# Research on LSTM-Based Industrial Added Value Prediction Under the Framework of Federated Learning

Pan Hu, Jun Qi(✉), Jue Bo, Yu Xia, Chuan-Ming Jiao, and Meng-Tong Huang

Information Communication Branch, State Grid Liaoning Electric Power Co., Ltd., No. 18 Ningbo Road, Heping District, Liaoning 110060, Shenyang, China
{hp,qj,bj,xiayu,jiaocm,hmt}@ln.sgcc.com.cn

**Abstract.** Industrial added value is an important indicator to measure the performance of the real economy. Scientifically predicting industrial added value helps the government to understand the economic situation in a timely and accurate manner, and to formulate practical and reliable economic policies. Existing research shows that the industrial value-added data from the Department of Industry and Information is related to the macroeconomic data within the government and the power big data within the State Grid, and the three parties are independent of each other, data is heterogeneous, and there is a "data island" problem. Therefore, this paper takes Liaoning Province as an example, under the framework of federated learning, using Long Short-Term Memory (LSTM) network to predict the industrial added value data of Liaoning Province. The results show that the method proposed in this study can effectively protect the privacy security between institutions, fully integrate and utilize the data value of each institution, and can more accurately predict the industrial added value.

**Keywords:** Federated learning · Industrial added value · LSTM

## 1 Introduction

The industrial added value is the final result of the industrial production activities of the industrial enterprises in the form of currency during the reporting period, reflecting the newly created value in the production process of the enterprises, an important indicator for monitoring the operation of the macro economy, and an important basis for implementing macro-control. Therefore, the prediction of industrial added value has become a very important subject. Timely and accurate prediction of industrial added value is the basis for the country to formulate reliable economic policies, and has important practical significance for individual enterprises to grasp the economic situation and understand the level of social development more quickly.

Domestic and foreign scholars have carried out a small amount of research on the prediction of industrial added value. Existing literature shows that the industrial added value data from the Department of Industry and Information is related to the macroeconomic data within the government and the power big data within the State Grid. For


Z. Zhan et al. (Eds.): ICBDIE 2022, AHCS 5, pp. 426–434, 2023.
https://doi.org/10.2991/978-94-6463-034-3_44

example, the industrial climate index is used to predict the industrial added value [6], the credit spread index of state-owned enterprises is constructed to predict the industrial added value [5, 9], and the industrial added value is predicted based on the consumer price index, value-added tax, etc. and electricity purchase data of different industries [2–4]. However, due to the independence of the Ministry of Industry and Information Technology, the government, and the State Grid, the data of the three parties is heterogeneous, contains a large amount of private information, and cannot be shared, thus forming a data island, resulting in a large amount of valuable data cannot be effectively mined and utilized. Therefore, a new method is urgently needed to fully integrate and utilize the data in each isolated island to assist the prediction of industrial added value under the premise of protecting data privacy.

In recent years, with the rapid development of artificial intelligence, federated learning has emerged as the times require. It is different from centralizing all data to a central server for modeling in the past. On the premise of protecting data privacy and security, parameters are uploaded by each data owner and aggregated by the central server, thereby realizing local training and global optimization of the model. Therefore, this paper takes the historical data of industrial added value in Liaoning Province as an example. Under the framework of federated learning, the long short-term memory network (LSTM) model is used to mine the macroeconomic data within the Liaoning Provincial People's Government and State Grid-Liaoning Electric Power Co., Ltd.'s monthly power data for different industries, and then forecast industrial added value. The results show that the method proposed in this study can effectively solve the data island problem formed by the protection of data privacy among the three parties, realize the effective mining and full utilization of the data of different institutions, and can more accurately predict the industrial added value.

The rest of the paper is structured as follows: Sect. 2 presents a long-short-term memory network-based industrial value-added prediction method under the federated learning framework. Section 3 conducts experimental tests, and the results show that the prediction method proposed in this paper performs well. Section 4 summarizes the main conclusions and future research prospects.

## 2 Model and Research Framework

### 2.1 LSTM Model

The long short-term memory network (LSTM) model is a special recurrent neural network (RNN) proposed by Hochreiter and Schmidhuber (1997), which can effectively solve the problems of gradient disappearance and gradient explosion in the long sequence training process of RNN. The structure diagram of its basic unit is shown in Fig. 1. There are three different gated structures, including forget gate, input gate and output gate.

