Prediction of Consumer Price Index based on Long Short-Term Memory Model

Xiqin Ao¹, Yujie Gong² and Beibei Zuo¹

¹Institute of information engineering, Anhui Xinhua University, Anhui, 230088, China
²Anhui Sun Create Electronics Co., Ltd. Anhui, 230088, China

Corresponding author’s e-mail: aoxiqin@axhu.edu.cn

Abstract. Consumer Price Index(CPI) is the main standard to identify inflation or deflation. Accurate prediction of CPI will help the government to implement macro-control and formulate price stabilization policies, so as to achieve the goal of building a moderately prosperous society. CPI is a non-stationary and non-linear time series, and has relevance in time dimension. In order to fully mine the correlation of CPI sequence in long and short time span, a method of predicting CPI using Long Short-Term Memory(LSTM) model is proposed. Taking historical CPI data of Anhui Province as the empirical analysis object, modeling and predicting are carried out. The prediction effect of LSTM is compared with classic time series model-Autoregressive Integrated Moving Average(ARIMA). According to the predicted results, LSTM model has significantly improvement in RMSE and MAE indicators compared with ARIMA model, indicating the LSTM model has higher prediction accuracy. The SDAE indicator of LSTM model is smaller than ARIMA model, indicating the LSTM model has better prediction stability.

1. Introduction

At present, China is in the decisive period of building a moderately prosperous society. As a binding index, CPI can be used to judge how to ensure that economic development and growth can fully support the improvement of people's livelihood [1]. CPI reflects the price change trend of service items and consumer goods purchased by urban residents and rural residents in a period of time, as well as the relative number of price change degree. It is the main standard for countries in the world to identify inflation or deflation. When the consumer price index drops, it indicates that the inflation rate drops, that is, the cost of living of consumers is reduced and the purchasing power of money is enhanced [2]. On the contrary, when the rise of the consumer index is too large, it shows that the inflation factor has seriously affected the stable development of social economy, so that the government can adjust the tight monetary policy and fiscal policy to effectively resist the inflation risk, so as to help the stability and control of the economic market. Therefore, a reasonable analysis and prediction of CPI will help the government to implement macro-control and formulate price stabilization policies, so as to achieve the goal of building a moderately prosperous society.

There are many models used to predict CPI, such as time series model, neural network model, support vector machine model [3,4]. Among them, the classic time series model, such as ARIMA model, has better prediction effect for the data set with strong trend, but if it encounters the data set with less strong trend, the effect is not ideal. Traditional neural network models, such as BP neural network, have good nonlinear mapping ability, self-learning and self-adaptive ability, generalization ability and fault tolerance ability, but it is easy to fall into local minimization problem, slow
convergence speed, contradiction between prediction ability and training ability, sample dependence and other problems. Support vector machine model has a good effect for small sample learning, but it is difficult to implement for large-scale training samples. Generally speaking, although these models have good effect on the analysis and prediction of the change trend of the sequence data, they are not enough for the mining depth of the implied relationship in the data sequence. Deep learning algorithms can gradually abstract the input data information and extract features to extract the implicit relationships contained in the data sequence[5,6]. Therefore, this paper proposes the use of deep learning models for the prediction of CPI in order to improve the prediction accuracy of CPI.

2. Theory and method

2.1. LSTM model

LSTM model is a kind of model which can better deal with the problem of time series data prediction in deep learning area at present. It is a kind of fully interconnected recurrent neural network (RNNs) model. There are feedback connections between neurons with time parameters. It can dynamically memorize historical information and keep it persistent while learning new information. The structure of LSTM model unit is shown in Figure 1[7].

![Figure 1. Structure of LSTM model unit](image)

Each LSTM hidden layer contains a memory unit, which is composed of input gate, output gate, forgetting gate and memory cells. LSTM controls the influence of historical information on current information through the above gate unit, so that the network model can save and transfer information for a long time.

At the current time \( t \), the inputs of the LSTM unit are: the input signal \( x_t \), at time \( t \), the output signal \( h_{t-1} \) (history signal) at time \( t-1 \), the state signal \( c_{t-1} \) (memory signal) of the memory unit at time \( t-1 \). The outputs are: the output signal \( h_t \) at time \( t \), and the state signal \( C_t \) at time \( t \).

