Journal of Computers Vol. 34 No. 2, April 2023, pp. 69-79
doi: 10.53106/199115992023043402006

Air Quality Index Prediction Based on a Long Short-Term Memory Artificial Neural Network Model

Chen Wang[1], Bingchun Liu[2*], Jiali Chen[2], Xiaogang Yu[3]

[1] Organization Department, Tianjin University of Technology,
Xiqing District, Tianjin, China
scxxingchen520@163.com

[2] School of Management, Tianjin University of Technology,
Xiqing District, Tianjin, China
tjutlbc@163.com, tjutcjl@163.com

[3] Tianjin 712 Mobile Communication Co., Ltd.,
TCB Science Park, Hebei District, Tianjin, China
15935142263@163.com

Received 22 May 2022; Revised 23 August 2022; Accepted 19 September 2022

Abstract. Air pollution has become one of the important challenges restricting the sustainable development of cities. Therefore, it is of great significance to achieve accurate prediction of Air Quality Index (AQI). Long Short Term Memory (LSTM) is a deep learning method suitable for learning time series data. Considering its superiority in processing time series data, this study established an LSTM forecasting model suitable for air quality index forecasting. First, we focus on optimizing the feature metrics of the model input through Information Gain (IG). Second, the prediction results of the LSTM model are compared with other machine learning models. At the same time the time step aspect of the LSTM model is used with selective experiments to ensure that model validation works properly. The results show that compared with other machine learning models, the LSTM model constructed in this paper is more suitable for the prediction of air quality index.

Keywords: Index of Air Quality, prediction, deep learning, LSTM

1 Introduction

Air pollution has become an important environmental problem that restricts economic and social development. Urban air pollution originates from transportation, industry, energy consumption, living, and many other factors. Its most serious short-term impact is on the health of urban residents, and it also brings about global warming by means of the greenhouse effect, which facilitates a series of long-term climate changes. Fine particulate matter ($PM_{2.5}$), SO_2, NO_2, O_3, and other air pollutants lead to the degradation of human health as they can cause headaches, difficulty breathing, heart obstructions, and other symptoms [1]. The Air Quality Index (AQI) is an indicator that assesses the impact of air pollution on human health, and it can be understood easily by the public [2]. The US Environmental Protection Agency (EPA) first proposed the Pollution Standard Index (PSI) in 1972, and in 1999 the EPA proposed the Air Quality Index (AQI), which is now widely used in the world. In China, the air quality index (AQI) is obtained by converting real-time monitoring data of air quality so it is very important to compute the AQI value of the next day after receiving the relevant data. The AQI is a policy tool employed by the Environmental Protection Department, and AQI predictions are used by the Department to regulate and control traffic, industry, urban construction, and other socio-economic activities that take place within areas that can be susceptible to unsatisfactory air pollution conditions.

Many previous studies on air pollution prediction have focused on the application of deterministic and computational models. The deterministic model does not require a large amount of historical data but does require adequate source knowledge, timely emission quantities, and spatiotemporal physical transformation of exhaust gases from the main chemical reactions. McKeen [3] and Chuang [4] applied deterministic models of on-line/off-line meteorological chemistry models to deterministic problems. It should be noted that, in the absence of knowledge of pollution sources, and where physical changes occur in the processing of insignificant conditions, larger deviations will appear in the application of deterministic models. Computational models often require a large number

* Corresponding Author

of historical measurements under a variety of meteorological conditions, which establishes a relationship between historical pollution data and predicted variables through regression and machine learning [5]. A variety of machine learning algorithms have been used in air pollution prediction, such as the ANN (Artificial Neural Network) algorithm used by Pérez [6] and Li [7], the FC (Fuzzy logic) algorithm used by Shad [8] and Alhanafy [9], the KF (Kalman filter) algorithm used by Zolghadri [10] and Hoi [11], and the HMM (hidden Markov model) algorithm used by Sun [12].

Compared to the dynamic characteristics of air pollution indices, recurrent neural networks (RNN) are effective to solve the adverse effects of the spatial and temporal evolution of an air pollution index. RNN is a deep learning method that can process any input sequence using memory cells across networks and therefore it has the ability to learn time series. Many RNN technologies have been proposed to solve the time series prediction problems, but studies have shown that the typical RNN models cannot solve the problem of the long time-dependence of input sequences. In order to solve this problem, a special RNN structure: a long-term and short-term memory artificial neural network (LSTM NN) has been proposed [13]. The LSTM network can learn the time series over a long time-span and can determine automatically the optimal time lag for accurate prediction. Over the past few years, LSTM has been applied successfully in handwriting recognition, human motion recognition, robot control, etc., but its application in air pollution prediction is still not commonplace. This study attempts to verify the effectiveness of the LSTM model in AQI prediction. First, select the air pollutants index and weather phenomena index that affect AQI, and use the information gain method to analyze their characteristics, and then construct the optimal index of IG (Information Gain). Secondly, the air pollutants index, weather phenomena index, optimal index of IG and air pollutant variable index are used as input indicators of the hybrid model. Finally, the prediction of AQI is realized through The AQI hybrid prediction model, and further analysis and demonstration is carried out according to the prediction results.

