bionic model based on physiology. Based on neuropsychology and cognitive science findings, it employs mathematical techniques to create a parallel distributed mode processing system [13]. It mimics the thinking ability of the human brain and finds out the rules through continuous learning, memory, calculation, and intelligent processing of many cases. Nonlinear modeling occurs naturally, eliminating the need to isolate the nonlinear connection; this greatly simplifies the procedure [14]. Regardless of the type of system models, the structure of neural networks used to express or describe these models is unchanged. Because of these characteristics, neural networks have strong learning abilities and adaptability. The complexity of the economic system has led to some drawbacks in the traditional methods of economic forecasting. At the same time, the good performance of the ANN technology has made the application of neural networks to economic forecasting research a valuable and potential research topic [15]. Therefore, the main contributions of this paper are as follows:

- First, this paper summarizes the development status of research on time series networks and economic forecasting at home and abroad.

- Second, we have referred to relevant literature, summarized regional economic development characteristics, and then designed a regional economic development prediction index system.

- Thirdly, this paper introduces the basic principles and structures of long and short-term memory (LSTM) and convolutional neural network (CNN) and builds a better CNN-LSTM model using an attention mechanism.

- Finally, this paper uses the improved CNN-LSTM model to predict the development trend of regional economies and obtains good prediction results.

The work arrangement of the whole article is as follows: Section 2 will introduce the research progress of economic forecasting; Section 3 will propose the design of a regional economic development prediction index system and show the principle and structure of LSTM and CNN; Section 4 will carry out experimental analysis on the proposed method; Section 5 will summarize all the previous work.

2 Related Work

Time series analysis and multiple regression are two of the most common methodologies used for economic forecasting. Due to technological advances, neural network models have recently been developed in economic forecasting. Time series data is often utilized for economic forecasting [16]. A time series is a data set representing a certain statistical variable plotted against time. Time series forecasting often uses statistical regression models for modeling and analysis. Two typical techniques for using statistical regression models to address time series issues are the autoregressive moving average model and the autoregressive summation moving average model [17]. The basic idea of these two models is to smooth the unstable time series data found after smooth detection. The time series is differentiated, and then the series is represented as a combination of white noise and the moving average of a previous point [18]. Another common prediction method is the vector autoregression (VAR) method. This method has been widely used in econometric research since it was proposed. VAR does not require any implementation constraints. The establishment of the model mainly depends on the statistical nature of the data, which has been widely used in economic forecasting in recent years [19]. The above two analysis methods are common algorithms that use linear models to solve economic forecasting problems, and they are difficult to fit the nonlinear characteristics of complex giant systems. In the face of complex nonlinear problems, researchers jump out of such traditional algorithms to find better solutions. Huang et al. [20] pioneer using ANN to address time series issues by using neural networks for economic forecasting. Through testing comparisons, Hwarng [21] demonstrates that neural network prediction performance for time series prediction is on par with that of the auto-regressive and moving average (ARMA) model. Al-Maqaleh et al. [22] conduct a comparative study on statistical models such as combination forecasting and neural network models. Evidence from this study demonstrates that neural network prediction results are superior to those obtained using more conventional statistical approaches, particularly when a multi-step prediction is involved. Chinese academics have also conducted numerous studies employing ANN for economic system prediction. Cao et al. [23] is China's first group of scholars to apply neural network methods to economic forecasting. They compared the neural network method with the traditional method.

Zhang et al. [24] apply a neural network to multi-step prediction. The findings demonstrate the neural network's approximation of any nonlinear function, which has significant benefits for tackling difficult issues like economic forecasting. The use of neural networks in economic forecasting has been growing with the rise of big data technologies. Yu [25] applies the radial basis function neural network to the macroeconomic prediction and makes a good prediction on the gross domestic product (GDP) of a province. The consumer price index (CPI) and a province's total exports and imports are two economic indicators that are predicted using a combination of deep belief nets (DBN) depth learning and a back propagation neural network (BPNN). From the standpoints of prediction accuracy, convergence speed, etc., they validated the outstanding performance of deep learning in economic forecasting [26]. At present, mainstream neural network models have been used to study the prediction of time series, including CNN, recurrent neural network (RNN), etc. Chuang et al. [27] use the combination of ANN and CNN to predict the relationship between events and stock trends. Chen et al. [28] use RNN to predict the fluctuation difference of the stock index. RNN is more suitable for analyzing and fitting sequential data because of its unique network structure. However, RNN has some defects, such as large time intervals and long sequence data. Because the continuous multiplication effect in gradient inverse multiplication will cause gradient disappearance, it is unable to learn long-term dependence. The emergence of LSTM has solved this problem well. It modifies the internal structure of the network unit and solves the defects [29]. Attention mechanism has also emerged in time series research many times. Hu et al. [30] suggested that the time series period should be included in the attention weight calculation of the original attention mechanism. Liu et al. [31] propose to adopt a two-stage attention mechanism, focusing on the data of time series and time steps, respectively. This has achieved better results than a single mechanism in financial data prediction. Cai et al. [32] assign varying weights to the semantic text coding; it takes the features produced from the bidirectional LSTM network and feeds them into the attention layer. With this, named entity recognition (NER) is much more precise. Meng et al. [33] use the attention layer to receive the output of the LSTM network and calculate the prediction value through the full connection layer. The experiment obtains high prediction accuracy by mining the information relationship between electricity price and load law in a certain area. In general, the attention mechanism is still a new concept, and the research on time series prediction needs to be deepened, which is worth mining. The effective prediction of time series data becomes urgent in the era of information explosion. To study the problem of neural sequence, many methods have been proposed by academic and financial circles. The concept of the deep learning method and its important attention mechanism has been widely used and has far-reaching research value. Aiming to forecast the regional economy's development trend, this paper combines an attention mechanism with a neural network to conduct research.