At time \( t \), the states of input gate, forgetting gate, output gate and memory cell of LSTM are \( i_t \), \( f_t \), \( c_t \), \( o_t \) respectively.

\[
i_t = \sigma(W_i h_{t-1} + W_x x_t + b_i) \quad (1)
\]

\[
f_t = \sigma(W_f h_{t-1} + W_x x_t + b_f) \quad (2)
\]

\[
o_t = \sigma(W_o h_{t-1} + W_x x_t + b_o) \quad (3)
\]

\[
c_t = f_t c_{t-1} + i_t \tanh(W_c h_{t-1} + W_x x_t + b_c) \quad (4)
\]

At time \( t \), the output signal of LSTM is: \( h_t = o_t \tanh(c_t) \)
In the formula (1-4), $W_i, W_f, W_o, W_c$ are the weight matrix of input gate, forgetting gate, output gate and memory cell respectively; $b_i, b_f, b_o, b_c$ are the bias of each gate state respectively; tanh is the activation function.

2.2. CPI prediction model based on LSTM

The LSTM model avoids the long-term dependency problem, and adopts special implicit units to inherit the characteristics of most RNN models while solving the vanishing gradient problem caused by the gradual reduction in the gradient back propagation process. It is suitable for non-linear regression variables, which can solve the problem of multiple input variables. The model has high accuracy, fast training speed, and strong parallel processing ability. LSTM is more suitable for dealing with problems that are highly related to short-term time series. It can process a large number of short-term time series data quickly and accurately by iterating in n sample batches. It is the most commonly used tool to solve the problem of time series prediction[8,9]. The CPI is very typical time series data. Based on this, a CPI prediction method based on the LSTM model is proposed. The method structure is shown in Figure 2.

3. Empirical research

3.1. Data source and preprocessing

The data of this paper comes from the website of the National Bureau of Statistics, in which the monthly CPI data of Anhui Province from January 2000 to July 2019 was selected as the experimental data set, with 228 pieces of data in total. The sequence diagram and correlation diagram of the experimental data set are shown in Figure 3 and Figure 4.
As can be seen from Figures 3 and 4, CPI is typical time series data, and the sequence has a certain periodicity in distribution. From the existing application fields and effects of LSTM model, the model has better applicability to this type of data.

Due to the needs of model modeling, the experimental data set is divided into training set and test set. The first 216 pieces of data belong to training set data, and the last 12 pieces of data belong to test set data.

3.2. Determination of model parameters

(1) Determination of model input and output: since there is only one CPI sequence data in the experiment, the method adopted is to use the CPI of the first 6 months in the sequence to predict the CPI of the seventh month during model training, so the model input is the CPI of the first 6 months and the output is the CPI of the seventh month.

(2) Determination of the number of hidden layer neurons: first, the accuracy of the calculation should be considered, and then the time needed for sample training and prediction should be considered. If the number of neurons in the hidden layer is too small, the learning accuracy will be low, and the number of training iterations will increase. However, if the number of neurons is too large, it will increase the complexity of the network, resulting in more training time and weights [10]. Therefore, the number of hidden neurons determines the final result of the neural network. In this paper, 10 neurons are selected as the number of hidden neurons.

(3) Determination of learning rate and training times: the learning rate and training times are important hyper parameters in supervised learning and deep learning. They determine whether the objective function can converge to a local minimum and when to converge to a minimum. The appropriate learning rate and training times can make the objective function converge to the local minimum value in the appropriate time. After multiple experiments, the learning rate in this paper is set to 0.0005, and the training times is set to 1000.

3.3. Model training and prediction

Based on the LSTM model in Figure 2, this project uses the Mini-Batch method to train the LSTM network. The goal of this paper is to predict the future CPI value, so mean square error (MSE) is selected as the loss function. Adam optimizer was used for optimization training. Adam optimizer is the most commonly used algorithm at present. Compared with other adaptive learning rate algorithms, Adam algorithm has faster convergence speed and more effective learning effect. With the help of tensorflow [11] deep learning framework, data sets are read in, parameters are set, training and prediction are carried out. Finally the prediction results of the test set are shown in Table 1.
4. Model effect verification

In order to verify whether the LSTM model has obvious advantages in the CPI prediction effect, the ARIMA model was used as the comparison model during the experiment.

