ISSN: 2664-9640

Ecosphere pollutant prediction based on multimodal information fusion

Shengwei Wang[a,*], Yulin Zhan [b], Hao Ji [c] and Ping Li [d]

College of Computer Science and Engineering, Northwest Normal University, Lanzhou (730070), China.

[a]wangsw@nwnu.edu.cn* (corresponding author); [b]2019211778@nwnu.edu.cn; [c]2018221753@nwnu.edu.cn; [d]2019221819@nwnu.edu.cn

Abstract

With the rapid development of modern industry and agriculture, the problem of ecosphere pollution has become an urgent problem to be solved. However, the prediction of ecosphere pollutants, as the key content of ecosphere pollution protection, is often neglected. The key to ecosphere pollution prediction is to extract the interaction relationship between different modes of ecosphere pollutants and to integrate and predict different modes in the ecosphere. This paper proposes a FC-LSTM model based on fully connected neural network and long short-term memory network. The model obtains multi-layer representation and abstraction of data through learning, and then converts the data into high-level abstract features of deep networks, and uses high-level abstract features for ecological Circular contamination forecast. The advantage of this method is that it has good scalability and high accuracy, so the FC-LSTM neural network can accurately predict the content of pollutants in the ecosphere.

Keywords

Multimodal information fusion; ecosphere pollutant prediction; full connected neural; Long-short term neural network.

1. Introduction

Ecosphere is a complex system in which living matter and non-living matter blend with each other. Nitrogen, phosphorus, As, Cr, and Pb in ecosphere pollutants are all modalities of different forms of ecosphere pollutants [1]. Each mode of pollutants in the ecosphere can provide certain information for the other modes. Therefore, how to perform heterogeneous complementation of multi-modal data and mine the correlation between modes is a hot research topic in today's scientific community [2]. Multimodal fusion technology can well solve the complementary fusion between different modalities in the pollutant ecosystem, and further improve the effectiveness of data tasks by jointly learning the potential shared information of each modal data [3]. Since the multimodal data of ecosphere pollutants exist in different subspaces, the main goal of the multimodal fusion technology of the ecosphere is to reduce the distribution gap of the multiple subspaces as much as possible and maintain the integrity of the modal information [4].

At present, researches have been carried out on the multimodality of ecosphere pollutants, but there are still some problems. Among the model-independent fusion methods, the early fusion methods can extract the correlation between different modalities and different forms of features, but this method needs to extract a large number of features and is difficult to represent the time synchronization between multimodal features [5]. The later fusion first classifies different modalities and then fuses them. This method has a corresponding improvement in data asynchronous processing, but it is more difficult for low-level fusion of different modalities.

Hybrid fusion has the characteristics of feature fusion and decision layer fusion, but it is not feasible for model training due to the integration of early fusion and late fusion [6]. The multi-core learning method in the model-based fusion method can flexibly fuse data of different forms, but this method consumes a lot of memory, so it cannot be widely used [7]. The image model method adopts joint probability modeling to make the spatiotemporal structure of the data clearer, but the generalization ability of the model is not strong [8]. In the neural network method, the fully connected neural network can achieve a better learning effect, and it has good scalability and is widely used[9]. Han Guang et al. established an air quality prediction method based on Feedforward Neural Network (FNN). Experiments verify that the prediction accuracy of the FNN neural network model can reach 87.94% [10]. Liu Yingjun et al. compared meteorological data as the input of the LSTM neural network model and the SWAT model and found that the LSTM neural network model is better than the SWAT model in capturing the complex nonlinear relationship between various factors in the first n moments [11]. Tu Jichang et al. established a water quality prediction model based on GRU neural network. Experiments show that GRU neural network model has smaller RSME value and MAPE value than ARIMA model and SVR model, and can predict water quality more accurately [12].

The ecosphere is an open ecosystem, and its formation is the result of the long-term mutual checks and balances between the living and non-living worlds on the earth. How to scientifically manage and control ecosphere pollutants requires the integration and prediction of different modes of ecosphere pollutants. The multimodal information sources of ecosphere pollutants provide rich but redundant scene information, which is also accompanied by uncertainty. Therefore, this paper proposes an FC-LSTM model based on a fully connected neural network and a long short-term memory network to fuse and predict the multimodal information of the ecosystem. Information, through the LSTM neural network to accurately predict the multi-modal data of the ecosphere pollutants after fusion. In this paper, the inorganic pollutants $N_2O$and the heavy metal pollutants As are used as the research objects to conduct several experiments. The experimental results show that the FC-LSTM model can effectively integrate the multi-modal data of the ecosystem and make predictions.

