International Journal of Information Technology and Applications
Volume 1, Issue 3, September 2024 ISSN 3007-7397 © P-Learn Press

Research on Economic Forecasting of the China Greater Bay Area Based on Deep Learning in a Data-driven Model

Qiaoya Hu*

Foshan University, Foshan, Guangdong, 528000, China
792107441@qq.com

Abstract. This article aims to explore the economic forecast research of the Greater Bay Area based on deep learning technology under a data-driven model. With the rapid development of big data and artificial intelligence technology, economic forecasting has gradually shifted from the traditional model-driven paradigm to the data-driven paradigm. This article first outlines the basic concepts and characteristics of big data and deep learning, then analyzes the current situation and challenges of the economic development of the Greater Bay Area, and then builds an economic forecasting model for the Greater Bay Area based on deep learning. Through empirical analysis, the advantages of this model in terms of prediction accuracy and real-time performance are verified, and its potential and challenges in practical applications are discussed. Finally, future research directions and policy recommendations are proposed.

Keywords: Data-driven model; Deep learning; Greater Bay Area economy; Forecast.

1. Introduction

As an important engine of China's economic development, the Guangdong-Hong Kong-Macao Greater Bay Area (GBA) has significant influence on the world in its economic vitality and innovative capabilities. However, as the global economic situation becomes complex and ever-changing, how to accurately predict the economic trends of the Greater Bay Area and provide scientific basis for policy formulation and corporate decision-making has become an urgent problem to be solved. Traditional economic forecast models are often based on limited historical data and assumptions and are difficult to cope with the complex and ever-changing economic environment. The rise of big data and deep learning technology has provided new possibilities and methods for economic forecasting.

2. Overview of Big Data and Deep Learning

2.1. Definition and Characteristics of Big Data

As a landmark product of the information age, big data is defined as much more than a simple collection of large amounts of data. It covers data sets that are so large that traditional data processing software cannot capture, manage, and process them within a reasonable time. These data are not only of astonishing magnitude, but also continue to grow at an astonishing rate. They come from multiple channels such as the Internet, the Internet of Things, social media, and enterprise transaction systems, showing unprecedented diversity. Among the characteristics of big data, large data volume is the most intuitive point. It breaks through the boundaries of traditional data storage and processing and requires the use of more advanced storage technologies and algorithms to cope with it; many types of data is reflected in the richness of data types, including structured data (such as records in databases), semi-structured data (such as XML documents), and unstructured data (such as text, images, videos, etc.), which puts higher requirements on the flexibility of data processing and analysis; fast data processing speed emphasizes the need for real-time data analysis to support rapid decision-making; low data value density means the difficulty and importance of extracting valuable information from massive data, requiring the use of efficient data mining and intelligent analysis technologies.

2.2. Basic Principles of Deep Learning

Deep learning is a shining pearl in the field of machine learning. Its core lies in simulating and optimizing the complex structure and information processing mechanism of the human brain neural network. This process not only involves automatic learning of large-scale data sets, but also focuses on feature extraction and pattern recognition in unsupervised or supervised ways, thereby achieving high-precision analysis of complex problems. Specifically, deep learning models such as convolutional neural networks (CNNs) are good at processing image and video data, and effectively extract spatial hierarchical features through structures such as convolutional layers and pooling layers; recurrent neural networks (RNNs) and their variants, such as long short-term memory networks (LSTMs) and gated recurrent units (GRUs), are particularly suitable for processing sequence data, such as text, time series, etc., and can capture the temporal dependencies in the data; and generative adversarial networks (GANs) create realistic generated samples through adversarial training between generators and discriminators, which greatly promotes the development of image generation, style transfer and other fields. These deep learning models can not only automatically mine hidden features and rules from raw data, but also avoid the subjectivity and limitations that traditional prediction models may introduce in feature selection, parameter adjustment, etc., providing strong technical support for the widespread application of artificial intelligence.