3 Proposed Method

3.1 LSTM Network Structure

Compared with RNN, LSTM has introduced a gated real-time memory control unit into the hidden gate layer. And it is used to save all the historical control data and the running status of the secret door in the layer for a long time. The hidden gate of LSTM can include three kinds of gates in the layer: input gate, forgetting gate, and output gate. These doors are system users without door locks and control doors who can directly use them as a complete information connection layer. LSTM's data storage and information update of various control information are directly designed and implemented by these gated control system applications. This gating control system is mainly realized by the function formula of the sigmoid and the length of the input point. Gating can be expressed in the following forms:

$$g(x) = \sigma(W_x + b) \tag{1}$$

where $\sigma(x)$ is a sigmoid function, which is a commonly used activation function in the field of neural networks, it has the advantages of smoothness and easy derivation, and its value range is between 0 and 1. If the output value is 0, no information passes; if the output value is 1, all information passes.

For the LSTM forward propagation process, we have Eqs. (2)–(4). The input gate's primary role is to determine what percentage of the network's current input data will be written to storage immediately. The value formula is defined as:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \tag{2}$$

where x_t and h_t represent the input vector and output vector of the hidden layer.

The forgetting gate plays a vital role in an LSTM-gated memory unit. Not only does it allow for fine-grained regulation of memory retention and recall, but it also protects against the disappearance and explosion of gradient that may result from the reversal of gradient propagation. The value f_t of the forgetting gate and the value c_t of the memory unit can be expressed as follows:

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \tag{3}$$

$$c_t = f_t \otimes c_{t-1} + i_t \otimes tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \tag{4}$$

The output gate's primary purpose is to regulate the impact of a memory processing unit on the input and output values of this data steadily and predictably. That is, the information may be continuously transmitted to each person at the same critical moment at a certain key position in the whole memory processing unit system. The value o_t of the output gate and the output h_t of LSTM at time t can be expressed as follows:

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \tag{5}$$

$$h_t = o_t \otimes tanh(c_t) \tag{6}$$

3.2 Activation Function

In order to convert multiple linear inputs into nonlinear relations, it is necessary to add excitation functions between the input and output of the hidden layer and the output layer. This section will introduce three activation functions and select the most suitable activation function for this algorithm according to its principle characteristics, advantages, and disadvantages.

1) Sigmoid function is a joint neural activation calculation function in the neural network application field. Its definition field is expressed as an actual number field, with a value range between (0, 1). A substantial number can be mapped to this range. The sigmoid function has the following advantages: smooth and easy to derive, monotonous output, stable between (0, 1), and suitable for the output layer. At the same time, the Sigmoid function also has some disadvantages: the calculation process is extensive, and division may be involved in derivation. Its soft saturation quickly leads to the disappearance of the gradient.

$$\sigma(x) = \frac{1}{(x + exp(-x))} \tag{7}$$

2) Tanh function is a hyperbolic function. It can be expressed as:

$$\tan h\left(x\right) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{8}$$

The tanh function has the following advantages: its convergence speed is better than the sigmoid function, saving computing time. Its output is zero-centric. At the same time, the tanh function also has shortcomings; its gradient vanishing problem has not been solved.

3) Relu function is a typical activation function, which can be expressed as:

$$f\left(x\right) = max(0, x) \tag{9}$$

Relu function is essentially a piecewise linear function for maximum value, in which the negative value is reduced to zero, and the positive value remains unchanged. In other words, the neuron will not be triggered if the input is negative and the output is zero. This will allow for a sparser network by activating fewer neurons simultaneously, speeding up calculations, and decreasing computational demands.

