J. Electrical Systems 20-2 (2024): 251-257

[1] **Hui Liu**
[2,*] **Mingfang He**
[3] **Shenghan Lai**
[4] **Xiaoying Zhong**

Assessing Performance of the Transformer Model in Predicting Hog Prices

Journal of Electrical Systems

Abstract: - The price of hogs has a significant impact on livelihoods, social development, and overall stability. Therefore, accurate prediction of hog prices is crucial for effective decision-making, breeding strategies, resource allocation, and risk mitigation. In this study, we compare the performance of Transformer and Recurrent Neural Network (RNN) models in predicting hog prices and evaluate their applicability in different scenarios. Additionally, we conduct a generalization test on the hog pig industry chain to assess the models' performance. Our findings indicate that Transformer models excel in parallel computing, context capture, and encoding/decoding tasks. On the other hand, RNN models demonstrate superior performance in predicting extreme events and localized tasks. Therefore, the choice of modeling method should be tailored to meet specific requirements based on the nature of the prediction task.

Keywords: Hog Price Forecast, RNN Model, Transformer Model, Performance Evaluation

I. Introduction

The old adage "as long as the supply of food and pork can be guaranteed, the society will be relatively stable" shows the importance of pork and food in people's daily diet. With China entering a well-off society in an all-round way and achieving the goal of the first century, the diet structure has changed in the past few decades, and the ratio of three categories of food-grain, meat, vegetables and fruits has changed from 8: 1: 1 in the past to 4: 3: 3 now. China is the largest producer and consumer of pork. The price of hog pigs is not only directly related to the livelihood of employees in the hog pig industry chain, but also related to the daily life of ordinary people and the development and stability of the whole society. The price of hog pigs is affected by natural and unexpected factors such as seasonal demand, plague, climate, supply, import and policy, and the price fluctuates constantly. Although hog pig futures were listed on Dalian Stock Exchange on January 8th, 2021, which played a certain role in price discovery, risk avoidance and hedging, the price discovery function still needs to be improved[1]. With the popularization of big data and machine learning technology, data mining methods based on machine learning find out the relevant information buried behind massive data, which is convenient for people to make scientific decisions. As a result, it improves the disadvantages of traditional price prediction models based on historical data, alleviates the impact of drastic fluctuations in pig prices, reduces the uncertainty of pig breeding, and finally provides model support for macro-control of pig industry development and maintains the stable and rapid development of market economy.

II. Literature Review

For the importance of hog pigs, Chinese and foreign scholars apply various quantitative models to optimize accurate pig price prediction. Jia et al. (2021) predicted the price of hog pigs by LSTM neural network model, and applied it to adjust the scale of pig breeding, reduce the cost of environmental pollution control and improve the production efficiency of pig industry [2]. According to the research results of Hamm and Brorsen (1997), the accuracy of BPNN model was higher than ARIMA model in monthly pork price prediction in different time span prediction; For the quarter, it was just the opposite [3]. More foreign literatures improve the accuracy of prediction results through combination models. For example, zhu (2022) adopted a new comprehensive model STL-SVR-ARMA model integrating three technologies in order to predict the future pork price [4]. Ding (2022) explained EEMD and Multi-LSTMs combination model. Firstly, the unstable original sequence was split into several smooth sub-sequences by EEMD model, and then the decomposed sub-sequences were predicted independently by a parallel structure model composed of several LSTMs. All the sub-results are combined by fusion function to obtain the final results, which improves the validity and reliability of the predicted data [5]. In addition foreign

