Method Article

LSTM Pig Price Prediction Model based on Attention Mechanism

Abstract: In many agricultural products, pig price fluctuation has a significant impact on price level and consumer price index, so accurate prediction of pig price is of great significance for pig market research and production. In order to predict the price of pigs more accurately in the short term, a short-term memory neural network (Attention-LSTM) prediction model based on attention mechanism is established. The results show that: compared with the traditional LSTM forecasting model, the Attention-LSTM model has higher prediction accuracy, and this model has a good effect on the short-term prediction of pig prices.

Keywords: Pig price Prediction Attention mechanism Attention-LSTM

Introduction

"Food is the first, grain is safe all over the world". As one of the main bulks traded agricultural products in China, live pigs are the necessities of life for residents' consumption. In recent years, affected by Infection with African swine fever virus and Corona Virus Disease 2019, pig prices fluctuated abnormally, which greatly affected the producers, operators, and national consumers of the pig industry.

Ren Qingshan [1] achieved better prediction results when using BP neural network to predict pig prices; RembiszW [2] used production efficiency and the relationship between pig prices and feed prices to study and analyze the profit coefficient of pig production; MolinaIR et al. [3] used X-12-ARIMA model to study Philippine pig prices and found that the price prediction effect was better after excluding seasonal fluctuations. Fu Lianlian [4] made an early warning analysis of pig price fluctuations; Luo Chuangguo [5] constructed an ARIMA model and made a short-term forecast of pig prices; Ernst et al. [6] found that there are multi-scale periodic characteristics in pork prices through the analysis of historical pork price data. Ling et al. [7] in order to solve the problem of information lag in the process of pig price prediction, a pig price prediction framework combined with online search technology is proposed, and the decomposition technology is combined to remove the noise in the online search text, so that the model has better performance in the medium-term and long-term prediction of pig price.Ye K, Piao Y, Zhao K, et al [8] propose Heterogeneous Graph-enhanced LSTM.ARIMA model is used by researchers to forecast the price of agricultural products [9-14].

Although scholars have achieved good expected results by using various models, they do not consider the correlation between input characteristics and prediction results, and not all input features are positively correlated with pig price prediction results, unrelated inputs often lead to a decline in prediction accuracy. The attention mechanism can overcome this challenge, and the attention mechanism allows the model to obtain key information from all inputs.

1. Factors affecting the price of live pigs

As can be seen from the pig breeding cycle in Fig. 1, the price of live pigs is directly determined by the price of piglets, and the price of piglets depends on the number of sows that can reproduce. In addition, the main cost of the whole pig farming industry is pig feed, while the main components of pig feed are corn and soybean meal.

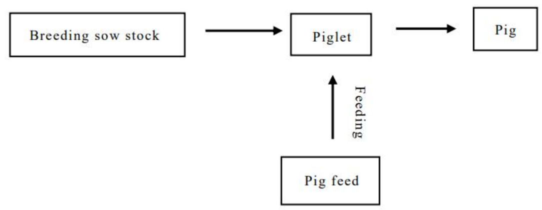

Fig. 1 Pig breeding cycle

Pig cycle is an economic phenomenon in which pig prices fluctuate periodically with the market economy. The specific period is shown in Fig. 2:

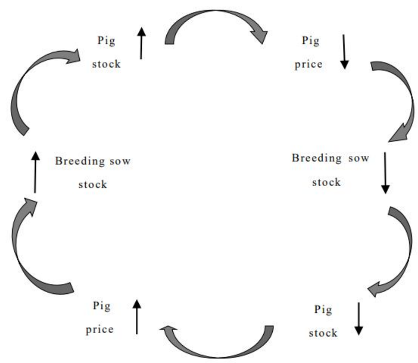

Fig. 2 Pig cycle

In summary, Pig price, Piglet price, Soybean meal price, Corn price, CPI (Consumer Price Index), Pig stock and Breeding sow stock were used as input variables in the experiment.