3.3 Model Optimizer Algorithm

When we choose the objective function J with a given parameter θ, we always hope to find aanathemamake the objective function J obtain the maximum or minimum value. The optimization algorithm is the way to find the θ. In the neural network model, the objective function J is generally the error between the predicted value and the tag. We hope to see theta that can make J the smallest. Several standard optimizer algorithms are described below.

1) Gradient descent (GD). GD is the first neural network optimizer algorithm used. Its main idea is to select the fastest descent direction of the current position, that is, the negative gradient direction, as the search direction.

2) Adagrad algorithm. In this algorithm, the update of low-frequency parameters is relatively large, but the update of high-frequency parameters is relatively small. It improves the robustness of GD and performs well on sparse data. The algorithm update rules are as follows:

$$\theta_{t+1,i} = \theta_{t,i} - \frac{\eta}{\sqrt{G_{t,ii} + \xi}} \cdot g_{t,i} \tag{10}$$

$$g_{t,i} = \nabla_{\theta} J(\theta_i) \tag{11}$$

where g is the gradient of parameter θ_i at time t.

Adagrad can reduce the learning rate manually, but the denominator will continue accumulating, and the learning rate may shrink to a minimal level.

3) Adam algorithm. Adam is an adaptive learning rate optimization method. The updated rules are as follows:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \cdot g_t \tag{12}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) \cdot g_t^2 \tag{13}$$

The main advantage of the Adam algorithm is that the learning rate of first-order moment estimation and second-order moment estimation will fall within a specific range after the offset check. This makes Adam's parameters more stable.

3.4 Convolution Neural Network

CNN's representational and learning technological capabilities have been the focus of much interest since the idea of substantial study in metaphorical learning was presented. The computer's fast development may be traced to the increase in its capacity to handle complicated numerical values and sequences. In the large-scale visual recognition competition of ImageNet, many complex CNN, including AlexNet, performed well. The overall input network management architecture of the convolutional input neural network is mainly divided into pooled input management layer, convolutional management layer, pooled input management layer, complete output connection input management system layer, and input/output connection management layer.

1) Input layer. Each data input layer of CNN must be a group of data that can be used to process multi-dimensional data simultaneously. Generally, each data input layer of one-dimensional CNN must accept one-dimensional or two-dimensional arrays at the same time. One dimension array is usually used for real-time sampling of a specific time or spectrum. Similarly, each data input layer of two-dimensional CNN must accept two-dimensional or three-dimensional arrays at the same time. Each data input layer of three-dimensional (3D) CNN must get the grouping of four-dimensional convolution functions simultaneously. The training algorithm of CNN is called GD; therefore, to improve the model's learning efficiency, it is necessary to standardize the channel or time dimension when inputting features. Normalization is the most common method to normalize data.

2) Convolutional layer. The layer implied by convolution in CNN is also widely referred to as the inner convolution layer, the essential part of all CNNs. The output precision data and information model of the convolution hidden layer is analyzed and extracted using the acquisition and processing function of the coating. Its inner workings comprise dual-core precision cells with several convolution layers. In the same way, neurons measure the coefficient and accuracy deviation, each precision element in the convolution layer monocyte must have the same or equal precision weight. The neurons in each single-layer network convolved in the bottom layer are connected with those in multiple single-layer networks in the network area with similar physical positions in the previous layer. The size of the material location deviation in this area mainly depends on the size of the material location deviation of each network convolution layer core. When the core processing system of each layer network is working, the convolution layer core processing system usually quickly traces the characteristics of each input signal of each layer network according to specific physical laws. The input network characteristic signs are calculated to make a quadratic matrix, and the quadratic multiplication of the elements is multiplied by the sum. A new deviation is calculated and added again. The basic parameters of the inner layer of the convolution level can mainly include three, namely, the area size, step size, and step filling of the image size of the convolution layer. Any tiny number or a value less than the size of the network's input information may be supplied directly as the size of the convolution layer core. The more complex the network's input information properties, the bigger the convolution layer core has to be. For every two iterations of the convolution layer kernel, an image of the network's feature pixels is automatically scanned, as indicated by the step size provided by the step size filling setting technique. When the pixel step involved in the layered kernel network is automatically set to $n-1$, the convolution layer kernel may automatically scan each feature element of the network feature pixel map one by one. If the step size is set to n, the convolution layer kernel may watch it once through an n-1 feature pixel step size after a jump. The method of step filling is defined as subjectively increasing the image size in the user network before the scanner in the convolution layer kernel inputs the image from the user network. And a scanning method that simultaneously counteracts the influence of network computing on the shrinking speed of the image size in the user's network. According to the actual application needs, the number of layers of network pictures and specific scanning targets. In the convolution layer

structure, there is an excitation function to help the complex features that are difficult to express be better understood by the machine, which can be expressed in the following form:

$$A_{i,j,k}^{l} = f(Z_{i,j,k}^{l}) \tag{14}$$

The excitation functions are often convolved to the kernel before they are designed. That is to say, some algorithmic designers may still adopt a technical operation mode of pre-expected activation and set the excitation function kernel to be placed in front of the convolved kernel.