The contributions of this research mainly include three aspects: (1) Apply LSTM in deep learning to the field of air pollution prediction, and achieve effective prediction of air quality index by building a hybrid model; (2) According to the research foundation of scholars, by using information The gain method extracts the features of the indicators that affect the AQI, improves the reliability of the input indicators, and realizes the selection and optimization of the input indicators of the prediction model. (3) This study uses four input indicators to divide into two branches, and uses the Merge method of the LSTM model to merge and transform, which greatly improves the correlation between indicators, improves the prediction accuracy, and provides new ideas for air quality prediction. (4) Compared with other prediction models, it is found that the LSTM model used in this study has higher performance in air quality index prediction, and it is determined that the LSTM model is suitable for the research of atmospheric environment.

2 Literature Review

2.1 Research on Air Quality Index

The Air Quality Index (AQI) is an important indicator for judging the impact of air pollution on human health. With the help of air quality indices, the public can clearly understand whether the air quality is good or bad for their health. Government agencies also can make decisions on pollution reduction and environmental management methods with supportive data from these indices. In recent years, many scholars have begun to pay attention to air quality index research. Bishoi (2009) attempted to calculate the Air Quality Index based on Factor Analysis (NAQI), which avoided the deficiencies of the USEPA method [14]. Gorai (2014) employed the fuzzy synthetic evaluation model to assess air quality in four monitoring stations situated in the Taj Trapezium Zone [15]. Xu (2017) proposed a new dynamic fuzzy synthetic evaluation and applied it to estimate air quality quotients and rank the primary pollutants [16]. Sarella (2015) proposed an air quality index (AQI) for the City of Vapi, India, for simplified public information and data interpretation [17]. Ruggieri (2013) focused on functional principal component analysis (FPCA) to investigate multiple pollutant datasets measured over time at multiple sites within a given urban area, and the approach was proved to effectively highlight relevant statistical features of the time series [18]. Van den Elshout (2014) described the logic of making an index, in particular, the CAQI, and its update with a grid for $PM_{2.5}$ [19]. Kirenga (2015) conducted air pollutant monitoring in Kampala and Jinja, Uganda, and found that air pollutant concentrations were dangerously high. Therefore, long-term studies were needed to characterize air pollution levels during all seasons [20]. Lv (2016) applied statistical methods to have better data comparisons, which showed that vehicles, dust, industry, biomass burning, coal combustion, and secondary products all were major sources of air pollution, and they had all continued to increase, except coal combustion and secondary products [21].

2.2 Research on Traditional Method Forecasting

At present, the prediction of air quality has become a significant topic in current studies. Cogliani (2001) used a regression model combined with meteorological variables to predict the air pollution index [22]. Hajek (2009) applied the fuzzy inference system to establish an air quality index forecasting model [23]. Zickus (2013) applied Gaussian and regression models to forecast urban air pollution and found that short-term forecasts based on statistically-determined relationships were more consistent with the measured data from a specific location than was the predicted value from a Gaussian model [24]. Combarro (2013) studied daily air pollution levels using the support vector machine (SVM) technique in the Oviedo urban area (Northern Spain) on a local scale [25]. Kassomenos (2013) attempted to create a suite of statistical models for predicting one-day-ahead maximum CO levels, based on both meteorological and pollutant data, which were recorded at six monitoring sites in the greater area of Athens, Greece. The results showed that the prognostic models managed to predict the CO maximum daily values with satisfactory accuracy [26]. Peng (2016) employed nonlinear updatable machine learning methods to evaluate hourly air quality forecasting in Canada and concluded that the nonlinear models often had a worse bias (mean error) and more severe underprediction of extreme events than did linear models [27].

2.3 Research on Neural Network Forecasting

In recent years, neural network models have been widely used in the field of atmospheric pollutants. Dahe (2004) applied the improved ANN model to predict the air pollution index in Shanghai [28]. With the help of Artificial Intelligence Methods, Karatzas (2009) understood and predicted air pollution levels in Athens, Greece [29]. Azid (2013) built a Feed-Forward Artificial Neural Network Model for Air Pollutant Index Prediction in the Southern Region of Peninsular Malaysia [30]. Kyriakidis (2012) employed a number of Computational Intelligence algorithms to study the forecasting of the hourly and daily changes in the CAQI (Common Air Quality Index) quotient, which utilized artificial neural networks, decision trees, and regression models [31]. A novel hybrid forecasting model combining a two-phase decomposition technique and an extreme learning machine (ELM) that was optimized by a differential evolution (DE) algorithm was developed for AQI forecasting by Wang (2016) to enhance forecast accuracy [32]. Feng (2015) built a hybrid model that combined air mass trajectory analysis and wavelet transformation to improve the artificial neural network (ANN) forecast accuracy of daily average concentrations of $PM_{2.5}$ two days in advance [33]. Kurt (2010) presented methods based on geographic forecasting models, using neural networks (GFM-NN), to forecast air pollutant indicator levels and the findings were quite satisfactory [34]. Plaia (2013) proposed a two-step aggregation that was related to space and pollutant levels in order to get a Multipollutant-Multisite Air Quality Index (AQI) time series [35].