2. LSTM Pig Price Forecast Model based on Attention Mechanism

2.1 LSTM

LSTM (Long-term and short-term memory network) is a kind of time-cyclic neural network, which is specially designed to solve the long-term dependence problem of general RNN (cyclic neural network). For time series, the proposal of RNN makes up for the disadvantage that the neural network cannot capture the time characteristics of the time series, but the RNN model is found to be unable to deal with the gradient problem well in multiple reverse conduction, and the gradient is easy to disappear or explode, resulting in the gradient cannot be effectively updated [15], while the proposal of LSTM makes up for its deficiency.

The core of LSTM lies in the addition of three gates and a memory element, as shown in Fig. 3, the horizontal line running through the top is called cell state, and the cell state is like a conveyor belt with only a little linear interaction, along which information can flow and remain unchanged to solve the problem of insufficient learning of long-distance information [16].

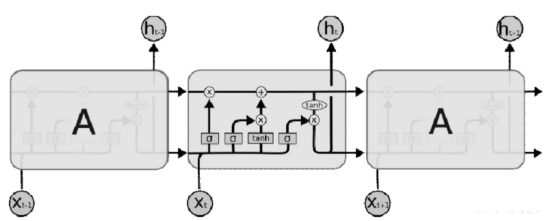

Fig. 3 LSTM structure diagram

The three different doors are the forgetting door, the input door and the output door. The forgetting gate chooses to selectively forget the information in the previous step of the cell state, the input gate decides to selectively record the new information to the cell state, and the output determines the information output of the cell state. The calculation formulas of forgetting door, input door and output door are as follows:

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

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

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{3}$$

In the formula, W_f,W_i,W_o is the weight matrix of the forgetting gate, the input gate and the output gate, respectively; $[h_{t-1}, x_t]$ means to connect these two vectors into a longer vector; b_f,b_i,b_o is the offset vector of the amnesia gate, the input gate and the output gate, respectively; σ is the Sigmoid activation function, and f_t,i_t,o_t is the state, cell state and output of each gate at t time.

The tanh layer creates a candidate vector $\widetilde{C_t}$ candidate, which will be added to the cell state. We will combine these two to update the cell state.

$$\widetilde{C_t} = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{4}$$

$$C_t = f_t * C_{t-1} + i_t * \widetilde{C_t} \tag{5}$$

Where tanh is a hyperbolic tangent activation function and * is a Hadamard product.

Finally, the output h_t of the current time of the LSTM unit is:

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

2.2 Attention mechanism

Attention mechanism is essentially a screening mechanism, which allows us to quickly filter out valuable information from a large amount of information. By giving different weights to the input characteristics of the model, Attention highlights the more critical factors and helps the model to make a more accurate judgment without increasing the extra overhead of the model.

We can look at the Attention mechanism this way (see Fig. 4): imagine that the constituent elements in Source are made up of a series of < Key, Value > data pairs. Given an element Query, by calculating the correlation between Query and each Key, the weight coefficient of each Key corresponding to Value is obtained, and then the Query output is obtained by weighted summing with the corresponding Value, and the final Attention value is obtained. So, the core of Attention mechanism is the weighted summation of the value of elements in Source, while Query and Key are used to calculate the weight coefficient of the corresponding Value.

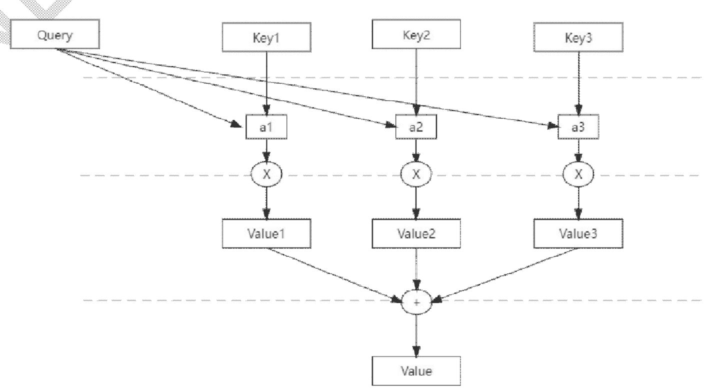

Fig. 4 essential idea of Attention mechanism

2.3 Attention-LSTM model

The LSTM pig price prediction model combined with Attention mechanism can judge the importance of information at each input time, and the training efficiency of the model can be improved. The Attention mechanism gives different weights to the input characteristics of the LSTM, highlights the key influencing factors, and helps LSTM to make an accurate judgment, so the Attention mechanism is introduced into the LSTM model to effectively highlight the factors that affect the pig price, so as to improve the forecasting effect.