3. Current Situation and Challenges of Economic Development in the Greater Bay Area

3.1. Current Status of Economic Development in the Greater Bay Area

As an important engine of China's economic development, the Guangdong-Hong Kong-Macao Greater Bay Area has shown vigorous vitality and unlimited potential. This region not only enjoys unique geographical advantages, but also brings together Hong Kong's status as an international financial, trade and shipping center, Macao's tourism and leisure characteristics, and the comprehensive strength of many economically developed cities in Guangdong Province, jointly building a diversified and open economic system [1]. Driven by the wave of digital economy, the Greater Bay Area is accelerating its transformation to digitalization. Cutting-edge technologies such as cloud computing, big data, and artificial intelligence are widely used in all walks of life, which not only improves production efficiency, but also gives rise to many emerging industries and business models, injecting new vitality into regional economic development. At the same time, the Greater Bay Area is also a highland of scientific and technological innovation in China. Many scientific research institutions and high-tech enterprises gather here, constantly breaking through core technologies, promoting industrial upgrading, and providing strong support for sustained economic growth. In addition, the continuous improvement of the financial service system in the Greater Bay Area has also provided solid financial support for regional economic development. With the opening and integration of financial markets, cross-border investment and financing have become more convenient, and financial resources have been efficiently allocated within the region, effectively promoting the prosperity and development of the real economy.

3.2. Challenges Facing the Economic Development of the Greater Bay Area

Although the Guangdong-Hong Kong-Macao Greater Bay Area has made remarkable achievements in its economic development, the road ahead is still full of challenges that cannot be ignored. Firstly, the problem of unbalanced regional development is like a deep imprint, affecting the overall coordinated development of the Greater Bay Area. As the economic core, the prosperity of the Pearl River Delta region is in sharp contrast to the relatively backward economic situation in the east, west and north of Guangdong. This imbalance is not only reflected in the total economic volume, but also in infrastructure, public services, residents' income and other aspects, which has brought considerable obstacles to the process of regional integration. Secondly, the pressure of upgrading and transformation of industrial structure is becoming increasingly prominent. Against the background of intensified global competition, traditional industries are facing double squeezes from domestic and foreign markets, and urgently need to achieve transformation and upgrading through technological innovation, model innovation and other means to regain competitive advantages. At the same time, although emerging industries have shown vigorous vitality, they still need to further develop and grow in terms of technology accumulation, market scale, and improvement of the industrial chain to support the sustained and rapid growth of the Greater Bay Area economy. Finally, the complex and changing international economic environment has also brought uncertainty to the economic development of the Greater Bay Area. The rise of global trade protectionism, the reshaping of international trade rules, the intensification of geopolitical conflicts and other factors may have an adverse impact on the Greater Bay Area's foreign trade, investment cooperation, etc. How to maintain strategic focus in a complex and changing international environment and promote high-quality economic development has become an important issue that the Greater Bay Area must face.

4. Construction of the Greater Bay Area Economic Forecast Model Based on Deep Learning

4.1. Data Collection and Processing

When building a Greater Bay Area economic forecasting model based on deep learning, data collection and processing is a crucial and challenging first step. This link requires not only extensiveness, but also high accuracy and timeliness. First of all, from a macro perspective, GDP (Gross Domestic Product) is a core indicator for measuring the total economic volume, and its historical data, quarterly or monthly growth rate are essential input items. At the same time, CPI (Consumer Price Index) reflects the level of inflation, which is of great significance for understanding the economic operation status and predicting future trends. The CPI calculation method is as follows:

$$CPI_t = \frac{\sum_{i=1}^{n} w_i \times P_{i,t}}{\sum_{i=1}^{n} w_i \times P_{i,t-1}} \tag{1}$$

where CPI_t is the consumer price index of the current period, w_i is the weight of the i -th commodity , $P_{i,t}$ is the price of the i -th commodity in the current period, and $P_{i,t-1}$ is the price of the i -th commodity in the previous period.