The pooling layer's job may be summarized as "downscaling" the original high-resolution picture. A smaller matrix and fewer nodes in the whole connection layer result from using a pooling layer, which reduces the parameters of the entire neural network. This can not only speed up training but also prevent overfitting. With a large amount of data processed by the convolution and pooling layers, we can determine whether the information has been abstracted into a feature with high information content. The full connection layer structure in CNN is equivalent to a hidden layer in ordinary feedforward neural networks, and it is the last hidden layer structure in CNN. And this is only transferred between full connection layers. As a result, the input picture is not present in the spatial topology of the entire connection layer, but rather, it is extended as a vector. Complete connection layers are not equipped to perform feature combinations on their own. The suggested feature combination of a convolution layer and a pooling layer is what allows it to perform its role of obtaining input nonlinearly.

3) Output layer. The structure and working principle of the CNN output layer are the same as those of the ordinary feedforward neural network. Different objective functions can be set in different situations to meet the needs.

The CNN-LSTM network flow chart of the regional economic forecast is shown in Fig. 1.

3.5 Attention Mechanism

In modern human cognitive science, people usually ignore the rest because they selectively focus on some of the information. This is mainly because different parts of the modern retina are sensitive to information processing and analysis. This mechanism is called the attention mechanism. Sometimes when a lot of information and vision is limited, we need to focus on the core part of the visual area. Therefore, introducing the attention mechanism can be roughly divided into two steps. The first is to set the information that needs to be focused on and considered, allocating limited information processing data resources. At present, there are mainly two forms of attention management mechanisms in practical tasks: soft attention and hard attention. The soft attention training mechanism refers to not simply selecting one of all the input information. Instead, the weighted average of all the input information is calculated and then input into the neural network for training. The hard attention mechanism refers to randomly selecting any input information or input with maximum probability. Generally speaking, the soft attention mechanism is widely used. In the practical application of the soft attention mechanism, we can define an attention variable to represent the index position $A \in [1, N]$ of the input information. $A = i$ represents the ith input information, and $s(X_i, q)$ represents the attention scoring function, mainly including the following.

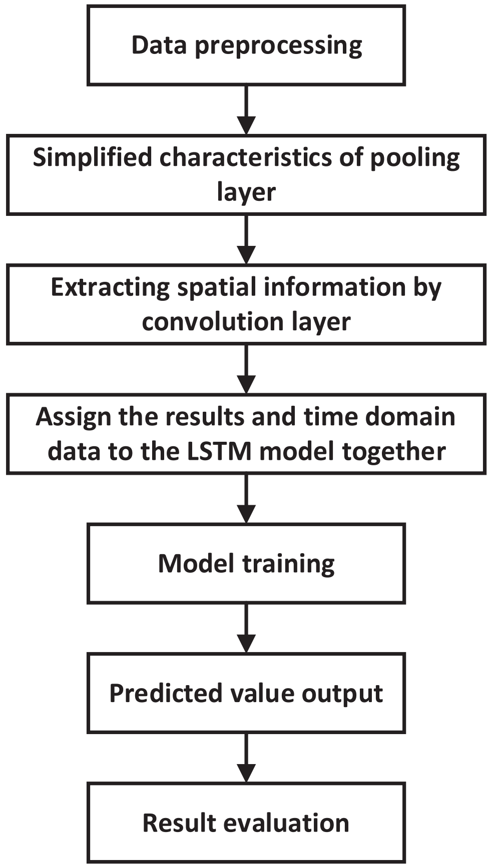

Figure 1: Flow chart of CNN-LSTM neural network prediction

Additive model:

$$s\left(X_i, q\right) = V^T tanh(WX_i + U_q) \tag{15}$$

Point product model:

$$s\left(X_i, q\right) = X_i^T q \tag{16}$$

Scale point product model:

$$s\left(X_i, q\right) = \frac{X_i^T q}{\sqrt{d}} \tag{17}$$

Bilinear model:

$$s\left(X_i, q\right) = X_i^T W_q \tag{18}$$

where W, U, and V are virtual network physical parameters that can be learned by themselves. d is the dimension of input information; this paper uses the attention mechanism under the additive model.