2. Datasets and preprocessing

2.1. Overview of the study area

This paper selects the Daxia River area in Linxia City, Gansu Province as the research object. This area has various types of landforms, mature tributaries, abundant vegetation and biological distribution, and frequent humanistic and economic activities. Taking ecosphere pollutants (nitrogen, phosphorus, As, Cr, Pb), surrounding ecological vegetation, biodiversity index, environmental factors, meteorological factors, and human and social factors as input data, comprehensively and systematically integrate the multi-level ecological environment pollutants mode, and select inorganic pollutants N_2O and heavy metal As content as prediction objects.

2.2. Data preprocessing

In this paper, the abnormal values in the meteorological and field data collected in the Daxia River Basin of Linxia are processed, and the missing data and abnormal data are replaced with the data of the near time. The collected and processed data are standardized so that the data distribution conforms to the normal distribution, and the standardization method is as follows.

$$x^* = \frac{x-u}{\sigma} \tag{1}$$

Among them, x^* represents the standardized data, x represents the actual sampled data, μ represents the mean value, σ represents the variance.

ISSN: 2664-9640

3. Ecosphere pollutant multimodal fusion

3.1. Dimensionality reduction for multimodal data

This paper uses biological surveys, social surveys, geographic remote sensing technology and other means to comprehensively collect multimodal data of ecosphere pollutants. In order to more accurately screen the multimodal data related to ecosphere pollutants, the Pearson correlation coefficient method was used to screen the main influencing factors. The Pearson correlation coefficient can be calculated according to Equation 2, and some of the multimodal data correlations are shown in Figure 1.

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \tag{2}$$

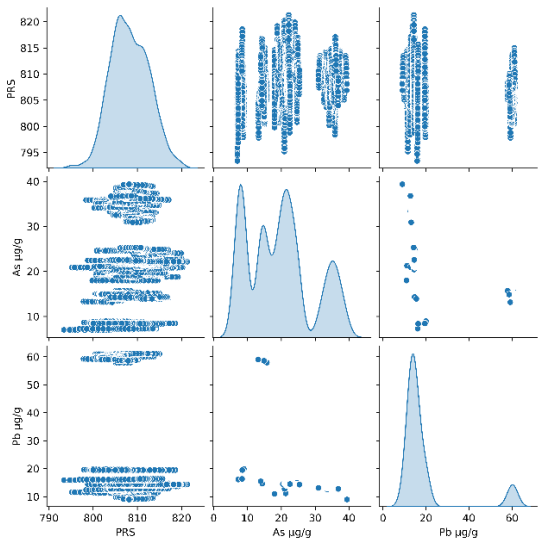

Figure 1. Ecosphere multimodal data correlation map

Where ***n*** is the total number of samples, ***X*** is the feature value, and ***Y*** is the target value.

3.2. Multimodal data fusion

A fully connected neural network is a layered computational model that learns multiple layers of abstract representations of data. A fully connected neural network uses a backpropagation algorithm to train its parameters, which can transform the raw input into an efficient task-specific representation. The multimodal fusion model of ecosphere pollutants based on fully connected neural network consists of input layer, hidden layer and output layer, each layer is composed of several neurons [16], and the model structure is shown in Figure 2. Specifically, the input layer input data after dimension reduction is multiplied by the weight matrix, and then the calculation result is input into the hidden layer, and the hidden layer data is multiplied by the weight and the result is input into the next hidden layer until the fused multimodal data is output.

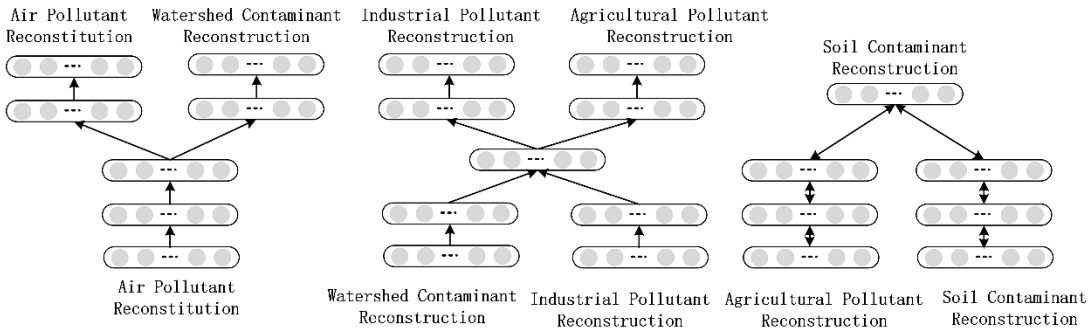

Figure2. Multimodal fusion graph based on fully connected neural network

3.3. Predicts the ecosphere pollutants

LSTM neural network is the earliest proposed gating-based recurrent neural network algorithm. LSTM neural network includes three basic gating units: input gate, forget gate and output gate [14]. The structure of LSTM neural network is shown in Figure 3. Specifically, the input gate g_i selectively retains the multi-modal input data of the ecosphere after fusion, and controls the influence of the current time step data on the output of pollutant content in the ecosphere. $\tilde{c}$ is obtained by nonlinear transformation of the current timestamp input x_t and the previous timestamp output h_{t-1}, g_i and $\tilde{c}$ are calculated as follows:

$$g_i = \sigma\left[w_i\left[h_{t-1}, x_t\right]\right] + bi \tag{3}$$

$$\tilde{c} = \tanh\left(w_c\left[h_{t-1}, x_t\right] + b_c\right) \tag{4}$$

The forget gate g_f selectively retains the internal state c_{t-1} of the ecosphere in the previous time step, and controls the impact of the internal state of the ecosphere on the content of pollutants in the ecosphere at the previous time step, and g_f is calculated as follows:

$$g_f = \sigma\left(w_f\left[h_{t-1}, x_t\right] + b_f\right) \tag{5}$$

Under the combined action of the input gate and the forgetting gate, the LSTM retains the partial information c_{t-1} of the previous time step and the screened current information $\tilde{c}_t$. At this time, the new state vector c_t is calculated as follows:

$$c_t = g_i\tilde{c}_t + g_f c_{t-1} \tag{6}$$

The output gate g_o makes a decision on the update of the content of pollutants in the internal state of the ecosphere, and g_o is calculated as follows:

$$g_o = \sigma\left(w_o\left[h_{t-1}, x_t\right] + b_o\right) \tag{7}$$

where $\boldsymbol{x_t}$ is the current time step input within the ecosphere pollutants, $\boldsymbol{w_c}$, $\boldsymbol{w_i}$, $\boldsymbol{w_f}$, $\boldsymbol{w_o}$ are the weight matrices corresponding to the gating. $\boldsymbol{b_c}$, $\boldsymbol{b_i}$, $\boldsymbol{b_f}$, $\boldsymbol{b_o}$ represents the offset matrix corresponding to the gating. Both the weight matrix and the offset matrix can be automatically optimized by the backward propagation algorithm. $\boldsymbol{\sigma}$ is a sigmoid function and ***tanh*** is an activation function with a value range of [-1, 1].

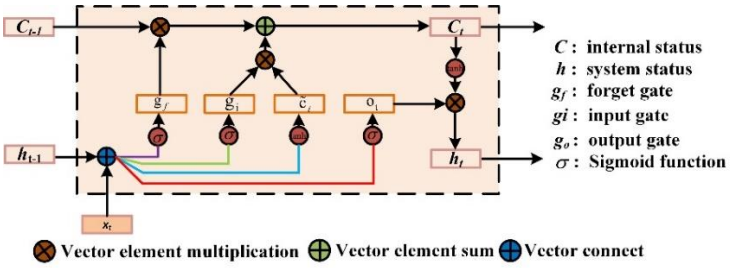

Figure3. LSTM neural network model

4. Experiment

4.1. Experiment environment

In this paper, the neural network framework is TensorFlow2.3.0, and python is used to implement the multimodal fusion prediction model of ecosphere pollutants.

Table1 Experiment parameters

Model	Hidden size	Batch size	Optimizer	Loss function
BP	32	32	Adam	MAE
LSTM	32	32	Adam	MAE
LAG	32	32	Adam	MAE

4.2. Model Evaluation Metrics

This paper uses Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) to evaluate the FC-LSTM model.

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|y_{(t)} - \hat{y}_{(t)}\right| \tag{8}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left|y_{(t)} - \hat{y}_{(t)}\right|^2} \tag{9}$$

Where $y(t)$ and $\hat{y}_{(t)}$ represent the real value and the predicted value at time $\boldsymbol{t}$ respectively, and $\boldsymbol{n}$ is the number of samples.

4.3. Experimental results

As shown in Figure 4, the FC-LSTM model has small errors in both MAE and RMSE evaluation indicators. And as the number of iterations increases, the error continues to decrease until it no longer changes. The experimental results show that the FC-LSTM model can fuse and accurately predict the multimodal data of ecosphere pollutants.

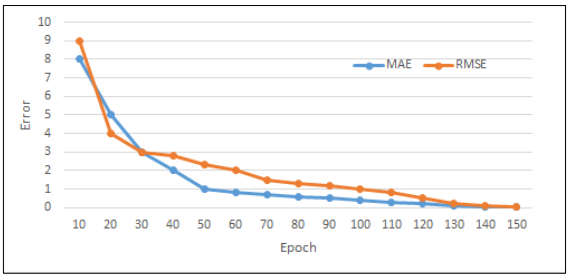

Figure4. FC-LSTM model prediction error

In the real data, ten inorganic pollutants N_2O and ten heavy metal pollutants As, respectively, were used to keep the experimental environment unchanged, and the FC-LSTM neural network

was used to simulate the real values. The experimental results are shown in Figure 5 and Figure 6.