[1] Business School, Nanfang College Guangzhou, Guangzhou, China
[2] School of Economics, Guangzhou College of Commerce, Guangzhou, China
[3] Business School, Nanfang College Guangzhou, Guangzhou, China
[4] Business School, Nanfang College Guangzhou, Guangzhou, China
*Corresponding author: Mingfang He

scholars have used joint prediction models such as EEMD-gray clustering-SVM [6], WOA-Light GBM-CEEMDAN [7], CEEMD and GA-SVR[8], graph enhanced LSTM network[9] , Traditional machine learning models[10], [11],[12]. Zhang Haifeng et al. (2021) analyzed the changing trend of the pig market from the aspects of pig production capacity development, price trend and policy incentives, and estimated that the price of hog pigs will fall in 2021 [13]. Ding Huijuan et al. (2018) used ARIMA and grey model to prediction pig price based on Zunyi pork price. The results showed that ARIMA model was more suitable for short-term predictioning, while grey model can better meet the medium-length predictioning demand [14]. In addition, Chinese scholars also studied X12-ARIMA model[15], ARIMA-GM-RBF combination model [16] to analyze the fluctuation law and predict the price of hog pig.

Scholars screen the econometric analysis methods of single or multiple models under the econometric framework of single variable or multivariate, study the price changes of Hog pigs at different time points, make accurate predictions in advance and equationte reasonable pig breeding schemes to maintain the stable meat price level. In addition, with the continuous maturity of deep learning technology, prediction algorithms based on deep learning, such as RNN, LSTM, GRU, GNN and CNN, have gradually replaced the previous traditional algorithms[17]. Because of the network structure composed of self-attention mechanism, Transformer naturally has unique potential such as super scalability and long-distance learning dependence. Besides replacing the network structures such as convolutional neural network (CNN) and recurrent neural network (RNN)[18], this method also occupies the fields of natural language processing (NLP) and CV, and is widely used in stock price prediction, quantitative investment and portfolio.

III. Framework and Principle of Transformer Model

Different from RNN series networks, which are based on recursive structure, they can only calculate in sequence in one direction, but cannot operate in parallel. Information gradually decays or disappears with the increase of time steps, which leads to gradient disappearance and gradient explosion. RNN and other networks are suitable for short-term prediction scenarios[17]. LSTM and Bidirectional RNN improve the structure of RNN, which is a variant of RNN. By introducing gating mechanism to control the flow of information, the model can better capture the long-term dependence. However, the computational complexity is high, the computational overhead is high, and it is easy to cause gradient disappearance and gradient explosion [19]. Taking a different approach, Transformer model abandons the use of RNN or CNN, and innovatively introduces Attention Mechanism, which can weigh and aggregate all positions of the input sequence at each position, so as to better obtain context information. A convolutional layer called Position-wise Feed-Forward Network is introduced to make nonlinear transformation of features at each position, so as to enhance the expressive ability of the model. The Transformer model consists of four parts: Input, Encoder, Decoder, and Output (see Figure 1)[20]. The encoder is composed of several identical encoder layers stacked, and each encoder layer includes two sub-layers: Multi-Head Self-Attention mechanism and Feed-Forward Neural Network. The decoder is also made up of a plurality of identical decoder layers stacked. Each decoder layer includes three sub-layers: multi-head self-attention mechanism, multi-head encoder-decoder attention mechanism and feedforward neural network. Each encoder and decoder layer puppet adds a residual connection to pass the input directly to the output of the layer. This connection can help information flow from the bottom to the top, and improve the gradient flow and training effect of the model. In order to train the model more steadily, the output of each sub-layer is processed by layer normalization. Because Transformer model does not use RNN or CNN, it cannot directly deal with sequence information. Transformer uses location encoding to add relative or absolute location information for each input location. Trigonometric function is a common way of position coding, which is used to encode the position of sequence. Transformer model can compute in parallel, and has good representation learning ability and context awareness, which achieves significant performance improvement in natural language processing and other tasks. This paper attempts to use five different models to fit the price of hog pigs, and test the prediction effect of Transformer and its extended model.