Moustris (2013) dealt with the development and application of ANN models as a tool to forecast daily concentration levels of PM_{10} in five different regions within the greater Athens area, and the model performance showed that the ANN models could successfully forecast the risk of daily PM_{10} concentration levels exceeding certain thresholds [36]. Moustris (2010) made use of Artificial Neural Networks to forecast the maximum daily value of the European Regional Pollution Index as well as the number of consecutive daily hours with at least one of the pollutants being above a threshold concentration, 24 to 72 h ahead. The results were in agreement with the real-time monitored data at a statistically significant level [37]. Dunea (2015) presented the screening of various feed-forward neural networks (FANN) and wavelet feed-forward neural networks (WFANN) applied to time series of ground-level ozone (O_3), nitrogen dioxide (NO_2), and particulate matter (PM_{10} and $PM_{2.5}$ fraction) concentrations recorded at four monitoring stations located in various urban areas of Romania to develop an optimized general sampling configuration. The results showed that both modes (i.e. FANN and WFANN) overestimated $PM_{2.5}$ forecasted values during the last quarter of the time series [38]. Beddows (2017) analyzed the emulation and sensitivity of the Community Multiscale Air Quality Model with the help of base case and found that the performance of the model appeared slightly better in Harwell than London [39]. In summary, these studies were attempts, qualitatively or quantitatively, to solve practical problems. However, their accuracy was quite low, and the present investigation aimed at to employ a method of deep learning to reduce prediction errors.

3 Methodology and Data

3.1 Design of the Model Structure

The AQI hybrid prediction model consists of three layers: the input layer, the hidden layer, and the output layer. The input layer consists of the air pollutant index, the air pollutant variable index, the weather phenomenon index, and the optimal characteristic index calculated by Information-gain method (IG). The hidden layer consists of information and information extraction. The total number of hidden layers is ten. The LSTM algorithm is used for information extraction and the CONCAT algorithm is used for information merging. The output layer of the model includes only one index of AQI for the next day. The specific model structure is shown in Fig. 1. The realization of the model can be broken down into two parts: the first is the design of the feature index, the second is the structure and calculation of the model.

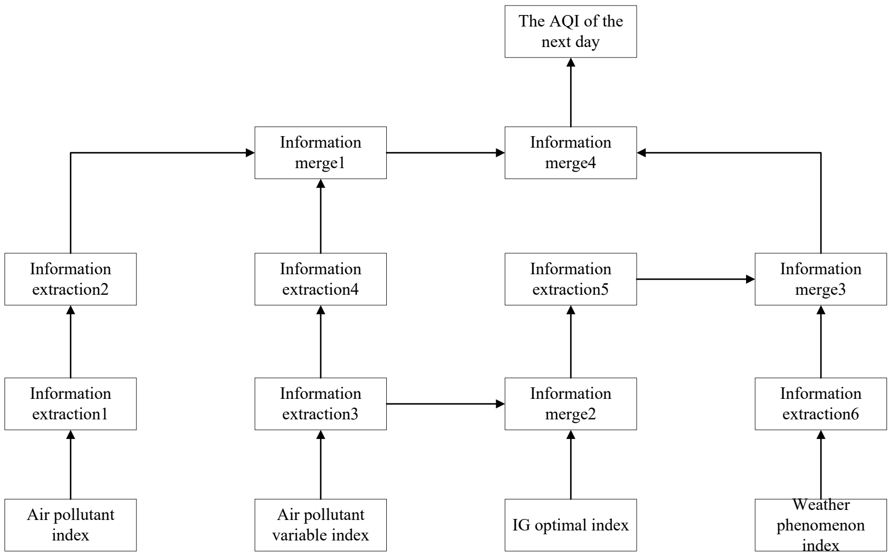

Fig. 1. The model framework of air quality index prediction based on IG-LSTM

3.2 Design of the Feature Index

In order to further improve the accuracy of AQI forecasting, four characteristic indices were selected to use as inputs to the model. The air pollutants index included $PM_{2.5}$, PM_{10}, SO_2, CO, NO_2, and O_3. The weather phenomena index included five characteristic indices: maximum temperature, minimum temperature, weather phenomenon (includes: cloudy, sunny, rainy, overcast sky, snowy, sand dust, haze, foggy), wind direction, and wind power.

The optimal index of IG (Information Gain) is the optimal characteristic index, which is selected from $PM_{2.5}$, PM_{10}, SO_2, CO, NO_2, O_3, maximum temperature, minimum temperature, weather phenomenon (i.e. cloudy, sunny, rainy, overcast sky, snowy, sand dust, haze, foggy), wind direction, wind power, and the AQI of the day before, to which is applied the calculation method for ‘information gain.