The attention distribution vector is represented by α_i:

$$\alpha_i = p\,(A = i|X, q) = softmax\,(s\,(X_i, q)) = \frac{exp(s\,(X_i, q))}{\sum_{j=1}^{N} exp(s\,(X_j, q))} \tag{19}$$

The weighted average attention value is:

$$Att\,(X, q) = \sum_{i=1}^{N} \alpha_i X_i \tag{20}$$

The model structure after applying the attention mechanism is shown in Fig. 2.

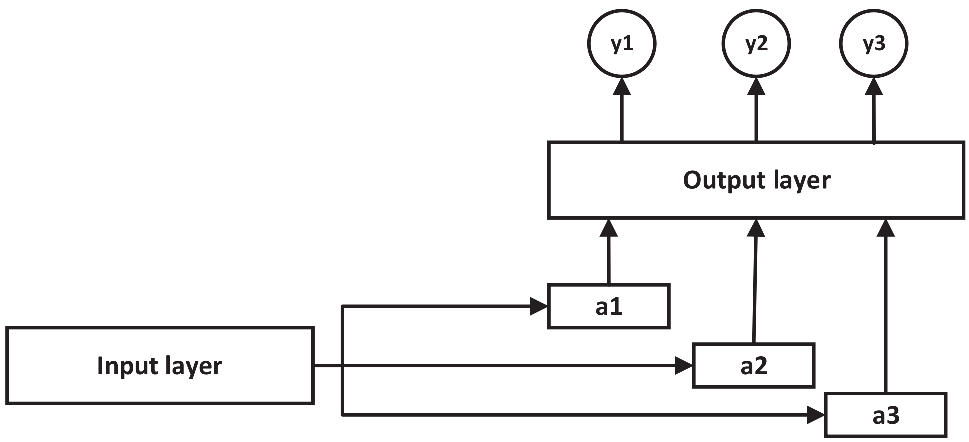

Figure 2: Schematic diagram of attention mechanism

3.6 Construction of Regional Economic Development Evaluation Index System

The evaluation index system of regional economic development (RED) should reflect the connotation of RED. The selected indicators should reflect the scientific content of economic development, namely, economic growth, economic structure optimization, improvement of economic relations, innovation of economic systems, and sustainability and coordination of economic development. The indicator system can comprehensively reflect these contents so that the indicator system can comprehensively reflect all the contents of RED. At the same time, the RED evaluation index system requires that the indicators we have established can accurately reflect the connotation of these aspects. Therefore, when establishing indicators, we should select or construct those indicators that can accurately reflect the connotation of economic development to evaluate RED. The particularity of economic development indicators is reflected in the strong correlation between many economic development and structural indicators. However, these indicators reflect different economic contents from the connotation of economic development. Therefore, it is necessary to eliminate the information related to indicators and fully reflect the economic content. This requires us to propose some indicators that can reflect economic development when building the indicator system. According to the basic laws of RED, the evaluation index system of RED we have constructed is shown in Table 1.

Table 1: Evaluation index system of regional economic development

Index	Label
Gross Domestic Product (GDP) growth index	E1
Per capita investment	E2
The growth rate of tourism foreign exchange income	E3
The growth rate of fiscal revenue	E4
The proportion of operating surplus in GDP	E5
The proportion of tertiary industry	E6
The proportion of rural residents' consumption	E7
Propensity to consume	E8
Price index	E9
The growth rate of trade volume	E10
Per capita income ratio between urban and rural areas	E11
The proportion of government consumption in GDP	E12
Per capita cultivated land m	E13
The growth rate of technical turnover	E14
The investment growth rate of renewal and transformation funds	E15

4 Experiments

4.1 Data Preprocessing

The data set used in the prediction experiment of the RED trend in this paper comes from the published data on the official website. The data includes the historical data of RED indicators from January 2010 to December 2020. The output of the model is regional GDP (RGDP). Values for the same characteristic may vary greatly depending on the underlying dimension it uses. The accuracy of predictions may suffer if distance calculations and gradient descent methods are not executed. It is important to standardize feature dimensions and value ranges such that there is no negative effect from having them vary; we need to normalize the data in advance. The formula is as follows:

$$x_{nor} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{21}$$

4.2 Model Parameter Selection

1) Selection of activation function: three activation function tests based on the experimental sample data selected in this paper are shown in Fig. 3.

According to the experimental results in Fig. 3, the performance of the tanh and the relu activation functions on the economic forecast data set used in this paper has little difference. For the convenience of the experiment, the tanh function is chosen as the activation function.

2) Selection of time step: in this section, the time step is set to 1 to 8, respectively, for comparative experiments to evaluate the impact of the time step on the model effect. Fig. 4 shows the training of the LSTM model under different time steps of the economic forecast dataset.

