

## RESEARCH ARTICLE

# Monitoring and early warning of a metal mine tailings pond based on a deep learning bidirectional recurrent long and short memory network

Zhanjie Jing<sup>1,2,3,4\*</sup>, Xiaohong Gao<sup>1,4\*</sup>

**1** Department of Geographic Information Science, Qinghai Normal University, Xining, China, **2** Institute of Resources and Ecology, Yili Normal University, Yining, Yili Kazakh Autonomous Prefecture, China, **3** College of Biology and Geography Sciences, Yili Normal University, Yining, Yili Kazakh Autonomous Prefecture, China, **4** Qinghai Province Key Laboratory of Physical Geography and Environmental Process, College of Geographical Science, Qinghai Normal University

\* [qhnujzhj@163.com](mailto:qhnujzhj@163.com) (ZJJ); [xiaohonggao226@163.com](mailto:xiaohonggao226@163.com) (XHG)



## OPEN ACCESS

**Citation:** Jing Z, Gao X (2022) Monitoring and early warning of a metal mine tailings pond based on a deep learning bidirectional recurrent long and short memory network. PLoS ONE 17(10): e0273073. <https://doi.org/10.1371/journal.pone.0273073>

**Editor:** Ardashir Mohammadzadeh, University of Bonab, ISLAMIC REPUBLIC OF IRAN

**Received:** September 7, 2021

**Accepted:** August 3, 2022

**Published:** October 13, 2022

**Copyright:** © 2022 Jing, Gao. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its [Supporting Information](#) files.

**Funding:** We declare that the sources of funding for this work was supported by the Natural Science Foundation of Qinghai Provincial Science and Technology Department under Grant 2021-ZJ-913.

**Competing interests:** We declare that no competing interests exist.

## Abstract

The effective monitoring and early warning capability of metal mine tailings ponds can improve the associated safety risk management level. The infiltration line is an important core index of tailings pond stability. In this paper, a tailings pond monitoring and early warning system, which provides technical support for the design and daily management of tailings reservoir early warning systems, is constructed. Based on a deep learning bidirectional recurrent long and short memory network, an infiltration line prediction model with univariate input and an infiltration line prediction model with multivariate input are proposed. The data adopted are those from four monitoring points of the same cross-section at different positions and data from one adjacent internal lateral displacement and internal vertical displacement monitoring point. Using the adaptive moment estimation (Adam) optimization algorithm and the root mean square error (RMSE) model evaluation metric, the multilayer perceptron model, univariate input model, and multivariate input model are compared. This work shows that their RMSEs are 0.10611, 0.09966, and 0.11955, respectively.

## Introduction

Mineral resources are an important material basis for promoting economic and social development and an indispensable support for ensuring social stability and sustainable economic development. In the process of mining mineral resources, the treatment of tailings should avoid causing environmental problems, and at the same time, the redevelopment and utilization of tailings resources should be considered. The solid-liquid mixed storehouse formed by the concentrated storage of mineral waste residues is called a tailings pond [1]. Many serious accidents related to tailings ponds have occurred worldwide. The basic causes of tailings dam failure can be divided into flood overtopping, dam cracks, seepage damage and dam landslides. However, the failure of tailings dams is often caused by many factors, which are essentially due to the influence of the external environment, such as the increased load through tailings dams,

earthquakes, rainfall, floods and dam foundation settlement [2]. The stress field and seepage field of the tailings reservoir change, which leads to instability of the dam body. The infiltration line of the seepage field in tailings ponds is called the "lifeline" of tailings ponds, and the determination of the seepage field is the basis of studying dam failure in tailings ponds. The position of the infiltration line affects the stability of the dam slope [3]. The consolidation speed of tailings below the infiltration line is slow, and tailings close to saturation increase the weight of the dam body, thus reducing the shear strength and effective stress of the dam body. In addition, rainstorms, floods and drainage facility failures usually lead to an increase in the saturation line in tailings dams, which leads to seepage damage. For tailings ponds, when the deformation conditions caused by seepage are met, a piping effect occurs in tailings dams. The material properties of tailings change after the piping effect, which leads to an increase in permeability and a decrease in shear strength and deformation modulus. Eventually, the tailings pond collapses, and the tailings dam breaks. Because of the complicated geological conditions and inaccurate boundary conditions of tailings ponds, it is difficult to find the exact solution of the seepage field and stress field of tailings ponds through theoretical research.