(1) Forget gate. The forget gate in LSTM determines how much of the unit state at the previous moment needs to be retained to the current moment, as shown by ① in Fig. 1. The calculation formula is as follows:

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

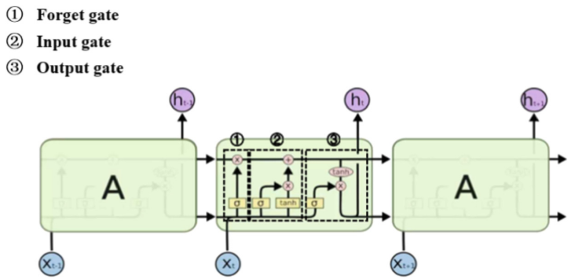


**Fig. 1.** LSTM basic unit structure diagram.

where $h_{t-1}$ is the output of the hidden layer at the previous moment, $X_t$ is the input at the current moment, $W_f$ is the $[h_{t-1}, X_t]$ weight matrix to be mapped to the forget gate, $b_f$ is the bias vector, σ is the neural network layer whose activation function is sigmoid, and its output value is 0–1 which determines how much state information is retained.

(2) Input gate. The input gate in LSTM determines how much of the input data of the network needs to be saved to the unit state at the current moment, as shown by ② in Fig. 1, the calculation formula is as follows:

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

$$\tilde{C}_t = \tanh(W_c * [h_{t-1}, X_t] + b_c) \tag{3}$$

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

Where $i_t$ represents how much of the input data of the network needs to be updated at the current moment, $\tilde{C}_t$ is the candidate unit state of the current input. By multiplying the state at the previous moment $C_{t-1}$ by the forgetting factor $f_t$, multiplying the candidate state at the current moment $\tilde{C}_t$ by $i_t$, and then adding the two product values to obtain the current unit state $C_t$, thus both preserving long-term memory and removing redundancy.

(3) Output gate. The output gate in LSTM determines the information to be output according to the unit state, as shown by ③ in Figure 1. The calculation formula is as follows:

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

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

Where $o_t$ represents the part of the current unit state that needs to be output, and then the current unit state $C_t$ is processed by the tanh layer, and the two are multiplied to obtain the final information to be output $h_t$.

## 2.2 Federated Learning Framework

In order to solve the data island problem among the three agencies of the Ministry of Industry and Information Technology, the government, and the State Grid, this paper

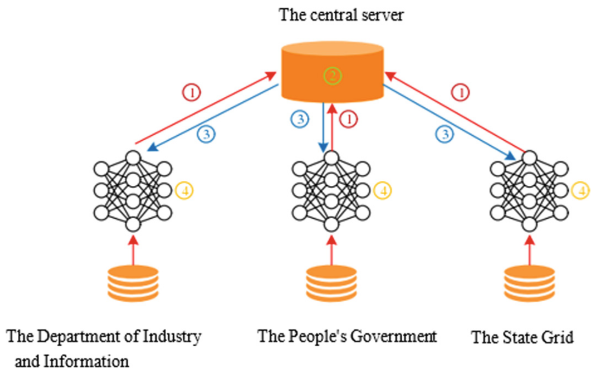


**Fig. 2.** Federated Learning System Architecture Diagram.

builds a horizontal federated learning system framework as shown in Fig. 2 to achieve complete protection of private data during model training.

In the system framework of federated learning, there are three clients and one central server. Each client uses local private data for model training, does not communicate with each other and does not transmit data to the server. The central server is mainly responsible for aggregating the model parameters trained by each client, updating the model, and feeding back the latest model parameters to each client. The specific workflow is as follows:

1) The Department of Industry and Information, the People's Government and the State Grid use their own data sets to train the prediction model, and transmit the obtained model gradients and parameters to the intermediate server through homomorphic encryption technology;
2) The central server aggregates the passed model parameters and model gradients using the federated average algorithm without contacting the data of each client to form an intermediate model;
3) The central server transmits the updated model parameters after aggregation to each client;
4) Each client decrypts the received gradient and uses the decrypted gradient result to update their respective model parameters.

Iterate through the above steps until the loss function of the central server converges, thus completing the entire training process.

## 3 Experimental Test

In this section, we first propose the research framework of this paper, and then introduce the corresponding methods.

### 3.1 Data Sources

The data used in this article come from three institutions: the Department of Industry and Information Technology of Liaoning Province, the People's Government of Liaoning Province, and the State Grid-Liaoning Electric Power Co., Ltd. All variables use the year-on-year growth rate, and the time span is from February 2016 to September 2021.