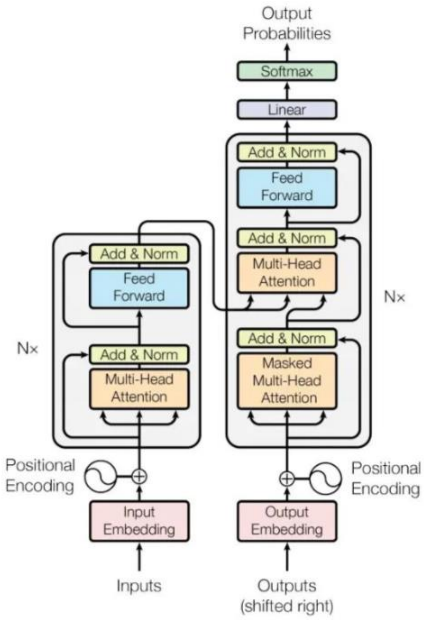

Figure 1: Transformer Model Basic Framework

IV. Empirical Analysis

A. Data Acquisition and Preprocessing and Analysis

The spot closing price of hog pigs from January 8th, 2021 to May 31st, 2023 was obtained from BAIINFO. Python's interpolate function was used to interpolate holiday data according to time interval, and 874 sample data were obtained. During this time, the price of hog pigs first experienced a surge in 2019 and two plunges in 2021 and 2022, and now it is at low and sideways in 2023 due to many influencing factors, especially the extreme historical events such as the three-year epidemic and the conflict between Russia and Ukraine. In turn, descriptive statistical pre-analysis was made (see Table 1). The Minimum and Maximum of this series were 11.02 and 36.4, respectively, which indicated that the price of hog pigs had a wide range. Variance was 33.26467, Stdev was 5.76755, which indicated that the data dispersion was high, Skewness was 1.11405, Kurtosis was 0.51238, which showed that the data distribution was right-sided and relatively flat. In addition, the Jarque-Bera test results showed that the X-squared value was 191.27, the degree of freedom was 2, and the p-value was 2.2e-16 at 5% significance level, which rejected the normal distribution. Based on these statistics, it can be concluded that the price of hog pigs fluctuates greatly and the possibility of price increase is greater. In order to reduce the problem that the model training is unstable due to the inconsistency of feature dimensions, the numerical difference between different features is too large, and the weight of some features is too large, so as to ensure the comparability of various variables and improve the convergence speed of the model, the maximum and minimum normalization is used for data preprocessing, and the data is scaled to [1,-1].

Tabel 1: Hog Price Description Statistics

Mean	**Median**	**Minimum**	**Maximum**	**Variance**	**Stdev**	**Skewness**	**Kurtosis**	**Jarque-Bera**
18.5700	15.9800	11.02000	36.4000	33.26467	5.76755	1.11405	0.51238	191.27

B. Model Selection and Construction

In order to verify the prediction effect of the detection Transformer and its extended model, this paper chooses five models of Simple RNN, LSTM, Bidirectional RNN, LSTM Transformer, and NS Transformer to model respectively. The models are divided into 3 groups, the first group is the traditional RNN model group. The Simple RNN model, which can capture the time dependence in sequence data, is especially suitable for processing sequence tasks with long-term dependencies[21]. The second group is the variant RNN model group. LSTM and Bidirectional RNN models are selected, which are variants of commonly used RNNs for processing sequence data and time series data, and are superior to traditional RNN models in dealing with long-term dependencies and memory problems[17]. Compared with LSTM, the Bidirectional RNN model considers both past and future information in one time step, which can better capture the context information of the sequence[22]. In the third Transformer model group, because of the self-attention mechanism of the basic Transformer model, the output of each position only depends on the input near the position, and does not consider the input of other positions, which leads to the loss of some context information and is not suitable for the spot price prediction of hog pigs.

Therefore, its extended models LSTM Transformer model and NS Transformer model are selected. LSTM Transformer model is suitable for dealing with long-term dependencies of time series, such as time series prediction[20], while NS Transformer model is a more suitable choice for non-normal time series data[23].

All models uniformly set the historical data window size to 24, the training epochs to 370, and the batch_size to 32. 874 samples were divided into training set, test set and verification set according to the ratio of 7: 2: 1, and the corresponding sample numbers of each set were 612, 175 and 87 respectively. Transformer Group Model Parameter Settings Table 2.