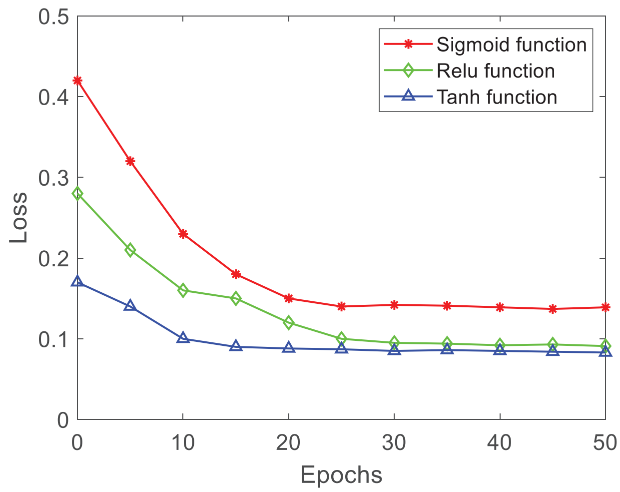

Figure 3: Performance of training set under three activation functions

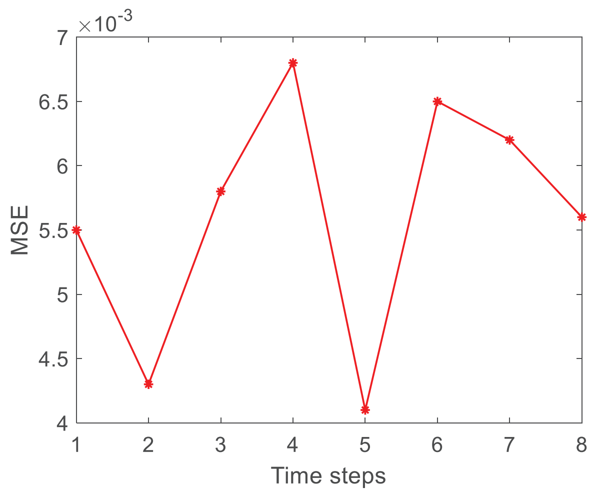

Figure 4: The training effect of the LSTM model under different time steps

According to the analysis of the experimental results in Fig. 4, it is found that the error fluctuation of the economic forecast sample data set is slight under different time steps. However, as the time step increases, the number of LSTM network loop layers increases, and the training time of each epoch will also increase. Based on training time and error considerations, the economic forecast dataset modeling time step is 1.

3) When the three optimizer algorithms are used as optimizer functions of loss functions, the simulation results on the experimental samples selected in this paper are shown in Fig. 5.

According to the analysis of the above experimental results, The Adam function is selected as the optimization function of the loss function in this paper.

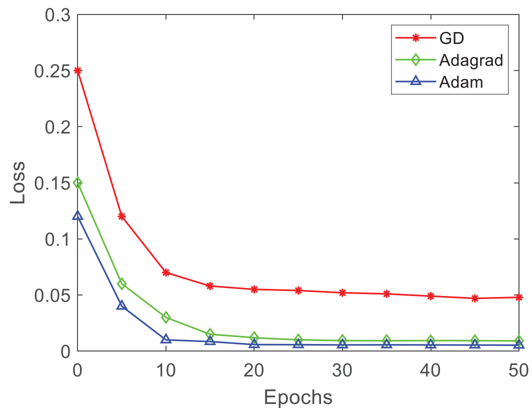

Figure 5: Experimental error rates of three different optimizer algorithms

4.3 Simulation Results

After the model parameters are selected, the trained model is used for prediction experiments. The results are shown in Table 2. The experimental results show that the model's prediction accuracy is very high.

Table 2: Comparison of prediction results and actual values under the RGDP dataset

Model (Trillion)	2017	2018	2019	2020
Actual value	5.25	5.75	5.98	5.86
Our model	5.24	5.76	5.96	5.85

4.4 Performance Comparison of Different Models

In this paper, accuracy and recall are selected as evaluation indicators to compare the prediction performance of different models. The experimental results are shown in Fig. 6. It can be seen that the CNN-LSTM model with the attention mechanism proposed in this paper has the best performance.