3.3 Model Structure and Computational Logic

The main structure of the model is composed of two main branches. One main branch utilizes inputs of the air pollutant index and the air pollutant variable value, and the corresponding key characteristic information is obtained by the two layers of the LSTM algorithm. The other main branch utilizes inputs of the weather phenomenon index and the optimal characteristic index obtained by the IG (information gain) method. In order to amplify the influence of the IG optimal index, the air pollutant variable index is added to the result of the first layer of the

LSTM calculation as a common input. The two-part input is transformed by a layer of the LSTM calculation to obtain the corresponding key characteristic information. Finally, the results of the two branches are combined as the final information output for the prediction of the AQI for the next day. The LSTM deep learning algorithm is used in the whole structure of the model, and the main calculation logic is shown in Fig. 2.

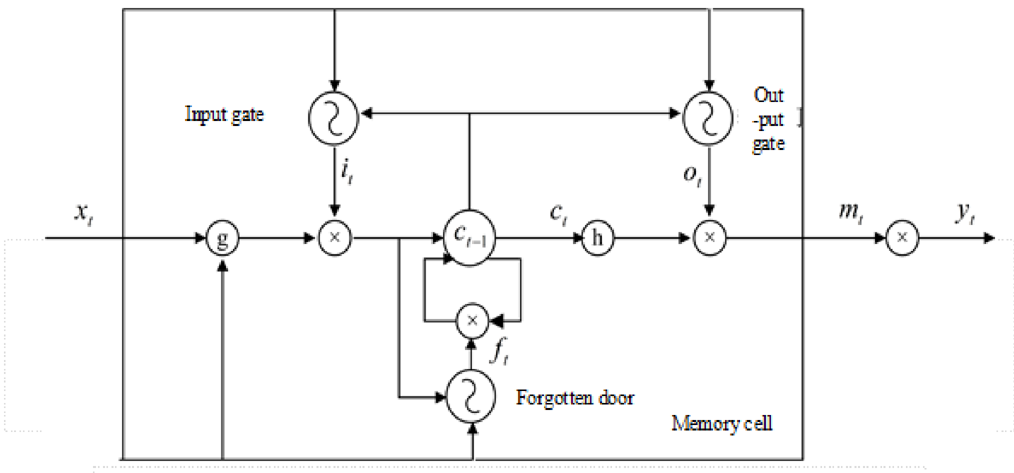

Fig. 2. The main calculation logic of LSTM

$$i_t = \sigma(w_{ix}x_t + w_{im}m_{t-1} + w_{ic}c_{t-1} + b_i). \tag{1}$$

$$f_t = \sigma(w_{fx}x_t + w_{fm}m_{t-1} + w_{fc}c_{t-1} + b_f). \tag{2}$$

$$c_t = f_t \otimes c_{t-1} + i_t \otimes g(w_{cx}x_t + w_{cm}m_{t-1} + b_c). \tag{3}$$

$$o_t = \sigma(w_{ox}x_t + w_{om}m_{t-1} + w_{oc}c_t + b_o). \tag{4}$$

$$m_t = o_t \otimes h(c_t). \tag{5}$$

$$y_t = w_{ym}m_t + b_y. \tag{6}$$

Where $\otimes$ denotes as the scalar product of two vectors, $\sigma(x)$, $g(x)$ and $h(x)$ is defined as the sigmoid function:

$$\sigma(x) = \frac{1}{1+e^{-x}},\ g(x) = \frac{1}{1+e^{-x}},\ h(x) = \frac{1}{1+e^{-x}}. \tag{7}$$

3.4 Data Source

The data used in this investigation were the air quality data of Tianjin (Jan 1, 2014 - Dec 30, 2016) and the meteorological data of Tianjin (Jan 1, 2017 - Sep 30, 2017). The air quality data of Tianjin were obtained from the open data released by the China Environmental Monitoring Station. The meteorological data of Tianjin were obtained from meteorological data released by Tianjin Meteorological Observatory. The data used was daily data. A total of 1367 data samples were collected from Tianjin data, and 15% of the total data samples were classified as training samples. The sample data for training were used to revise the training of the planning model, and the sample data of the test was used to judge the accuracy of model prediction. The experiment uses a total of 1095 pieces of information (Jan 1, 2014 – Dec 30, 2016) as a sample for training, and a total of 272 pieces of information (Jan 1, 2017 - Sep 30, 2017) as the sample for the test.

4 Results and Discussion

4.1 Prediction Model Evaluation Index

To accurately evaluate the validity of the experimental model, this paper adopts the mean absolute percentage error MAPE (Mean Absolute Percent Error) to quantitatively evaluate the predictive performance of the proposed model. The evaluation index is used to measure the difference between the simulated data and the model data. The smaller the value, the higher the model accuracy and the better the prediction performance. The error calculation formula was as follows:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y-\hat{y}}{y}\right| \times 100. \quad \textbf{(8)}$$

Where y is the actual value, $\hat{y}$ is the corresponding predicted value, and n is the number of indicator variables.