Tabel 2: LSTM Transformer & NS Transformer Model Parameter

Parameter Name	Parameter Setting
Number of heads of Transformer	8
Model Learning Rate (lr)	0.001
Hidden Layer size of the network (h)	50
Key vector dimension	16
Optimization function	Adam
Loss function ($loss$)	mean_squared_error

C. Evaluation and Comparison of Prediction Results

Three indexes, RMSE (Root Mean Squared Error), MAPE (Mean Absolute Percentage Error) and AAE (Average Absolute Error), are selected to analyze the prediction results of the five models.

RMSE is used to measure the average difference between the predicted value and the real value of the model, and the prediction accuracy of the model is evaluated by calculating the root mean square error. The calculation equation is shown in Equation (1).

$$RESE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y\_pred_t - y\_true_t)^2} \tag{1}$$

MAPE is used to measure the percentage error between the predicted value and the true value. The calculation equation is shown in Equation (2).

$$MAPE = \sum_{t=1}^{N} \left| \frac{y\_pred_t - y\_true_t}{y\_true_t} \right| \times \frac{100}{N} \tag{2}$$

AAE is used to measure the average absolute error between the predicted value and the true value of the model. The calculation equation is shown in Equation (3).

$$AAE = \frac{1}{N} \times \sum_{t=1}^{N} |y\_pred_t - y\_true_t| \tag{3}$$

The data in Table 3 show that for the three sets, the LSTM model with extended RNN group performs best in the training set, and it has more accurate prediction results and smaller errors. For the training set and validation set, the Bidirectional RNN and Simple RNN models performed better, and for the test set, Simple RNN and Bidirectional RNN performed better. The Transformer Group model is relatively poor in the three sets, but RMSE, MAPE and AAE data are reasonable, and the prediction results can be accepted.

Tabel 3: Performance Evaluation of Five Models

Set	Evaluation Indicators	Traditional RNN	Variant RNN		Transformer	
		Simple RNN	LSTM	Bidirectional RNN	LSTM Transformer	NS Transformer
Training Set	RMSE	0.19901	0.18128	0.19563	0.45802	0.56261
	MAPE	0.81673%	0.70293%	0.79187%	1.90334%	2.84176%
	AAE	0.14340	0.12226	0.13787	0.33880	0.46987
Test Set	RMSE	0.25835	0.24578	0.26490	0.64670	0.73288
	MAPE	0.98173%	0.92254%	1.05111%	2.55440%	3.02201%
	AAE	0.19548	0.18536	0.21361	0.52867	0.57277
Validation Set	RMSE	0.10592	0.08007	0.09768	0.15347	0.43043
	MAPE	0.60174%	0.38847%	0.55109%	0.86515%	2.83181%
	AAE	0.08621	0.05557	0.07892	0.12362	0.40560

According to the prediction results, the following continues to compare and evaluate the prediction results of LSTM and LSTM Transformer model. Figure 2 shows the spot price prediction chart and Q-Q detection chart of LSTM and LSTM Transformer model. From the chart, the LSTM model on the left side of the upper row fits better, especially the train set. The data points in the lower row of Q-Q diagram are basically distributed along the reference line, which shows that the data fit well with the theoretical distribution, but the fitting effect in the tail LSTM model is better.