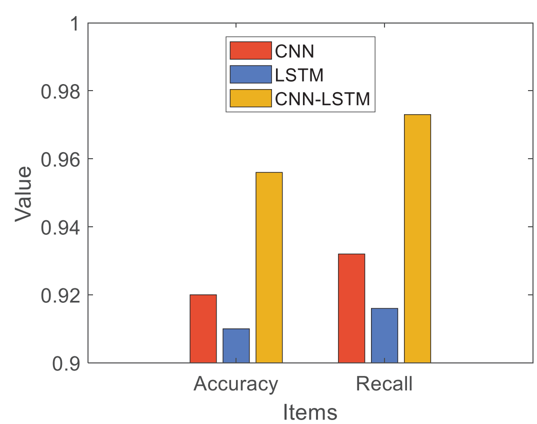

Figure 6: Comparison of accuracy and recall of different models

5 Conclusion

The traditional economic development prediction focuses on predicting the growth level of the total economic volume. In the primary stage of Western economic development, economic growth is used to evaluate the level of economic development. In the early days of reform and opening up, China also took economic growth as the leading indicator of economic development. When the economy develops to a certain level, the optimization of economic structure, the improvement of economic relations, and the innovation of systems have also become crucial economic development goals. In the past, paying attention to economic growth while neglecting optimization, relationship improvement, and system innovation within the economy will inevitably affect the overall development of the economy. And it will cause a series of contradictions, which will affect the speed and level of economic growth. In practice, forecasting is an essential work to guide macroeconomic development and observe economic activities from a macro perspective. This is of great value to avoid and reduce the phenomenon of direct government intervention or excessive regulation of economic activities. This article takes the regional economic development trend as the research object and constructs the regional economic development index to conduct the theoretical analysis. Then this paper uses the attention mechanism based on digital twins and the temporal network model to verify the actual data. Finally, the regional economy is predicted according to the theoretical model. The specific research work mainly includes the following aspects: 1) This paper introduces the development status of research on time series networks and economic forecasting at home and abroad, which provides a theoretical basis for the later model. 2) This paper introduces the basic principles and structures of LSTM and CNN networks, constructs an improved CNN-LSTM model combined with the attention mechanism, and then constructs a regional economic prediction index system through data screening. 3) The best parameters of the model are selected through experiments, and the trained model is used for simulation experiment prediction. The results show that the CNN-LSTM model based on the attention mechanism proposed in this paper has a high accuracy in predicting the regional economy.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare they have no conflicts of interest to report regarding the present study.

References

[1] X. Chao, G. Kou, Y. Peng and F. E. Alsaadi, "Behavior monitoring methods for trade-based money laundering integrating macro and micro-prudential regulation: A case from China," *Technological and Economic Development of Economy*, vol. 25, no. 6, pp. 1081–1096, 2019.

[2] S. H. Lee, J. H. Kim and J. H. Huh, "Land price forecasting research by macro and micro factors and real estate market utilization plan research by landscape factors: Big data analysis approach," *Symmetry*, vol. 13, no. 4, pp. 616, 2021.

[3] M. P. Clements and D. F. Hendry, "Macro-economic forecasting and modeling," *The Economic Journal*, vol. 105, no. 431, pp. 1001–1013, 1995.

[4] R. Batchelor and P. Dua, "Improving macro-economic forecasts: The role of consumer confidence," *International Journal of Forecasting*, vol. 14, no. 1, pp. 71–81, 1998.

[5] O. A. Lamont, "Macroeconomic forecasts and microeconomic forecasters," *Journal of Economic Behavior & Organization*, vol. 48, no. 3, pp. 265–280, 2002.

[6] R. Rich, J. Bram, A. Haughwout, J. Orr, R. Rosen *et al.,* "Using regional economic indexes to forecast tax bases: Evidence from New York," *Review of Economics and Statistics*, vol. 87, no. 4, pp. 627–634, 2005.