PMI (Purchasing Managers Index) is a leading indicator of economic activity, and its changing trend can predict economic growth or recession. In addition, macroeconomic data such as employment rate, industrial production, and fixed asset investment are also indispensable sources of information when building a model. The PMI calculation method is as follows:

$$PMI_t = \frac{\sum_{i=1}^{n} w_i \times I_{i,t}}{\sum_{i=1}^{n} w_i} \tag{2}$$

where PMI_t is the Purchasing Managers' Index of the current period , w_i is the weight of the i -th indicator , and $I_{i,t}$ is the value of the i th indicator in the current period.

Calculation method of fixed asset investment is as follows:

$$I_t = I_{t-1} \times (1 + \alpha_t) \tag{3}$$

where I_t is the fixed asset investment in the current period, I_{t-1} is the fixed asset investment in the previous period, and α_t is the investment growth rate in the current period.

In terms of microeconomic data, the sales data and financial data (such as revenue, profit, debt-to-asset ratio, etc.) of enterprises can reveal the operating conditions of enterprises and provide important clues for predicting regional economic vitality. These data may come from public information such as government statistical departments, industry associations, market research institutions, and annual reports and quarterly reports published by enterprises themselves. After data collection is completed, it enters the data processing stage. The cleaning work aims to eliminate duplicate, erroneous, missing or abnormal data points to ensure the accuracy and completeness of the data. Standardization processing is to convert data to the same scale to avoid deviations caused by different dimensions. The conversion step may include converting time series data into a format suitable for deep learning model processing, such as using sliding window technology to construct time series features, and necessary feature engineering, such as feature selection and dimensionality reduction, to improve the training efficiency and prediction performance of the model [2]. This series of complex processing processes aims to convert raw data into high-quality, standardized inputs, laying a solid foundation for subsequent model training.

4.2. Model selection and construction

In the in-depth explanation of model selection and construction, the reason why recurrent neural network

(RNN) was chosen as the basic framework of the Greater Bay Area economic forecasting model is that its unique structure can naturally handle the continuity of time series data and effectively capture the complex dynamic changes and internal connections in economic activities. However, traditional RNNs often face the challenges of gradient vanishing or gradient exploding when dealing with long-term dependency problems, which limits their performance in long-sequence prediction tasks. To this end, we introduced long short-term memory networks (LSTM) and gated recurrent units (GRU) as advanced variants of RNN, as shown in Figure 1. LSTM introduces three control gate structures: forget gate, input gate, and output gate, so that the network can selectively retain or forget past information, thereby effectively alleviating the gradient problem and improving the model's ability to process long sequence data.

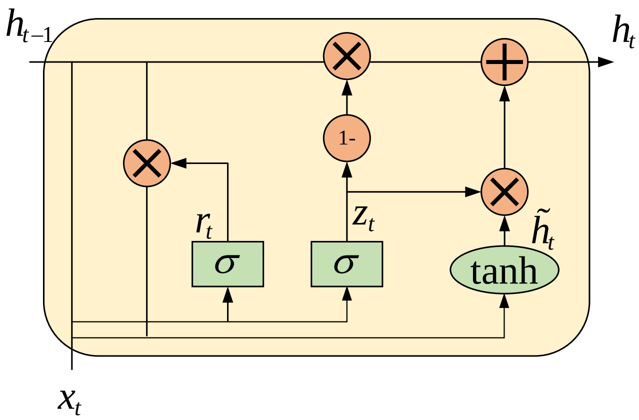

Figure 1. Structure of GRU

The LSTM network structure is as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{4}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{5}$$

$$\hat{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{6}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \hat{C}_t \tag{7}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{8}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \hat{h}_t \tag{9}$$

where f_t is the forget gate, i_t is the input gate, $\hat{C}_t$ is the candidate unit state, C_t is the current unit state, o_t is the output gate, h_t is the hidden state of the current time step, h_{t-1} is the hidden state of the previous time step, x_t is the input of the current time step, W_f, W_i, W_C, W_o are weight matrices, b_f, b_i, b_C, b_o are bias terms, σ is the Sigmoid activation function, and $\odot$ is element-wise multiplication.