Among them, the data of the Department of Industry and Information Technology is the historical data of industrial added value in Liaoning Province; the government data is macroeconomic data, including a total of 12 economic variables such as consumer price index and value-added tax; State Grid's data are monthly electricity data for major industries, including a total of 6 economic variables such as Electricity consumption in petroleum and other fuel processing industries, Electricity consumption of non-metallic mineral products industry, Electricity consumption of chemical raw materials and chemical products manufacturing, Electricity consumption of ferrous metal smelting and rolling industry, Electricity consumption of non-ferrous metal smelting and rolling processing industry and the overall industrial electricity consumption in Liaoning Province.

### 3.2 Data Preprocessing

#### 3.2.1 Stationarity Test

In order to ensure the stationarity of variables, the above 19 variables are tested for stationarity. If the test result shows that the time series is non-stationary, take the difference and test its stationarity again; if the variables are still non-stationary after differencing, perform the difference again until all variables are stationary.

#### 3.2.2 Data Normalization

The data used in this paper come from three independent agencies, the Department of Industry and Information Technology, the People's Government and the State Grid. In this paper, the maximum and minimum normalization method is used, and the calculation formula is as follows:

$$\widehat{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{7}$$

Where x is the original data, $x_{min}$ is the minimum value of the original data, $x_{max}$ is the maximum value of the original data, and $\hat{x}$ is the normalized data.

#### 3.2.3 Feature Screening

This paper intends to use the economic variables of the people's government, the power variables of the State Grid and the historical data of industrial added value from the Ministry of Industry and Information Technology to predict the industrial added value under the framework of federated learning. From Sect. 3.1, it can be seen that there are many features used in this paper, the calculation is more difficult, and the model is more complex, so feature engineering is needed to screen out the least and best independent variables.

Therefore, this paper adopts the Filter filtering method, that is, each feature is scored according to the correlation, and the feature selection is performed by setting a threshold or the number of thresholds to be selected. There are 9 independent variables finally selected, The growth rate of electricity consumption in petroleum and other fuel processing industries, the growth rate of electricity consumption in non-ferrous metal smelting and rolling processing industries, the growth rate of industrial electricity consumption in Liaoning Province, and the growth rate of industrial added value in Liaoning Province lag by one period.

### 3.3 Performance Metrics and Model Parameters

In order to test the accuracy of the model prediction, this paper selects the following indicators to measure the prediction effect of the prediction model: root mean square error (RMSE), coefficient of determination ($R^2$). The calculation methods of each method are as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{8}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \tag{9}$$

Where $y_i$ represents the actual value of the forecast target, $\widehat{y_i}$ represents the forecast value of the forecast target, $\overline{y}$ represents the average value of the actual value of the forecast target, and n is the number of samples contained in the forecast set. The smaller the RMSE value, the larger the value of $R^2$, which means that the fitting effect of the model is better and the prediction accuracy is higher.

In the datasets for all experiments, the training, validation, and test sets are split in a ratio of 7:2:1. Set the batch-size to 10, the epoch of the training period to 50, and use the BayesSearchCV optimizer to train the optimal model. The best model consists of 7 layers, an input layer, 5 LSTM layers and 1 fully connected layer. The specific performance is shown in Fig. 3.

### 3.4 Result Analysis

The evaluation index results of this paper are RMSE = 4.05, $R^2 = 0.66$. The predicted value of industrial added value based on LSTM under the federated learning framework proposed in this paper is shown in Fig. 4.