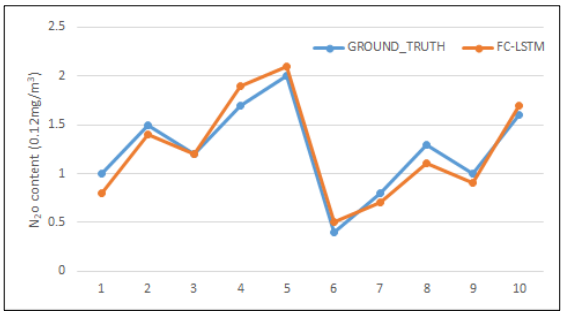

Figure5. Comparison of FC-LSTM predicted value and Ground_truth(N_2O)

Figure6. Comparison of FC-LSTM predicted value and Ground_truth(As)

It can be seen from Figure 5 and Figure 6 that the FC-LSTM model has a good simulation effect on both the inorganic pollutant $N_2O$and the heavy metal pollutant As.

4.4. Analyze

The multimodal data of ecosphere pollutants are fused together under the action of fully connected neural network. Eliminate some meaningless multimodal data while strengthening the modal data that are important for the prediction of ecosphere pollutants. The LSTM neural network controls the input and output of pollutants in the ecosphere in each time step through its excellent gating mechanism, so that the model can selectively remember historical information. Under the combined effect of multimodal fusion based on fully connected neural network and LSTM model, FC-LSTM can accurately predict multimodal pollutants in the ecosystem based on multimodal fusion.

5. Conclusion

The concentration changes caused by the transport of pollutants in the ecosphere have a great negative impact on personal health and social development. Multimodal data prediction of ecosphere pollutants requires a comprehensive and quantitative assessment to better simulate the impact of ecosphere pollutant multimodal data on the ecological environment. Collect different characteristics of pollutants in the ecosphere, and use neural network-based multimodal feature fusion prediction technology to integrate heterogeneous and disconnected multimodal data from the ecosphere to generate more reliable predictions. Ultimately, a more effective ecosphere environmental pollution control will be achieved.

References

[1] Kraemer G, Camps-Valls G, Reichstein M, et al. Summarizing the state of the terrestrial biosphere in few dimensions[J]. Biogeosciences, 2020, 17(9): 2397-2424.

[2] Jun H, Caiqing Z, Xiaozhen L, et al. Survey of research on multimodal fusion technology for deep learning[J]. Computer Engineering, 2020, 46(5): 1-11.

[3] Gao J, Li P, Chen Z, et al. A survey on deep learning for multimodal data fusion[J]. Neural Computation, 2020, 32(5): 829-864.

[4] Yoo J H, Kim Y, Kim J, et al. 3d-cvf: Generating joint camera and lidar features using cross-view spatial feature fusion for 3d object detection[C]//European Conference on Computer Vision. Springer, Cham, 2020: 720-736.

[5] Martínez H P, Yannakakis G N. Deep multimodal fusion: Combining discrete events and continuous signals[C]//Proceedings of the 16th International conference on multimodal interaction. 2014: 34-41.

[6] Murphy R R. Computer vision and machine learning in science fiction[J]. Science Robotics, 2019, 4(30): eaax7421.

[7] Selfe C L. Multi-Modal Composition[J]. Cresskill: Hampton, 2007.

[8] Peng F, McCallum A. Information extraction from research papers using conditional random fields[J]. Information processing & management, 2006, 42(4): 963-979.

[9] Wang W, Tran D, Feiszli M. What makes training multi-modal classification networks hard?[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 12695-12705.

[10] Han G, Xu L, Sun X, et al. A feedforward neural network modeling method with improved learning rates for API prediction[J]. computers and applied chemistry, 2016, 2: 147-151.

[11] Y . Liu, K. Wang and L. Li, Prediction of watershed pollutant flux based on Long Sho rtTerm Memory neural network, Journal of Hydroelectric Engineering 39(10) (2020), 72–81.

[12] T. Jichang, Y . Xueqin, C. Chaobo, G. Song, W. Jingcheng and S. Cheng, Water quality prediction model based on GRU hybrid network, in: 2019 Chinese Automation Congress (CAC), IEEE, 2019, pp. 1893–1898.

[13] Kumar S, Mishra P K, Kumar J. Synergistic use of artificial neural network for the detection of underground coal fires[J]. Combustion science and technology, 2017, 189(9): 1527-1539.

[14] Yan H, Qin Y, Xiang S, et al. Long-term gear life prediction based on ordered neurons LSTM neural networks[J]. Measurement, 2020, 165: 108205.