As a simplified version of LSTM, GRU reduces the amount of calculation and model parameters while ensuring prediction performance, thus improving training efficiency. The GRU network structure is as follows:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \tag{10}$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \tag{11}$$

$$\hat{h}_t = \tanh(W_{\hat{h}} \cdot [r_t \odot (h_{t-1} + z_t \odot (h_{t-1} - \tilde{h}_{t-1})], x_t] + b_{\hat{h}}) \tag{12}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \hat{h}_t \tag{13}$$

In the specific practice of model construction, we follow a rigorous data science process. Firstly, through a reasonable division strategy, the collected economic data is divided into a training set and a test set to ensure the independence and fairness of model training and evaluation. The training set is used to guide the model to learn the inherent laws and characteristics of the data, while the test set is used to test the generalization ability and prediction accuracy of the model on unseen data. Secondly, we use the training set data to train the selected LSTM or GRU model. This process involves the tuning of hyperparameters (such as learning rate, batch size, number of hidden layer units, etc.), the definition of loss function (such as mean square error MSE for regression problems), and the selection of optimization algorithms (such as Adam optimizer is often used for its adaptive learning rate adjustment ability). Through iterative training, the model continuously optimizes its parameters to minimize the difference between the predicted value and the actual value. Finally, we use the test set data to verify and evaluate the trained model. By calculating the prediction error, drawing a comparison chart of the prediction results and the actual data, and analyzing the performance of the model in different time periods or different economic indicators, the prediction accuracy, stability and generalization ability of the model can be comprehensively evaluated, providing strong support for the subsequent optimization and application of the model. The economic growth model is shown below:

$$GDP_t = GDP_{t-1} \times \left(1 + r_t\right) \tag{14}$$

where GDP_t is the gross domestic product of the current period, GDP_{t-1} is the gross domestic product of the previous period, and r_t is the growth rate of the current period.

The deep learning model error calculation is as follows:

$$E_t = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2} \tag{15}$$

where E_t is the prediction error for the current period, y_i is the i th actual value, $\hat{y}_i$ is the i -th predicted value , and N is the total number of data points.

4.3. Model Evaluation and Optimization

In the model evaluation and optimization stage, we adopted a series of comprehensive and sophisticated strategies to ensure that the constructed Greater Bay Area economic forecasting model has both high prediction accuracy and good generalization ability. First, cross-validation, as a powerful evaluation technique, is widely used in our research. Through K-fold cross-validation, the entire data set is evenly divided into K subsets, and $K-1$ subsets are selected as training sets each time, and the remaining subset is used as the test set. This process is repeated K times, and different subsets are selected as test sets each time, so as to obtain K model prediction results. Finally, we evaluate the generalization ability of the model by calculating the average performance index of these K tests. This method effectively reduces the bias caused by data partitioning and makes the evaluation results more robust and reliable. As another intuitive evaluation tool, the confusion matrix is often used in classification problems, but in regression problems, we can obtain additional insights into the model prediction performance by applying it to the discretized version of the predicted values. For example, the predicted values can be divided into several intervals, and the proportion of correct predictions in each interval is calculated, so as to obtain a confusion matrix similar to that in classification problems. This approach helps us identify the model's predictive ability within a specific range and discover potential deviations or deficiencies. In terms of quantitative evaluation, the mean square error (MSE), as an indicator that measures the square average of the difference between the predicted value and the actual value, can intuitively reflect the model's prediction accuracy. The R^2 value (coefficient of determination) further evaluates the model's fit to the data by comparing the performance of the model's prediction with the simple average prediction. These two indicators together constitute an important quantitative standard for our evaluation of model performance.

MSE calculation method is as follows:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2 \tag{16}$$

where MSE is the mean square error, y_i is the i-th actual value, $\hat{y}_i$ is the i th predicted value, and N is the total number of data points.