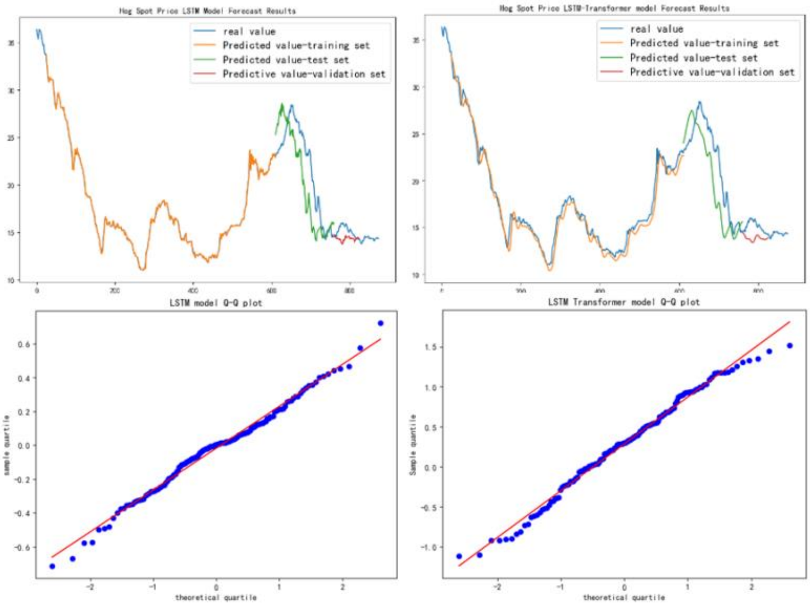

Figure 2: Hog Spot Price Forecast Chart and Q-Q Chart

Figure 3 shows the residual distribution diagram and residual sequence diagram of LSTM model and LSTM Transformer model. From the residual distribution diagram, it can be seen that the error range of LSTM model is [-0.75, 0.75], which is narrower than that of LSTM Transformer model [-1.0, 1.5]. It means that the prediction error of LSTM model is concentrated and relatively stable. There are 5 blue columns with frequencies over 10, one more than LSTM Transformer model, and the maximum frequency of LSTM Transformer model is close to 20, which implies that LSTM model is slightly better than LSTM Transformer model in some error intervals. It can be seen from the residual sequence diagram that the error range of LSTM model is [-0.6, 0.6], which is narrower than that of LSTM Transformer model [-1.0, 1.5], and fluctuates up and down around 0, and the diagram is relatively convergent. It can also be concluded that the prediction error of LSTM model is small, concentrated and relatively stable for pig price.

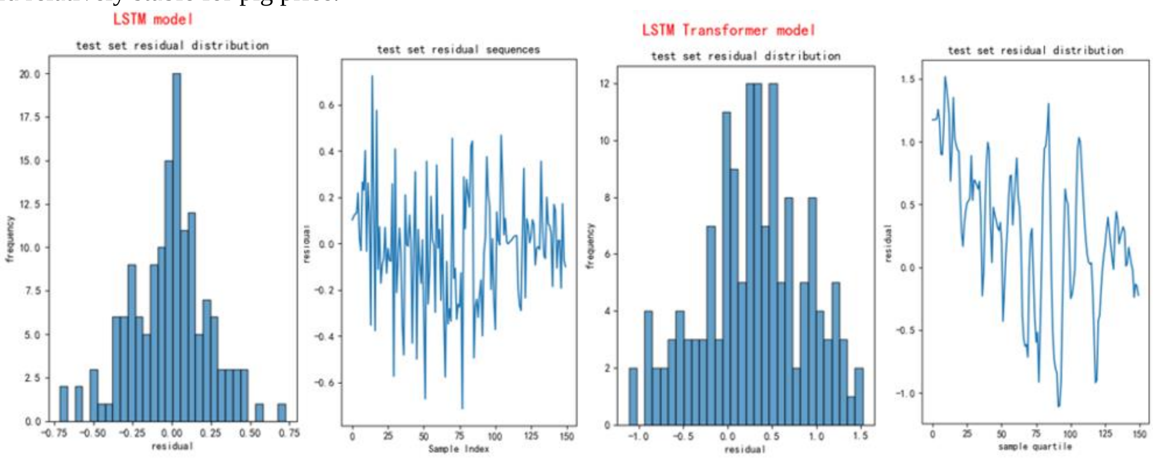

Figure 3: Hog Spot Price Test Set Residual Distribution & Residual Sequence Chart

D. Model Generalization Detection

In order to verify the generalization of the above conclusions, we continue to predict the spot closing prices of soybean meal and corn and the futures prices of hog pigs in the upstream of the hog pig industry chain. In the same industry chain, part of the reasons for their price fluctuations are the same. It can be seen from Table 4 that for soybean meal spot, the Simple RNN model of the first group of traditional RNN models is the best choice, followed by the LSTM model of the second group of RNN variants. LSTM model is the best choice for corn spot and hog pig futures variant RNN group. Corn spot is followed by Bidirectional RNN model, and hog pig futures is Simple RNN model. On the whole, the LSTM model in Variant RNN group performs well in most cases, while the Transformer model performs poorly in some cases.