4.1. CPI prediction model based on ARIMA model

ARIMA is a classical and common model among time series model, which is suitable for non-stationary time series data modeling. The structure of ARIMA (p, d, q) model is as follows:

\[
\Phi(B)\nabla^d \chi_t = \Theta(B)\epsilon_t
\]

\[
E(\epsilon_t) = 0, \text{Var}(\epsilon_t) = \sigma^2, E(\epsilon_t\epsilon_s) = 0, s \neq t
\]

\[
E(\chi_t\epsilon_s) = 0, \forall s < t
\]

In formula (5) \( \nabla^d = (1-B)^d \), \( \Phi(B) = 1-\phi_1B-\cdots-\phi_pB^p \), it is the autoregressive coefficient polynomial of the stationary invertible ARMA (p, q) model. \( \Theta(B) = 1-\theta_1B-\cdots-\theta_qB^q \), it is the moving smooth coefficient polynomial of the stationary invertible ARMA (p, q) model.

Modeling idea of ARIMA model: First, judge the stability of the sequence. If it is not stable, carry out differential processing until it is a stable sequence. The number of differences is the value of parameter d; Next judge the values of the parameters p and q based on the autocorrelation graph and the partial correlation graph; Then the parameter estimation and validity test of the model are carried out, and the final prediction model is determined.

In order to ensure the effectiveness of the comparative experiment, the training data set and test data set of ARIMA model are exactly the same as those of LSTM model. By means of econometrics software Eviews [13], the ARIMA model modeling process has been completed. And according to Akaike Information Criterion (AIC) and Schwarz Criterion (SC), combined with the R-squared optimization principle [14], the optimal prediction model is ARIMA (4,1,5). The model is used for prediction, and the prediction results are shown in Table 2.

| Table 1. Prediction results of LSTM model |
|------------------------------------------|
| time | Aug-18 | Sep-18 | Oct-18 | Nov-18 | Dec-18 | Jan-19 |
| Predictive value | 101.90 | 102.16 | 102.48 | 102.34 | 102.23 | 101.97 |
| time | Feb-19 | Mar-19 | Apr-19 | May-19 | Jun-19 | Jul-19 |
| Predictive value | 101.21 | 101.36 | 102.28 | 102.55 | 102.67 | 102.52 |

| Table 2. Prediction results of ARIMA model |
|------------------------------------------|
| time | Aug-18 | Sep-18 | Oct-18 | Nov-18 | Dec-18 | Jan-19 |
| Predictive value | 101.92 | 102.02 | 101.67 | 101.70 | 101.46 | 101.58 |
| time | Feb-19 | Mar-19 | Apr-19 | May-19 | Jun-19 | Jul-19 |
| Predictive value | 101.48 | 101.67 | 101.67 | 101.88 | 101.90 | 102.09 |

4.2. Evaluation index

It is difficult to judge the prediction effect of the two models directly from the prediction value of the two models. In order to further compare and verify the prediction effect of the two models. In this paper, two indexes, mean absolute error (MAE) and root mean square error (RMSE), are used to evaluate the prediction accuracy; Standard deviation of absolute error (SDAE) is used for the prediction stability evaluation.

\[
\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - g_i)^2} \quad (6)
\]

\[
\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - g_i| \quad (7)
\]

\[
\text{SDAE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(|y_i - g_i| - \text{MAE})^2} \quad (8)
\]
$y_i$ is the real value and $g_i$ is the predicted value. The smaller the RMSE value, the closer the predicted value is to the real value, indicating the higher the prediction accuracy; the smaller the MAE value, the closer the predicted value is to the real value, indicating the higher the prediction accuracy; the smaller the SDAE value, indicating the higher the prediction stability.

4.3. Comparison of prediction effects

The predicted value of ARIMA model and LSTM model are compared with the real value, the results are shown in Figure 5; the evaluation index results of the two models are shown in Table 3.