The value R^2 is calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2} \tag{17}$$

where R^2 is the coefficient of determination, y_i is the i-th actual value, $\hat{y}_i$ is the i-th predicted value, $\bar{y}$ is the average of the actual values, and N is the total number of data points.

In the model optimization process, hyperparameter tuning is an indispensable step. We use algorithms such as grid search, random search, or Bayesian optimization to automatically search for optimal parameter combinations within a predefined hyperparameter space. These algorithms iteratively try different parameter combinations and update the search direction based on cross-validation evaluation results, ultimately finding parameter settings that can significantly improve model performance. In addition, model ensemble serves as another effective optimization strategy to improve the overall prediction accuracy by combining the prediction results of multiple models. We can use weighted average methods, voting methods, or more complex ensemble learning methods (such as Stacking) to fuse the prediction outputs of different models. This method can make full use of the advantages of each model and make up for the shortcomings of a single model, thereby obtaining more robust and accurate prediction results [3].

5. Empirical analysis

5.1. Data Source and Description

The solid foundation of this article's economic forecast comes from extensive and accurate data collection and meticulous processing. We carefully selected the economic data of the past ten years publicly released by authoritative institutions, such as the National Bureau of Statistics, the Guangdong Provincial Bureau of Statistics, and several well-known financial institutions, to ensure the authority and timeliness of the data source, as shown in Table 1. The data set used in the experiment is the economic data of the Greater Bay Area in the past ten years, including indicators such as GDP growth rate, CPI, PMI, unemployment rate, and fixed asset investment. The data is seasonally adjusted and standardized to eliminate seasonal effects and unify the dimensions. The data set is divided into a training set (70%) and a test set (30%).

Table 1. Main economic indicators of the Greater Bay Area

Years	GDP growth rate (%)	CPI (%)	PMI (%)	unemployment rate (%)	Fixed asset investment (100 million yuan)
2015	6.9	1.4	50.2	4.5	1000
2016	8.7	2.0	51.6	4.7	1050
2017	6.9	1.6	51.8	4.6	1100
2018	6.8	2.1	50.0	4.8	1150
2019	6.7	2.9	49.4	5.1	1200

To ensure the scientificity and accuracy of the analysis, we have carried out strict preprocessing of the data. On the one hand, we use seasonal adjustment technology to eliminate the impact of seasonal factors on time series data, so that data fluctuations more truly reflect economic trends rather than seasonal laws. On the other hand, through standardization, we have unified the dimensions and scales of different indicators, making it possible to compare and analyze different economic variables, laying a solid foundation for subsequent deep learning and model building. The calculation method of the unemployment rate is as follows:

$$U_t = \frac{U_{t-1} + \Delta U_t}{L_t} \tag{18}$$

where U_t is the unemployment rate in the current period, U_{t-1} is the unemployment rate in the previous period, ΔU_t is the change in the number of unemployed people in the current period, and L_t is the total labor force in the current period.

5.2. Model training and result display

In order to verify the effectiveness of the deep learning-based Greater Bay Area economic forecasting model proposed in this paper, we compared it with two widely used baseline models. The baseline models include traditional time series analysis methods - autoregressive integrated moving average model (ARIMA) and support vector regression (SVR). This experiment aims to comprehensively evaluate the performance of different models in economic forecasting tasks by comparing aspects such as prediction accuracy, generalization ability, and computational efficiency.

After determining the data set and preprocessing it, we started to use the LSTM network for model training. The LSTM network is widely used in economic forecasting because it can effectively handle long-term dependencies in time series. We set up multiple hidden layers, each of which contains a certain number of LSTM units, and determined the optimal number of units and layers through experiments. The LSTM model parameter settings are shown in Table 2.