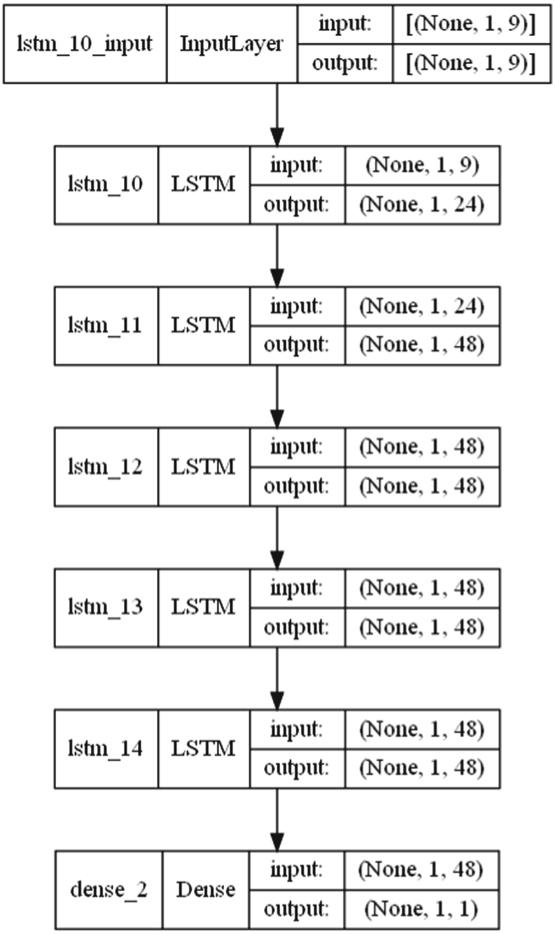


**Fig. 3.** Best Model Structure Diagram.

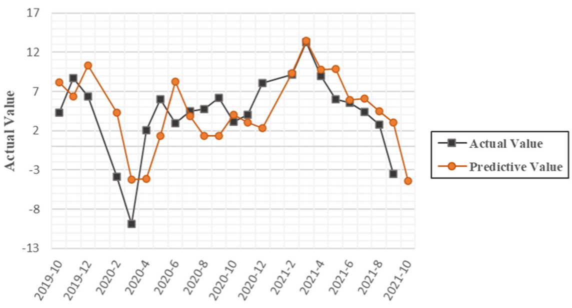


**Fig. 4.** Prediction result graph.

## 4 Conclusions

This study proposes a new method of industrial added value prediction based on LSTM under the framework of federated learning. Taking Liaoning Province as an example, under the premise of protecting privacy, the data existing in the three-party island of Liaoning Provincial Department of Industry and Information Technology, Liaoning Provincial People's Government, and State Grid-Liaoning Provincial Electric Power Co., Ltd. are effectively used. The industrial added value of Liaoning Province from October 2019 to October 2021 is predicted, and the root mean square error and determination coefficient of the model are calculated. The results show that the prediction method under the proposed federated learning framework has RMSE = 4.05 and R2 = 0.66. This method realizes the intelligent and accurate prediction of industrial added value and the fusion and utilization of multi-source heterogeneous data, which is of great significance for the government to formulate economic policies and industrial enterprises to make production decisions.

Future research work can also be expanded in the following two aspects:

(1) Indicator filtering. The experimental research in this paper uses 12 macroeconomic variables, 6 industry power variables and historical data of industrial added value. There may also be other factors related to industrial added value. Therefore, more effective features should be constructed and selected in the future.
(2) Model application. This paper mainly studies the forecasting of industrial added value. The proposed model can also be applied to other time series forecasting problems, such as stock price, power load, etc. Therefore, testing and improving the universality of the model is an important research direction in the future.

**Acknowledgements.** This work was supported by Science and Technology Project of State Grid Liaoning Electric Power Supply Company Limited under Grant No. 2021YF-84.

## References

1. Moradzadeh, H. Moayyed, B. Mohammadi-Ivatloo, A. P. Aguiar and A. Anvari-Moghaddam, "A Secure Federated Deep Learning-Based Approach for Heating Load Demand Forecasting in Building Environment," in IEEE Access, vol. 10, pp. 5037-5050, 2022, doi: https://doi.org/10.1109/ACCESS.2021.3139529.
2. Chen H B, Pei L L, Zhao Y F. Forecasting seasonal variations in electricity consumption and electricity usage efficiency of industrial sectors using a grey modeling approach [J]. Energy, 2021, 222: 119952.
3. Dilaver Z, Hunt L C . Industrial electricity demand for Turkey: A structural time series analysis [J]. Energy Economics, 2011, 33(3): 426-436.
4. Fang PENG, Gaoqun PENG, Yaru QI, Tiantian LIU, Xiaolei ZHOU. Nowcasting of China's industrial added value based on electric power big data [J]. Telecommunications Science, 2021, 37(7): 115-125.
5. Guangtong G U, Bing X U. Half-month forecast of China's industrial added value: Based on the macro monthly data [J]. Systems Engineering-Theory & Practice, 2018.

6. He Y, Tan J, Liu Y. Study on Combination Forecasting Model of the Industrial Added Value Combined with PMI and Prosperity Index [J]. Advances in Intelligent Systems and Computing, 2015.
7. Li, J.; Cui, T.; Yang, K.; Yuan, R.; He, L.; Li, M. Demand Forecasting of E-Commerce Enterprises Based on Horizontal Federated Learning from the Perspective of Sustainable Development. Sustainability 2021,13, 13050.
8. Lei, Kuang, Chengyu, et al. Prediction to Industrial Added Value Based on Holt-Winters Model and ARIMA Model [C].
9. Yang, Y., Kong, J., Yang, L., et al. Sequential Big Data-Based Macroeconomic Forecast for Industrial Value Added. Commun. Math. Stat. 7, 445-457 (2019).