Table 4: comparison Table of Model Generalization Detection

Commodity	Evaluation Indicators	Traditional RNN	Variant RNN		Transformer	
		Simple RNN	LSTM	Bidirectional RNN	LSTM Transformer	NS Transformer
Soybean Meal Spot	RMSE	29.38879	37.72110	41.22913	159.24319	94.66630
	MAPE	0.42532%	0.61107%	0.59929%	2.72285%	1.69249%
	AAE	21.40334	30.68388	30.88050	140.01313	84.60457
Corn Spot	RMSE	4.53615	3.28840	3.36981	8.18538	15.19828
	MAPE	0.12512%	0.07602%	0.07643%	0.18422%	0.47951%
	AAE	3.65606	2.22548	2.24327	5.41670	14.04740
Hog Futures	RMSE	433.52212	413.48599	446.90105	1633.46287	797.13822
	MAPE	1.67134%	1.58523%	1.72327%	7.24927%	3.66327%
	AAE	292.35433	280.43169	305.64802	1361.83478	635.67989

V. Conclusions and Recommendations

In this study, we compare and explore the Transformer model and the RNN model in the hog pig price prediction problem. Through empirical validation, we find that the RNN model shows higher flexibility and accuracy in dealing with extreme events in time series data, while the Transformer model has some advantages in dealing with local dependencies.

However, there are some shortcomings in this study. First, our empirical validation is based on a specific dataset and time period, and all the data from January 8, 2021 to May 31, 2023, after the listing of the hog futures to the present time, do not go through a complete hog price cycle, which is usually 3~4 years, which means that the forecasting model for hog prices may not be able to fully take into account the price fluctuations during the whole cycle and the trends throughout the cycle. Another limitation is that this study only compared the performance of Transformer and RNN models in hog price prediction, and at the same time, we only considered the single-root prediction problem without involving the analysis and modeling of other related factors. Therefore, in future research, the scope of the dataset can be further extended and more influencing factors can be considered to improve the accuracy and effectiveness of the model.

In terms of future prospects, we can combine the advantages of the Transformer model and the RNN model, focusing on exploring hybrid modeling approaches or enhancing the Transformer model to improve its performance in capturing extreme events and global dependencies in time series data. In addition, we can try to introduce other advanced machine learning techniques and algorithms, such as Deep Learning Networks and Augmented Learning, to cope with more complex time series data analysis and prediction tasks.

Acknowledgements

This work was supported by two projects grant one from Nanfang College, Guangzhou A study on the influence of hog futures on hogs and the industry chain price of hogs (Grant No. 2021XK25), the other from Guangdong Key Construction Discipline Research Capacity Enhancement Project, Research on Digital Inclusive Financial Development System in Guangdong, Hong Kong and Macao Greater Bay Area under Rural Revitalization Strategy(Grant No.2021ZDJS129).

Reference

[1] Wang Tao, Gui Cheng, Research On The Price Discovery Function Of China's Hog Futures Market. Price: Theory & Practice. Beijing, 2023, 03: 140-143+206.

[2] Yunfei Jia, Zhaohui Zhang, et al. Production Efficiency Prediction of Pig Breeding Industry By Optimized LSTM Computer Algorithm under Environmental Regulation. Scientific Programming. Holland, Vol. 2021: 3074167. 1-12, DOI:10.1155/2021/3074167.

[3] Hamm, Lonnie; Brorsen, B. Wade, Forecasting Hog Prices with a Neural Network. Agribusiness. USA, 1997, 154 (1): 37–54.

[4] Zhu H, Xu R, Deng H. A novel STL-based hybrid model for forecasting hog price in China Computers and Electronics in Agriculture. UK, 2022, 198 (7): 1-15. DOI:10.1016/j.compag.2022.107068.

[5] Lianlian Fu, Xinsheng Ding, Yuehui Ding, Ensemble empirical mode decomposition-based preprocessing method with Multi-LSTM for time series forecasting: a case study for hog prices. Connection Science. UK, 2022, 34 (1): 2177-2200. DOI:10.1080/09540091.2022.2111404.