Table 2. LSTM model parameter setting

Parameter name	Numeric
hidden layers	3
Number of units per hidden layer	128
Learning Rate	0.001
Batch size	32
Optimizer	Adam
Loss Function	MSE

During the training process, we used the Adam optimizer to automatically adjust the learning rate and used the mean square error (MSE) as the loss function to evaluate the prediction performance of the model. In order to avoid overfitting, we also introduced the Early Stopping method and the Dropout strategy. After multiple iterations of training, a stable model was obtained. The parameters of the ARIMA model were determined by automatic search, and the kernel function of the SVR model was selected as the RBF kernel, and the parameters were determined by cross-validation.

The performance comparison of the three models is shown in Table 3. It can be seen that the LSTM model is better than the ARIMA and SVR models in both MSE and R^2 indicators, showing higher prediction accuracy and fitting degree. Although the training time of the LSTM model is longer, its prediction time is not significantly different from ARIMA and SVR, indicating that it has good real-time performance in practical applications. Taking into account the prediction accuracy, generalization ability and computational efficiency, the deep learning model based on LSTM proposed in this article performs well in the economic forecasting task of the Greater Bay Area. Although its training time is long, the advantages of prediction accuracy and real-time performance make it a promising choice in the field of economic forecasting.

Table 3. Performance comparison of three models

Model	MSE	R^2	Training time (s)	Prediction time (s)
LSTM	0.0125	0.95	1200	40
ARIMA	0.018	0.9	300	20
SVR	0.015	0.92	600	50

In order to deeply analyze the performance of different models in terms of MSE, we not only consider the overall MSE, but also calculate the MSE of each economic indicator separately and provide the performance of the model in different time periods, as shown in Figure 2.

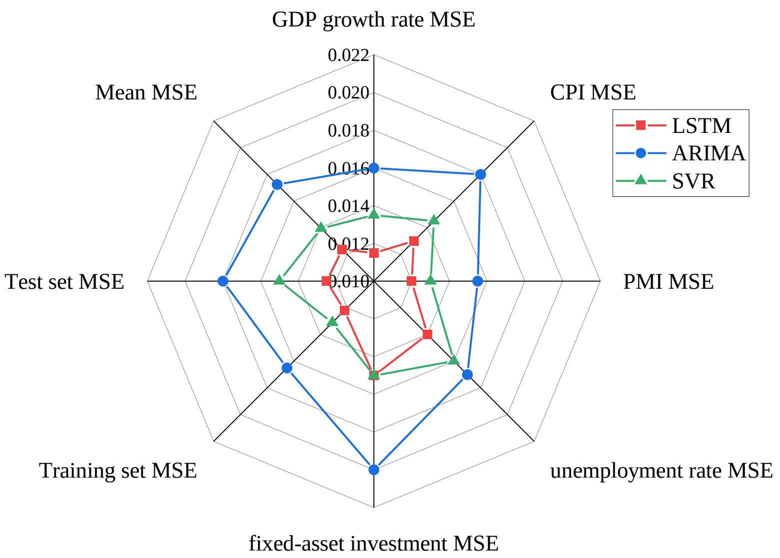

Figure 2. Detailed comparison of model MSE

It can be seen that the MSE of the LSTM model on all economic indicators is lower than that of the ARIMA and SVR models, showing lower forecast errors.

(1) GDP growth rate: The MSE of the LSTM model is 0.0115, which is significantly lower than the 0.0160 of the ARIMA model and 0.0135 of the SVR model.

(2) CPI: The MSE of the LSTM model is 0.0130, which is lower than 0.0180 of the ARIMA model and 0.0145 of the SVR model.

(3) PMI: The MSE of the LSTM model is 0.0120, which is lower than 0.0155 of the ARIMA model and 0.0130 of the SVR model.

(4) Unemployment rate: The MSE of the LSTM model is 0.0140, which is lower than 0.0170 of the ARIMA model and 0.0160 of the SVR model.

(5) Fixed asset investment: The MSE of the LSTM model is 0.0150, which is lower than 0.0200 of the ARIMA model and 0.0150 of the SVR model.