4.2 Results of the IG

The first step of the experiment used the IG (Information Gain) method to select the optimal features from the existing features. The direct collection features included $PM_{2.5}$, PM_{10}, SO_2, CO, NO_2, O_3, maximum temperature, minimum temperature, weather phenomena, wind direction, wind power, and the AQI quotient of the day before; whose total number – i.e. the number of direct collection features - was 12. The calculation results for the characteristic entropy of the information gain method are shown in Table 1.

Table 1. The calculation results for the characteristic entropy of the IG method

Index	Entropy value	Sequence
The AQI quotient of the day before	4.55	3
$PM_{2.5}$	4.71	2
PM_{10}	3.56	6
SO_2	4.43	4
CO	3.79	5
NO_2	4.84	1
O_3	3.49	7
Maximum temperature	2.07	8
Minimum temperature	2.03	9
Weather phenomena	0.71	11
Wind direction	0.85	10

The entropy values that are more than 4 are featured in the IG optimal index, namely NO_2, $PM_{2.5}$, the AQI quotient of the day before, and SO_2. For the final calculation and selection, the final indicators and characteristics of the inputs of the model are shown in Table 2.

Table 2. The final indicators and characteristics of the inputs of the model

Feature index	Index	Explain
Air pollutant index	$PM_{2.5}$, PM_{10}, SO_2, CO, NO_2,O_3	China Environmental Monitoring Station releases the open data of air pollutant concentration, but it currently announces only 6 types.
Air pollutant variable index	$\Delta PM_{2.5}$, ΔPM_{10}, ΔSO_2, ΔCO, ΔNO_2, ΔO_3	The air pollutant variable index is based on the data of air pollutant concentration.
Weather phenomenon index	maximum temperature minimum temperature, weather phenomena, wind direction, wind power	Weather data are released by the Meteorological Observatory of Tianjin. but it currently announces only 5 types.
IG optimal index	NO_2, $PM_{2.5}$, the AQI quotient of the day before, SO_2	The characteristic index whose characteristic entropy value is more than four is obtained with the method of IG.

4.3 Results and Comparison of the LSTM

In order to more comprehensively evaluate the prediction performance of the LSTM model and achieve the optimal goal of the final model, this paper compares the LSTM model with some machine learning models, choosing SVR (Support Vector Programming), RDG (Ridge Regression), LASSO (Least Absolute Shrinkage and Selection Operator), DT (Decision Tree), GBR (Lifting Tree Regression), and KRR (Kernel Ridge Regression) six models are used as comparison models. With other conditions unchanged, the final results are shown in Table 3.

In order to improve the efficiency of model training with LSTM, the deep learning program KERES, which uses time steps as the key parameters that influence the final predicting results, was used to compare the prediction results for different time steps. According to the basic principle of the LSTM algorithm, the selected range of time steps was from 2 to 6. Each model operation was performed 10 times to reduce randomness. The final results are shown in Table 3, with the calculation of various models.

Table 3. The compare of predicting result on various models l

Serial number	The model name	MAPE	
		Training values	Test value
1	SVR	0.2624	0.2326
2	RDG	0.2669	0.2123
3	LASSO	0.2813	0.369
4	DT	0.2428	0.2052
5	GBR	0.2139	0.1835
6	KRR	0.2827	0.2016
7	LSTM (steps = 3, epoch = 20000)	0.1579	0.0987

According to the results, the LSTM model proposed in this paper is more effective than other prediction models. Among the seven experimental models, the LSTM model had the lowest error, with a MAPE of 0.0987 and a time step of 3, and was the only model with a MAPE value below 0.1. Among the machine learning models, the GBR model is the best among the six machine learning models, with a MAPE of 0.1835, but it can be clearly seen that its results are far inferior to the LSTM model. The LASSO model has the lowest accuracy, with a MAPE of 0.369. The results are shown in Fig. 3 and Fig. 4.