![Figure 5. Model prediction results](image)

According to Figure 5, the comparison results between the predicted values of the two models and the real values, it can be roughly judged that the fitting degree of the LSTM model is better, and the prediction effect of ARIMA is more gentle.

| Table 3. Model index evaluation results |
|----------------------------------------|
| RMSE | MAE | SDAE |
|------|-----|------|
| ARIMA model | 0.644945868 | 0.572283362 | 0.2974 |
| LSTM model | 0.476455007 | 0.379551317 | 0.2880 |

According to table 3, the RMSE index of LSTM model was 0.1685 lower than that of ARIMA model. And the MAE index of LSTM model decreased by 0.1927 over the ARIMA model; Those results of two indexes show that the RMSE model has better prediction accuracy. According to the SDAE index results, the prediction stability of the LSTM model is significantly better than the ARIMA model.

5. Summary and conclusions

According to the nonlinear and typical time series characteristics of CPI data, a prediction model based on LSTM framework is proposed to further improve the prediction accuracy of CPI. In the empirical study, the CPI data of Anhui Province is used to model and predict, and the prediction effect of LSTM is compared with ARIMA model. Three evaluation indicators show that the prediction accuracy and stability of the LSTM model are significantly better than the ARIMA model. But there are still some shortcomings in the experiment, for example, only historical data of CPI is used as the model input, and the change of CPI is affected by many factors. If the main influencing factors can be summarized and used as the input of the model, it is expected that the accuracy of CPI prediction will be further improved.

Acknowledgements

This research was financially supported by Outstanding Youth Visit Program (Grant No. gxgnfx2019069); Anhui Province Innovative Training Program for College Students.
CPI prediction of Anhui Province based on deep learning); School-level research team(Grant No kytd201902); 2019 Anhui University Scientific Research Projects (Grant No. KJ2019A0877); 2017 School-level Teaching Team Project for Information Management and Information System Major (Grant number: 2017jxtdx07).

References
[1] Xiao G Q (2015). Connotation, strategic status and institutional guarantee of a well-off society. Leading Journal of Ideological and Theoretical Education, 09:62-67.
[2] Xu G J, Xu K, Lu Q Q (2019). CPI weight, structural price rise and price perception. Price: Theory and Practice. http://kns.cnki.net/kcms/detail/11.1010.F.20191126.1127.002.html.
[3] Fan J Y, Zou Y M, Wu C D (2016), An empirical study on the prediction model of SVM and ARMA based on wavelet analysis--Taking CPI as an example. Science & Technology and Economy, 05:6-9.
[4] Yang J,Fei J J(2014), Empirical analysis and prediction of the trend of China's consumer price index. Statistics & Decision, 07:112-116.
[5] Zhan X K, Li Y H, Li R X, et al (2018). Stock Price Prediction Using Time Convolution Long Short-Term Memory Network. In:Knowledge Science, Engineering and Management Conference. Changchun. pp.461-468
[6] Hoang M N, Gaurav K, Daeyoung K (2019). Host load prediction in cloud computing using Long Short-Term Memory Encoder–Decoder. The Journal of Supercomputing, 75 (11): 7592-7605
[7] Wang X, Wu J,Liu CH, et al(2018). Fault time series prediction based on LSTM recurrent neural network. Journal of Beijing University of Aeronautics and Astronautics, 44(04):772-784.
[8] Ma J, Ding Y X, C.P J. et al(2019). A temporal-spatial interpolation and extrapolation method based on geographic Long Short-Term Memory neural network for PM2.5. Journal of Cleaner Production, 237.
[9] Ghimire S, C R, Raj N, et al (2019). Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms. Applied Energy, 253.
[10] Yu N G, Weber C, Hu X L (2019). Learning Sparse Hidden States in Long Short-Term Memory. Artificial Neural Networks and Machine Learning. pp.288-298
[11] Wang X M, Pan M, Li J J (2018). Deep learning and tensorflow practice. Posts and telecom press, Beijing
[12] Zhang W T (2018), SPSS Statistical Analysis Advanced Course (3rd Edition). higher education press, Beijing
[13] Li Y Y, Liu R, Dai W D, et al (2013). Eviews statistical analysis and Application. Electronic Industry Press, Beijing
[14] Gao T M (2016). Econometric Analysis Methods and Modeling. Tsinghua University Press, Beijing.