[7] S. Yaseen, S. M. A. Abbas, A. Anjum, T. Saba, A. Khan *et al.,* "Improved generalization for secure data publishing," *IEEE Access*, vol. 6, pp. 27156–27165, 2018.
[8] K. Ketova and I. Rusiak, "Identification and forecast of generalized indicators of regional economic system development," *Applied Econometrics*, vol. 15, no. 3, pp. 56–71, 2009.
[9] H. Shi, "Application of backpropagation (BP) neural network in the marine regional economic forecast," *Journal of Coastal Research*, vol. 98, no. SI, pp. 67–70, 2019.
[10] M. D. Partridge and D. S. Rickman, "Computable general equilibrium (CGE) modeling for regional economic development analysis," *Regional Studies*, vol. 44, no. 10, pp. 1311–1328, 2010.
[11] A. E. Kremin and K. A. Gulin, "Managing the functioning of small business at the municipal level as a driver of regional economic development," *Economic and Social Changes: Facts, Trends, Forecast*, vol. 10, no. 4, pp. 94–110, 2017.
[12] T. Tarasova, O. Usatenko, A. Makurin, V. Ivanenko and A. Cherchata, "Accounting and features of mathematical modeling of the system to forecast cryptocurrency exchange rate," *Accounting*, vol. 6, no. 3, pp. 357–364, 2020.
[13] J. V. Pérez-Rodríguez, S. Torra and J. Andrada-Félix, "STAR and ANN models: Forecasting performance on the Spanish "Ibex-35" stock index," *Journal of Empirical Finance*, vol. 12, no. 3, pp. 490–509, 2005.
[14] J. P. Teixeira and P. O. Fernandes, "Tourism time series forecast-different ANN architectures with time index input," *Procedia Technology*, vol. 5, pp. 445–454, 2012.
[15] J. Chen, J. Wu, H. Liang, S. Mumtaz, J. Li *et al.,* "Collaborative trust blockchain-based unbiased control transfer mechanism for industrial automation," *IEEE Transactions on Industry Applications*, vol. 56, no. 4, pp. 4478–4488, 2019.
[16] D. G. Loomis and J. E. Cox, "A course in economic forecasting: Rationale and content," *The Journal of Economic Education*, vol. 31, no. 4, pp. 349–357, 2000.
[17] J. Bai, "On the partial sums of residuals in autoregressive and moving average models," *Journal of Time Series Analysis*, vol. 14, no. 3, pp. 247–260, 1993.
[18] H. T. Pham and B. S. Yang, "A hybrid of the nonlinear autoregressive model with exogenous input and autoregressive moving average model for long-term machine state forecasting," *Expert Systems with Applications*, vol. 37, no. 4, pp. 3310–3317, 2010.
[19] J. P. Elhorst, M. Gross and E. Tereanu, "Cross-sectional dependence and spillovers in space and time: Where spatial econometrics and global VAR models meet," *Journal of Economic Surveys*, vol. 35, no. 1, pp. 192–226, 2021.
[20] W. Huang, K. K. Lai, Y. Nakamori, S. Wang and L. Yu, "Neural networks in finance and economics forecasting," *International Journal of Information Technology & Decision Making*, vol. 6, no. 1, pp. 113–140, 2007.
[21] H. B. Hwarng, "Insights into neural-network forecasting of time series corresponding to ARMA (p, q) structures," *Omega*, vol. 29, no. 3, pp. 273–289, 2001.
[22] B. M. Al-Maqaleh, A. A. Al-Mansoub and F. N. Al-Badani, "Forecasting using artificial neural network and statistical models," *International Journal Education and Management Engineering*, vol. 3, pp. 20–32, 2006.
[23] Q. Cao, K. B. Leggio and M. J. Schniederjans, "A comparison between Fama and French's model and artificial neural networks in predicting the Chinese stock market," *Computers & Operations Research*, vol. 32, no. 10, pp. 2499–2512, 2005.
[24] R. Zhang, Z. Y. Chen, L. J. Xu and C. Q. Ou, "Meteorological drought forecasting based on a statistical model with machine learning techniques in Shaanxi Province, China," *Science of the Total Environment*, vol. 665, no. 70, pp. 338–346, 2019.
[25] Z. J. Yu, "RBF neural networks optimization algorithm and application on tax forecasting," *TELKOMNIKA Indonesian Journal of Electrical Engineering*, vol. 11, no. 7, pp. 3491–3497, 2013.
[26] M. L. Shen, C. F. Lee, H. H. Liu, P. Y. Chang and C. H. Yang, "An effective hybrid approach for forecasting currency exchange rates," *Sustainability*, vol. 13, no. 5, pp. 2761, 2021.

[27] H. Chung and K. Shin, "Genetic algorithm-optimized multi-channel convolutional neural network for stock market prediction," *Neural Computing and Applications*, vol. 32, no. 12, pp. 7897–7914, 2020.
[28] W. Chen, C. K. Yeo, C. T. Lau and B. S. Lee, "Leveraging social media news to predict stock index movement using RNN-boost," *Data & Knowledge Engineering*, vol. 118, no. 1, pp. 14–24, 2018.
[29] J. Cao, Z. Li and J. Li, "Financial time series forecasting model based on CEEMDAN and LSTM," *Physica A: Statistical Mechanics and Its Applications*, vol. 519, no. 2, pp. 127–139, 2019.
[30] J. Hu and W. Zheng, "A deep learning model to effectively capture mutation information in multivariate time series prediction," *Knowledge-Based Systems*, vol. 203, no. 4, pp. 106139, 2020.
[31] Y. Liu, C. Gong, L. Yang and Y. Chen, "DSTP-RNN: A dual-stage two-phase attention-based recurrent neural network for long-term and multivariate time series prediction," *Expert Systems with Applications*, vol. 143, pp. 113082, 2020.
[32] X. Cai, S. Dong and J. Hu, "A deep learning model incorporating part of speech and self-matching attention for named entity recognition of Chinese electronic medical records," *BMC Medical Informatics and Decision Making*, vol. 19, no. 2, pp. 101–109, 2019.
[33] A. Meng, P. Wang, G. Zhai, C. Zeng, S. Chen *et al.,* "Electricity price forecasting with high penetration of renewable energy using attention-based LSTM network trained by crisscross optimization," *Energy*, vol. 254, pp. 124212, 2022.