Attention-LSTM includes Input layer, LSTM layer, Attention layer and Output layer. The input feature attribute starts from the input layer, passes through the LSTM layer, and outputs the processed vector. As the input of the Attention layer, the Attention layer calculates the weight vector according to the input vector value of the current layer, and then combines the weight vector with the current input layer vector to get a new vector input to the full connection layer, and finally calculates the predicted value. The Attention-LSTM structure is shown in Fig. 5:

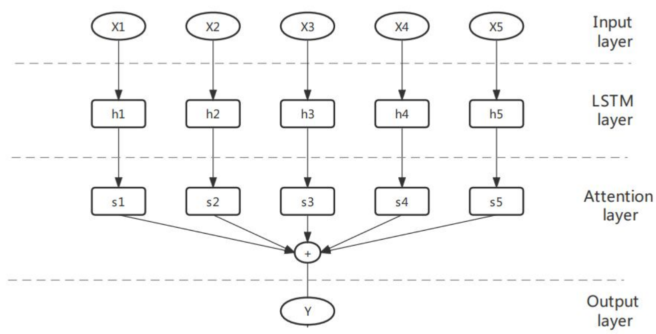

Fig. 5 Attention-LSTM structure diagram

3. Empirical analysis

3.1 Data source and preprocessing

The original data of the study came from the Wind database (financial data and analysis tools service provider), National Bureau of Statistics (http://www.stats.gov.cn/) and Prospective database. Selected Pig price, Piglet price, Soybean meal price, Corn price, CPI, Pig stock and Breeding sow stock from January 2009 to December 2021. Since Pig stock, Sow stock and CPI are much larger than the price value, the logarithm of the three data is chosen in order to reduce the heteroscedasticity of the sample and increase the stationarity. Namely: Pig price (ppig), Piglet price (ppiglet), Soybean meal price (psmp), Corn price (pcorn), CPI (lnCPI), Pig stock (lnspig), Breeding sow stock (lnssow).

The experiment selects the first 65% of the data as the training set and the last 35% as the test set. First, the data is normalized. Normalization refers to a way to simplify the calculation. Different evaluation indicators often have different dimensions and dimensional units. This situation will affect the results of data analysis. In order to eliminate the dimensional impact between indicators, data standardization processing is needed to solve the comparability between data indicators. After processing, the data is limited to a certain range, and after normalization calculation, the data can be limited to $[0,1]$. Data normalization can accelerate the fast convergence of the algorithm, and it is more convenient in subsequent data processing. The formula is as follows (7):

$$X^* = \frac{X - \min}{\max - \min} \tag{7}$$

Max is the maximum value of sample data and min is the minimum value of sample data.

3.2 Test environment

All tests were conducted on Lenovo laptops, and the hardware environment configuration is shown in the following Table 1:

Table 1 Hardware environment configuration

Configuration Item	Parameter
CPU	11th Gen Intel(R) Core (TM) i5-1135G7
RAM	16G
Video card	NVIDIA® GeForce MX450

This test was conducted in python (version number 3.8) under Windows operating system, and data processing and model building were conducted in PyCharm (version number 2019.3.3), the professional version. The library and version number of machine learning used are shown in Table 2:

Table 2 Machine learning library version number

Configuration Item	Parameter
Tensorflow	2.8.0
Numpy	1.22.3
Pandas	1.4.1
Keras	2.8.0
Sklearn	1.0.2