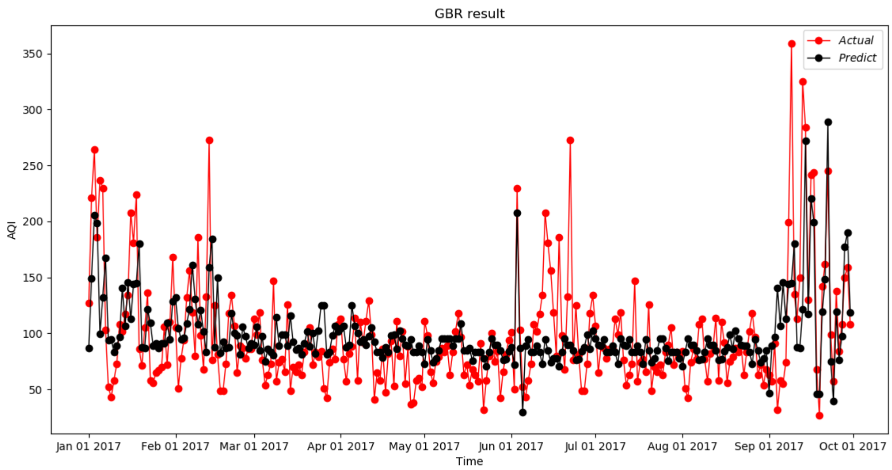

Fig. 3. The test result of GBR model

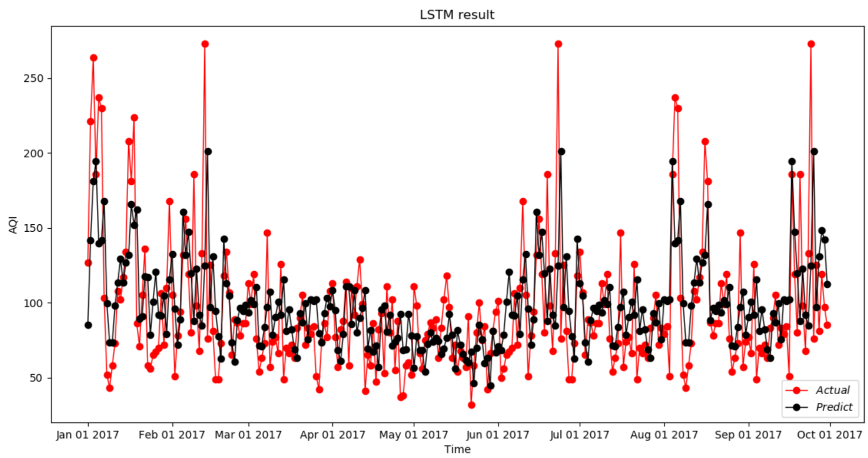

Fig. 4. The test result of LSTM model

Based on an analysis of the results, the following three discussion results are obtained.

(1) Overall, the LSTM algorithm is superior to the machine learning model for the processing of time series data. With the combination of multiple LSTM layers and MERGE layers, a more accurate prediction model for the next-day AQI quotient can be established.

(2) The characteristics index of the model's input can be further selected for optimization. The feature index is the basis of the model's prediction and judgment. Feature selection or extraction with the IG method can improve data validity very effectively and can help the model to improve its judgment ability.

(3) The number of time steps is a crucial parameter in the LSTM model, and the correct time step setting can improve prediction accuracy. As the number of time steps is related to the amount of data in the back-trace time, it can be stated that both excessively-large and too-small time steps in the experiment can affect the accuracy of the prediction.

4.4 Application Analysis of Predictive Models

Accurate prediction of air quality index is very important in air pollution control. The accuracy of the LSTM prediction model constructed in this paper has been verified, and it has also made important contributions to relevant policy formulation and early warning to the public.

(1) Predict the AQI quotient with the help of LSTM, fully exploit the ambient air quality data resource information, meet the real-time and validity requirements of ambient air quality forecast, and minimize the air quality forecast error.

(2) The experimental results will meet the air quality monitoring and forecasting needs of the government and residents, enabling them to mine the effects of air pollutant levels. Governments may be inspired by the results of experiments to apply such tests in the formulation and implementation of related policies such as energy. At the same time, it will help relevant government departments to issue early warning information in real time and help residents adjust their travel plans.

(3) The research results can help to grasp the changing trend of AQI quotient, so that reasonable actions can be taken. In addition, in addition to macro policies, the improved accuracy of AQI forecasts can allow governments to formulate comprehensive and detailed policies that provide intuitive guidance for activities such as production and living.

Therefore, the LSTM prediction model proposed in this paper uses the IG method to conduct a more reasonable analysis of air quality-related indices, which can be used as a monitoring and early warning tool in some areas to improve the prediction level of air pollutants.

5 Conclusions

The LSTM algorithm is a long-short-term memory-based sorting method for machine learning of time series data. Based on its advantages in processing time series data, this study established a long short-term memory artificial neural network model for predicting the combined air pollution index quotient. The neural network model mainly consists of two parts: one is to select and process the existing index data by using the IG method based on the existing air pollutant indexes and weather phenomenon index values; secondly, a two-branch LSTM prediction model is established. The experimental results are as follows:

(1) Air quality is related to air pollutant index, air pollutant variable index, weather phenomenon index, and can be used as the input variable of the model to effectively predict the air quality index.

(2) Using the information gain method, select and extract the optimal eigenvalues from indicators such as air pollutant index, weather phenomenon index, etc. NO_2, $PM_{2.5}$, the AQI quotient of the day before, SO_2 as the eigenvalue with the highest quotient has the greatest relationship with the air quality index, and entering it into the model as an input indicator helps to improve the performance accuracy.

(3) This paper selects the LSTM model to predict the air quality index of Tianjin. The experimental results show that the MAPE of the prediction model is 0.0987. Compared with the machine learning model, the model can be more precise and accurate, and the prediction accuracy can be further improved with proper and effective feature design and parameter settings.

It can be seen that the LSTM model used in this paper is feasible in the prediction of air quality index, and the prediction accuracy is high. However, due to the limitations of experimental equipment and data sources, there are still many areas for improvement. For example, in the future, the structure of the LSTM prediction model can be deeply developed, the parameter settings can be optimized, and the accuracy of the prediction results can be further improved. Furthermore, considering the correlation of air quality in different cities, data from multiple cities can be leveraged to facilitate collaborative prediction.Therefore, in the future research process, we will continue to conduct in-depth research in this area to break through the limitations of the article.