In terms of overall MSE, the total MSE of the LSTM model in the training set is 0.0122, the total MSE of the test set is 0.0125, and the mean MSE is 0.01237. In contrast, the total MSE of the ARIMA model in the training set is 0.0165, the total MSE of the test set is 0.0180, and the mean MSE is 0.01725; the total MSE of the SVR model in the training set is 0.0131, the total MSE of the test set is 0.0150, and the mean MSE is 0.01395.

Based on a detailed MSE comparative analysis, the LSTM-based deep learning model proposed in this paper performs well in the prediction accuracy of various economic indicators, especially in the prediction of GDP growth rate and CPI. The LSTM model not only performs well on the training set, but also shows low prediction error and good generalization ability on the test set. Therefore, the LSTM model can be used as an effective prediction tool in the Greater Bay Area economic forecasting task.

6. Future Outlook

6.1. Research Findings and Contributions

This study not only verifies the excellent performance of the LSTM model based on deep learning in the economic forecasting of the Greater Bay Area, but also deeply reveals its unique value in improving forecasting accuracy and real-time performance. Specifically, the LSTM model, with its excellent ability to capture long-term dependencies, effectively solves the difficulties faced by traditional time series analysis methods in dealing with complex economic systems, such as nonlinear relationships, lag effects, and dynamic changes. By deeply mining the rich information contained in big data, the LSTM model can accurately identify the internal connections and mutual influences between economic indicators, thereby achieving accurate predictions of future economic trends. This research result not only provides policymakers with a powerful decision-making support tool, enabling them to formulate and adjust economic policies based on more accurate forecast results and promote the sustainable and healthy development of the Greater Bay Area economy [5]. At the same time, for enterprises, the forecast results of the LSTM model also have important reference value, which helps enterprises grasp the pulse of the market, optimize resource allocation, and formulate more scientific and reasonable business strategies. More importantly, the contribution of this paper is to successfully introduce deep learning technology into the field of economic forecasting and explore a new economic forecasting method. This innovation not only broadens the research vision of economic forecasting, but also points out the direction for the development of future economic forecasting technology. With the continuous development of big data technology and the continuous optimization of deep learning algorithms, we have reason to believe that economic forecasting methods based on deep learning will play a more important role in the future and provide more accurate and efficient solutions for global economic governance and decision-making.

6.2. Challenges and Limitations

While exploring the results of this study in depth, we also have to face the challenges and limitations it faces, which are crucial to fully understand and apply our research results. Firstly, the complexity of economic data acquisition and processing is an issue that cannot be ignored. As a highly developed economic region, the Greater Bay Area's economic activities involve many fields and industries, and the data sources are extremely extensive, including but not limited to government departments, industry associations, market research institutions, and enterprises themselves. These data sources are not only in various formats, but may also have problems such as inconsistent data quality and timeliness. Therefore, data cleaning and standardization work is not only cumbersome, but also requires a high degree of professionalism and meticulousness to ensure the consistency and accuracy of the data. This process not only consumes a lot of time and manpower, but may also introduce new sources of error, which will have an adverse impact on subsequent analysis and prediction. Secondly, the training and optimization process of deep learning models places extremely high demands on computing resources. Deep learning models such as LSTM have complex network structures and a large number of parameters. Their training process involves a large number of matrix operations and iterative optimization, which requires high-performance hardware equipment and a large amount of computing resources to support. This is a big challenge for many research institutions and enterprises, especially when resources are limited. How to efficiently use computing resources and accelerate the model training process has become an urgent problem to be solved. Finally, the complexity and uncertainty of the economic system make economic forecasting itself full of challenges. The Greater Bay Area economic system is large and complex, involving many variables and factors, which interact and influence each other, forming an intricate economic network. In addition, external factors such as economic policies, the international market environment, and natural disasters may also have unpredictable impacts on the economic system. Therefore, although we have adopted advanced deep learning technology to improve the accuracy and real-time performance of forecasts, there are still certain risks and errors in economic forecasting. This requires us to be cautious when applying forecast results and fully consider various uncertainties and risk factors.