3.3 Evaluation index

Considering the characteristics and limitations of the evaluation criteria in the field of practical application, a single evaluation index is difficult to comprehensively measure the training results of the model [17]. In the experiment, two evaluation indexes are used to test the training effect of the model, that is MAE (Mean Absolute Error) and RMSE (Root Mean Square Error) to show the pig price prediction error of different methods, RMSE uses the average error, MAPE is based on the principle of relative error, the two errors take into account different factors, so the experiment combines the two as one of the bases to judge the advantages and disadvantages of the model, the smaller the measurement number of these evaluation indicators, the higher the average prediction accuracy of the model. As shown in formulas (8) and (9).

$$\mathrm{MAE}(X, h) = \frac{1}{m}\sum_{i=1}^{m}|h(x_i) - y_i| \qquad (8)$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2} \qquad (9)$$

Among them, $\hat{y} = \{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n\}$ is the predicted value, which is equal to $y = \{y_1, y_2, \dots, y_n\}$ is the true value.

3.4 Comparative test and analysis

In the process of the comparative test, the parameters of the comparative test are corrected, and the optimal parameters of the comparative model are selected.

Modeling according to a single variable as an input cannot reflect the influence of other factors on the response variable, and the factor affecting the pig price is not only the pig price itself, so comprehensively considering the idea of multivariable and multi-factor, we choose the multivariable LSTM forecasting model, which will also take the main factors affecting the pig price as input, which is more accurate than the single variable forecasting model.

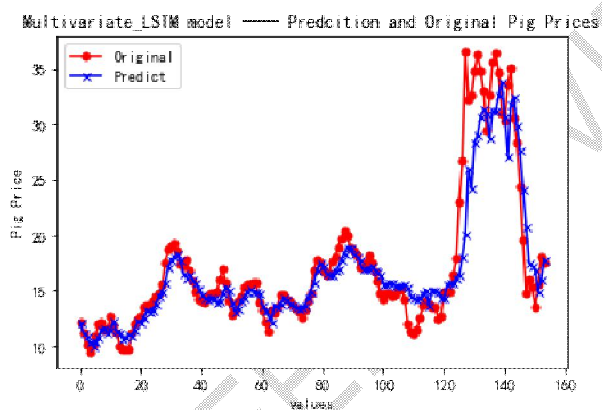

Fig. 6 Multivariable_LSTM model

Based on the multivariable forecasting model, the Attention mechanism is introduced into the model to effectively highlight the factors that affect the pig price, so as to improve the forecasting effect.

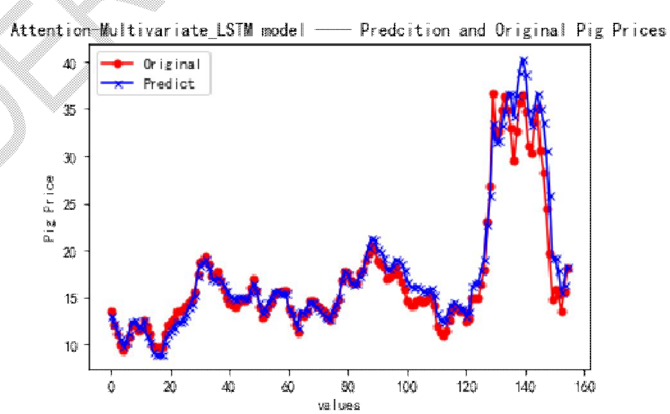

Fig. 7 Attention- Multivariable_LSTM model

Observing Fig. 6-7, we can see that the multivariable prediction model with attention mechanism in the two prediction models performs better. In order to get more accurate results, RMSE and MAE are used to quantify the prediction accuracy of the model. Table 3 shows the prediction performance of traditional LSTM multivariable prediction model and LSTM multivariable prediction model with Attention mechanism on pig price data sets.

Table 3 average performance of the prediction model

	Evaluation index			
	Training set		Test set	
Model	MAE	RMSE	MAE	RMSE
Multivariable_LSTM model	0.70	0.88	2.93	4.10
Attention- Multivariable_LSTM model	0.50	0.64	2.01	2.53

From the evaluation indicators in Table 1, we can see that the prediction performance of Attention-LSTM is the best of the two models, indicating that the attention mechanism plays a positive role in the prediction of LSTM model.