Acknowledgement

This work supported by the Teaching Fund Key Program of Tianjin University of Technology (grant number ZD19-06).

References

[1] N. Künzli, R. Kaiser, S. Medina, M. Studnicka, O. Chanel, P. Filliger, M. Herry, F. Horak, V. Puybonnieux-Texier, P. Quénel, J. Schneider, R. Seethaler, J.-C. Vergnaud, H. Sommer, Public-health impact of outdoor and traffic-related air pollution: a European assessment, The Lancet 356(9232)(2000) 795-801.

[2] A. Plaia, M. Ruggieri, Air quality indices: a review, Reviews in Environmental Science and Bio/Technology 10(2)(2011) 165-179.

[3] S. McKeen, G. Grell, S. Peckham, J. Wilczak, I. Djalalova, E.-Y. Hsie, G. Frost, J. Peischl, J. Schwarz, R. Spackman, J. Holloway, J. de Gouw, C. Warneke, W. Gong, V. Bouchet, S. Gaudreault, J. Racine, J. McHenry, J. McQueen, P. Lee, Y. Tang, G. R. Carmichael, R. Mathur, An evaluation of real-time air quality forecasts and their urban emissions over eastern Texas during the summer of 2006 Second Texas Air Quality Study field study, Journal of Geophysical Research 114(D7)(2009) 1-26.

[4] M.-T. Chuang, Y. Zhang, D. Kang, Application of WRF/Chem-MADRID for real-time air quality forecasting over the Southeastern United States, Atmospheric Environment 45(34)(2011) 6241-6250.

[5] R. Stern, P. Builtjes, M. Schaap, R. Timmermans, R. Vautard, A. Hodzic, M. Memmesheimer, H. Feldmann, E. Renner, R. Wolke, A. Kerschbaumer, A model inter-comparison study focussing on episodes with elevated PM10 concentrations, Atmospheric Environment 42(19)(2008) 4567-4588.

[6] P. Perez, J. Reyes, An integrated neural network model for PM_{10} forecasting, Atmospheric Environment 40(16)(2006) 2845-2851.

[7] V. Villanyi (Eds.), Air Pollution, IntechOpen, 2010 (Chapter 9).

[8] R. Shad, M.-S. Mesgari, A. Abkar, A. Shad, Predicting air pollution using fuzzy genetic linear membership kriging in GIS, Computers Environment and Urban Systems 33(6)(2009) 472-481.

[9] T.-E. Alhanafy, F. Zaghlool, A.S.E.D. Moustafa, Neuro fuzzy modeling scheme for the prediction of air pollution, Journal of American Science 6(12)(2010) 605- 616..

[10] A. Zaharim, M. Shaharuddin, M.J.M. Nor, O.-A. Karim, K. Sopian, Relationships between airborne particulate matter and meteorological variables using non-decimated wavelet transform, European Journal of Scientific Research 27(2)(2009) 308-312.

[11] K.-I. Hoi, K.-V. Yuen, K.-M. Mok, Kalman filter based prediction system for wintertime PM_{10} concentrations in Macau, Global Nest Journal 10(2)(2008) 140-150.

[12] W. Sun, H. Zhang, A. Palazoglu, A. Singh, W. Zhang, S. Liu, Prediction of 24-hour-average $PM_{2.5}$ concentrations using a hidden Markov model with different emission distributions in Northern California, Science of the Total Environment 443(2013) 93-103.

[13] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Computation 9(8)(1997) 1735-1780.

[14] B. Bishoi, A. Prakash, V.-K. Jain, A comparative study of air quality index based on factor analysis and US-EPA methods for an urban environment, Aerosol & Air Quality Research 9(1)(2009) 1-17.

[15] A.-K. Gorai, Kanchan, A. Upadhyay, P. Goyal, Design of fuzzy synthetic evaluation model for air quality assessment, Environment Systems and Decisions 34(3)(2014) 456-469.

[16] Y. Xu, P. Du, J. Wang, Research and application of a hybrid model based on dynamic fuzzy synthetic evaluation for establishing air quality forecasting and early warning system: A case study in China, Environmental Pollution 223(2017) 435-448.

[17] G. Sarella, A.-K. Khambete, Ambient air quality analysis using air quality index—A case study of Vapi, International Journal for Innovative Research in Science & Technology 1(10)(2015) 68-71.

[18] M. Ruggieri, A. Plaia, F. Di Salvo, G. Agró, Functional principal component analysis for the explorative analysis of multisite–multivariate air pollution time series with long gaps, Journal of Applied Statistics 40(4)(2013) 795-807.

[19] S. Van Den Elshout, K. Léger, H. Heich, CAQI common air quality index - update with $PM_{2.5}$ and sensitivity analysis, science of the total environment (488-489)(2014) 461-468.