[6] Liwen Ling, Chaomin Cai, Dabin Zhang, An Ensemble Model with Grey Clustering for Hog Price Prediction. Journal of Grey System. UK, 2019, 31 (3): 14-28.

[7] Xiang Wang, Shen Gao, Yibin Guo, et al, A Combined Prediction Model for Hog Futures Prices Based on WOA-LightGBM-CEEMDAN. Complexity. USA, Vol.2022:3216036.1-15, DOI:10.1155/2022/3216036.

[8] D.B. Zhang, C.M. Cai, L.W. Ling, S.Y. Chen, Pork Price Ensemble Prediction Model Based on CEEMD and GA-SVR. J. Syst. Sci. Math. Sci. 2020,40 (6):1061-1073.

[9] Kai Ye, Yangheran Piao, Kun Zhao, Xiaohui Cui, A heterogeneous graph enhanced LSTM network for hog price prediction using online discussion, Agriculture. Agriculture. Switzerland, 2021, 11 (4): 359, DOI:10.3390/agriculture11040359.

[10] Yiran Liu, Qingling Duan, Dongjie Wang, et al, Prediction for hog prices based on similar sub-series search and support vector regression. Computers and Electronics in Agriculture. UK, 2019, 157:581-588.

[11] Liwen Ling, Dabin Zhang, Shanying Chen, et al, Can online search data improve the forecast accuracy of pork price in China?. Journal of Forecasting. UK, 2020, 39 (4): 671-686.

[12] Ma Z, Chen Z, Chen T, et al, Application of machine learning methods in pork price forecast. Proceedings of the 2019 11th International Conference on Machine Learning and Computing. New York, ACM, 2019.

[13] Zhang Haifeng, Lin Zhenji, Guan Zhihui, Analysis of China's Hog Price Trend and Future Market Changes in 2021. Chinese Journal of Animal Science. Beijing, 2021, 57 (04): 237~240.liv

[14] Ding Hui-juan, Zhang Jin-lei,Chen Jian-zhong , et al. Comparison of ARIMA Model and Grey Model in Agricultural Product Price Forecasting. Journal of Anhui Agricultural Sciences. Anhui, 2018, 46 (24): 191-194.

[15] Fan Qingqing, uan Yanhong, Study on Pork Price Fluctuation Law Based on X12-ARIMA Model. Chinese Journal of Animal Science. Beijing, 2018, 54 (06): 138~142.

[16] Wu pei, Li Zheming. Forecast Researchon China's Pork Price Based on the
ARIMA-GM-RBF Hybrid Model. Price:Theory & Practice. Beijing, 2019 (01): 75-78.

[17] Sepp Hochreiter, Jürgen Schmidhuber, Long Short-Term Memory. Neural Computation. USA, 1997, 9 (8): 1735-1780.

[18] Radford A, Narasimhan K, Salimans T, Sutskever I. Improving language under- standingwith unsupervised learning. Technical Report, OpenAI, 2018.

[19] Cho K, Merrienboer B V, Bahdanau D, et al, On the Properties of NeuralMachine Translation: Encoder-Decoder Approaches. Computer Science. Holland, 2014: 1-9.

[20] Ashish Vaswani, Noam Shazeer, Niki Parmar, et al, Attention Is All You Need. last revised 24 Jul 2023:1-15, arXiv:1706.03762v6.

[21] Jeffrey L. Elman, Finding Structure in Time. Cognitive Science. USA, 1990, 14 (2): 179-211.

[22] Mike Schuster and Kuldip K. Paliwal, Bidirectional recurrent neural networks. IEEE Transactions on Signal Processing. 1997, 45 (11): 2673-2681.

[23] Yong Liu, Haixu Wu, Jianmin Wang, Mingsheng Long, Non-stationary Transformers: Exploring the Stationarity in Time Series Forecasting. last revised 18 Oct 2022:1-21, arXiv: 2205.14415v3.