4. Conclusions

In order to solve the problem of many factors affecting pig price, a short-term price forecasting model based on Attention-LSTM network is established by making use of the time series characteristics of pig price, combining price factors, stock factors, household consumption index and historical pig price data, and different weights are given to the input characteristics by using Attention mechanism to highlight the characteristics that play a key role in pig price prediction. The experimental results show that the accuracy of LSTM prediction model with attention mechanism is higher than that of traditional LSTM prediction model.

References

[1] Ren Qingshan, Fang Kui, Zhu Xinghui. Pig price prediction model based on BP neural network based on multiple regression [J]. Jiangsu Agricultural Sciences, 2019, 47 (14): 277281.

[2] Rembisz W, Zawadzka D. Profitability coefficient of pig production–analytical and empirical analysis for the period between 2001 and 2016[J]. Zagadnienia Ekonomiki Rolnej, 2017 (3 (352)): 116-132.

[3] Molina I R, Reyes J A D, Gordoncillo P U. Analysis of seasonality in monthly pork prices in the Philippines based on X-12 ARIMA[J]. Journal of International Society for Southeast Asian Agricultural Sciences, 2017, 23(2): 215-226.

[4] Fu Lianlian, Weng Zhenlin, Zhang Yayan. The cause and early warning Analysis of Pig Price fluctuation in Jiangxi Province-- based on Grey relational and LS-SVM Model [J]. Acta Agriculturae Zhejiangensis, 2016, 28 (9): 1624-1630.

[5] Luo Chuangguo, Zhang Meizhen, Xue Jiliang. Short-term Forecast of Chinese Pig Price based on ARIMA Model [J]. World Agriculture, 2010 (10): 45-48.

[6] Berg E, Huffaker R. Economic dynamics of the German hog-price cycle[J]. International Journal on Food System Dynamics, 2015, 6(2): 64-80.

[7] Ling L, Zhang D, Chen S, et al. Can online search data improve the forecast accuracy of pork price in China? [J]. Journal of Forecasting, 2020, 39(4): 671-686.

[8] Darekar A, Reddy A A. Cotton price forecasting in major producing states[J]. Economic Affairs, 2017, 62(3): 373-378.

[9] Jadhav V, CHINNAPPA R B V, Gaddi G M. Application of ARIMA model for forecasting agricultural prices[J]. 2017.

[10] Pardhi R, Singh R, Paul R K. Price forecasting of mango in Lucknow market of Uttar Pradesh[J]. International Journal of Agriculture, Environment and Biotechnology, 2018, 11(2): 357-363.

[11] Li W, Ding W, Sadasivam R, et al. His-GAN: A histogram-based GAN model to improve data generation quality[J]. Neural Networks, 2019, 119: 31-45.

[12] Gu Y H, Yoo S J, Park C J, et al. BLITE-SVR: New forecasting model for late blight on potato using support-vector regression[J]. Computers and electronics in agriculture, 2016, 130: 169-176.

[13] BV B P, Dakshayini M. Performance analysis of the regression and time series predictive models using parallel implementation for agricultural data[J]. Procedia computer science, 2018, 132: 198-207.

[14] Minghua W, Qiaolin Z, Zhijian Y, et al. Prediction model of agricultural product's price based on the improved BP neural network[C]//2012 7th International Conference on Computer Science & Education (ICCSE). IEEE, 2012: 613-617.

[15] Chen Taotao. Study on Prediction Model of Pig Price and yield based on Deep Learning [D]. Huazhong Agricultural University, 2020.

[16] Lin Xin, Zhu Xiaodong. LSTM stock price forecasting model based on Attention mechanism [J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2022.

[17] Jing Nan, Shi Bauhinia, Shu Yumin. High-frequency price prediction of Shanghai copper futures based on attention mechanism and CNN-LSTM model [J/OL]. Chinese Journal of Management Science: 1-13 [2022-05-22].