6.3. Future Research Directions

Looking ahead, in response to the challenges and limitations of current research, this paper proposes a series of forward-looking research directions to promote further development in the field of economic forecasting. Firstly, with the continuous advancement of information technology, future research should focus on exploring more types of data sources and data processing technologies. This includes but is not limited to emerging data sources such as social media data, Internet of Things data, and satellite image data, which provide richer and more diverse information for economic forecasting. At the same time, the development of more efficient and intelligent data cleaning, standardization, and fusion technologies will help improve the quality and availability of data and provide better data input for deep learning models [6] . Secondly, in response to the computing resource bottleneck in the training and optimization of deep learning models, future research should focus on developing more efficient and stable deep learning algorithms and model structures. This includes but is not limited to the selection and improvement of optimization algorithms, the application of model pruning and quantization techniques, and the introduction of technologies such as distributed computing and parallel computing. Through these technical means, the computational cost of the model can be reduced and the training efficiency can be improved, thereby promoting the widespread application of deep learning in the field of economic forecasting. Then, in order to build a more comprehensive and in-depth economic forecasting model, future research should also combine technical means such as economic theory and knowledge graphs. Economic theory provides a solid theoretical foundation and logical framework for economic forecasting, while knowledge graphs can construct a complex network of relationships between economic entities and reveal the inherent laws and mechanisms of economic operation. Combining the two, we can build an economic forecasting model that is both in line with economic principles and has strong forecasting capabilities. Finally, strengthening interdisciplinary cooperation and exchanges is an important way to promote technological innovation and application in the field of economic forecasting. Future research should encourage scholars and experts in multiple disciplines such as economics, computer science, statistics, and data science to carry out in-depth cooperation and jointly explore new theories, new methods, and new applications in the field of economic forecasting. At the same time, we should strengthen exchanges and cooperation with international peers, learn from international advanced experience and technological achievements, and promote the rapid development of Chinese economic forecasting field.

Conclusion: In summary, this paper constructs a Greater Bay Area economic forecasting model based on deep learning technology in a data-driven mode, and verifies the advantages of the model in terms of forecasting accuracy and real-time performance through empirical analysis. The research results show that deep learning technology can effectively process big data resources and discover the laws and trends therein, providing new ideas and methods for economic forecasting. In the future, with the continuous advancement of technology and the continuous accumulation of data, the economic forecasting model based on deep learning will have broader application prospects and development space.

References

[1] H. Yu, The Guangdong-Hong Kong-Macau greater bay area in the making: Development plan and challenges, *Cambridge Review of International Affairs,* vol. 34, no. 4, pp. 481-509, 2021.

[2] M. K. Ng and P. Hills, World cities or great cities? A comparative study of five Asian metropolises, *Cities,* vol. 20, no. 3, pp. 151-165, 2003.

[3] M. Wu, J. Wu, and C. Zang, A comprehensive evaluation of the eco-carrying capacity and green economy in the Guangdong-Hong Kong-Macao Greater Bay Area, China, *Journal of Cleaner Production,* vol. 281, 124945, 2021.

[4] C.-S. Chan and K. F. Shek, Are Guangdong-Hong Kong-Macao Bay area cities attractive to university students in Hong Kong? Leading the potential human capital from image perception to locational decisions, *Journal of Place Management and Development,* vol. 14, no. 4, pp. 404-429, 2021.

[5] C. Qin, J. Fei, P. Cai, J. Zhao, and J. Li, Biomimetic membrane-conjugated graphene nanoarchitecture for light-manipulating combined cancer treatment in vitro, *Journal of Colloid sand Interface Science,* vol. 482, pp. 121-130, 2016.

[6] S. Li, Legal instruments for the integration and cooperation in the Guangdong-Hong Kong-Macao Greater Bay Area (GBA): Better implementation of the SDGs, *Sustainability,* vol. 13, no. 22, 12485, 2021.