[20] B.-J. Kirenga, Q. Meng, F. Van Gemert, H. Aanyu-Tukamuhebwa, N. Chavannes, A. Katamba, G. Obai, T. Van der Molen, S. Schwander, V. Mohsenin, The state of ambient air quality in two Ugandan cities: a pilot cross-sectional spatial assessment, International journal of environmental research and public health 12(7)(2015) 8075-8091.

[21] B. Lv, B. Zhang, Y. Bai, A systematic analysis of $PM_{2.5}$ in Beijing and its sources from 2000 to 2012, Atmospheric Environment 124(B)(2016) 98-108.

[22] E. Cogliani, Air pollution forecast in cities by an air pollution index highly correlated with meteorological variables, Atmospheric Environment 35(16)(2001) 2871-2877.

[23] P. Hájek, V. Olej, Air Quality Indices and their Modelling by Hierarchical Fuzzy Inference Systems, WSEAS Transactions on Environment and Development 10(5)(2009) 661-672.

[24] M. Zickus, K. Kvietkus, Urban air pollution forecast based on the Gaussian and regression models, WIT Transactions on Ecology and the Environment 28(1998) 515-523.

[25] P.-G. Nieto, E.-F. Combarro, J.-J. del Coz Díaz, E. Montañés, A SVM-based regression model to study the air quality at local scale in Oviedo urban area (Northern Spain): A case study, Applied Mathematics and Computation 219(17)(2013) 8923-8937.

[26] P.-A. Kassomenos, A.-K. Paschalidou, A. Vlachogianni, One-day-ahead prediction of maximum carbon monoxide concentration in urban environments, Stochastic Environmental Research & Risk Assessment 27(2)(2013) 561-572.
[27] H. Peng, A.-R. Lima, A. Teakles, J. Jin, A.-J. Cannon, W.-W. Hsieh, Evaluating hourly air quality forecasting in Canada with nonlinear updatable machine learning methods, Air Quality Atmosphere & Health 10(2)(2017) 195-211.
[28] D. Jiang, Y. Zhang, X. Hu, Y. Zeng, J. Tan, D. Shao, Progress in developing an ANN model for air pollution index forecast, Atmospheric Environment 38(40)(2004) 7055-7064.
[29] K.-D. Karatzas, G. Papadourakis, I. Kyriakidis, Understanding and forecasting air pollution with the aid of artificial intelligence methods in Athens, Greece, in: C. Koutsojannis, S. Sirmakessis (Eds.), Tools and Applications with Artificial Intelligence. Studies in Computational Intelligence, vol 166, Springer, Berlin, Heidelberg, 2009 (pp. 37-50).
[30] A. Azid, H. Juahir, M.-T. Latif, S.-M. Zain, M.-R. Osman, Feed-forward artificial neural network model for air pollutant index prediction in the southern region of Peninsular Malaysia, Journal of Environmental Protection 4(12A)(2013) 1-10.
[31] I. Kyriakidis, K. Karatzas, G. Papadourakis, A. Ware, J. Kukkonen, Investigation and forecasting of the common air quality index in Thessaloniki, Greece, in: IFIP International Conference on Artificial Intelligence Applications and Innovations, 2012.
[32] D. Wang, S. Wei, H. Luo, C. Yue, O. Grunder, A novel hybrid model for air quality index forecasting based on two-phase decomposition technique and modified extreme learning machine, Science of the Total Environment 580(2017) 719-733.
[33] X. Feng, Q. Li, Y. Zhu, J. Hou, L. Jin, J. Wang, Artificial neural networks forecasting of $PM_{2.5}$ pollution using air mass trajectory based geographic model and wavelet transformation, Atmospheric Environment 107(2015) 118-128.
[34] A. Kurt, A.-B. Oktay, Forecasting air pollutant indicator levels with geographic models 3 days in advance using neural networks, Expert Systems with Applications 37(12)(2010) 7986-7992.
[35] A. Plaia, F. Di Salvo, M. Ruggieri, G. Agró, A Multisite-multipollutant air quality index, Atmospheric Environment 70(2013) 387-391.
[36] K. Moustris, I. Larissi, P.-T. Nastos, K. Koukouletsos, A.-G. Paliatsos, 24-Hours ahead forecasting of PM10 concentrations using artificial neural networks in the greater athens area, Greece, in: Proc. advances in meteorology, Climatology and Atmospheric Physics, 2012.
[37] K.-P. Moustris, I.-C. Ziomas, A.-G. Paliatsos, 3-Day-ahead forecasting of regional pollution index for the pollutants NO_2, CO, SO_2, and O_3 using artificial neural networks in Athens, Greece, Water Air & Soil Pollution 209(1-4)(2010) 29-43.
[38] D. Dunea, A. Pohoata, S. Iordache, Using wavelet-feedforward neural networks to improve air pollution forecasting in urban environments, Environmental Monitoring and Assessment: An International Journal 187(7)(2015) 1-16.
[39] A.-V. Beddows, N. Kitwiroon, M.-L. Williams, S.-D. Beevers, Emulation and sensitivity analysis of the community multiscale air quality model for a UK Ozone pollution episode, Environmental Science & Technology 51(11)(2017) 6229-6236.