Li et al. [4] evaluated tailings reservoir disasters by the dynamic hierarchical gray relational analysis method and established an evaluation index and dynamic early warning index. Li et al. [5] studied the safety monitoring and early warning of tailings by examining the spatial evolution process of sediment flow and simulated the dam failure process of tailings dams in three-dimensional space. Wang et al. [6] designed and implemented a tailings pond monitoring system based on the Internet of Things and realized the real-time collection of monitoring data. Recently, an increasing number of researchers have applied data-driven methods in risk prediction, such as the artificial neural network (ANN). Through machine learning, we can use more data information, including nonlinear, mutual relations, and even hidden information in imperceptible data. Through model training, the safety trend of tailings ponds can be predicted, and the safety assessment and risk prediction performance of tailings ponds can be greatly improved. With the development of 5G networks, artificial intelligence, big data and other technologies, industrial production and processing have become more intelligent. Increasingly more monitoring data are being collected. It is very difficult to automatically process a large amount of data through a shallow network processing method.

The deep learning method has a strong feature extraction performance. In the first stage, it reduces the requirement of data feature description and provides a new solution for processing massive data. Deep learning methods have achieved great breakthroughs and have been used as advanced models in the field of artificial intelligence (AI) [7], such as for image semantic analysis and recognition [8,9], natural language processing [10] and speech recognition [11]. Deep learning has been widely used in various fields, such as earthquake and emergency response [12–15], biomedicine [16–18], mechanical fault diagnosis management [19–21], public transportation [22,23], energy [24–26], novel electrodes for future devices [27], and tailings dam failure and risk management [28–31]. In deep learning, the bidirectional long and short memory network model can combine the historical state and current memory to address time series problems. It is mainly used to describe the relationship between current data and previous input data, and its memory is used to save the internal information of previous data. The bidirectional cyclic long-short memory network model can not only learn the shallow nonlinear network structure but also approximate complex functions, extract the essential features of input time series data, and remember information for a long time. On this basis, the following topics are studied in this paper:

1. A monitoring and early warning system is constructed for tailings ponds that integrates a deep learning bidirectional cyclic long and short memory network.

2. Based on a deep learning bidirectional recurrent long-short memory network, a tailings pond infiltration line prediction model is proposed that has a single-variable input and a multi-variable input.
3. Through experiments, the multilayer perceptron model is compared with the model based on a bidirectional recurrent long short-term memory network.

## Materials and methods

### Deep learning bidirectional recurrent long and short memory network

A recurrent neural network is the foundation of a bidirectional cyclic long and short memory network. The calculation of the cyclic neural network forms a directed graph, and the expansion calculation diagram of the training loss of the basic cyclic neural network is shown in Fig 1. For time step  $t$ ,  $h$  is the hidden state of the cyclic neural network,  $x$  is the input time series vector, and  $y^{\wedge}$  is the output vector of the cyclic neural network mapping the input vector of the  $x$  value.  $L$  is a measure of the loss function between each model output  $y^{\wedge}$  and the corresponding training target  $y$ . Comparing it with the target  $y$ , the bias vectors  $b$  and  $c$  and the weight matrices  $U$ ,  $W$ , and  $V$  are the links from the input to the hidden layer, the hidden layer to the hidden layer, and the hidden layer to the output layer, respectively. Its calculation expression is as follows: (1)–(3), where  $\theta$  is the parameter learning function with minimum

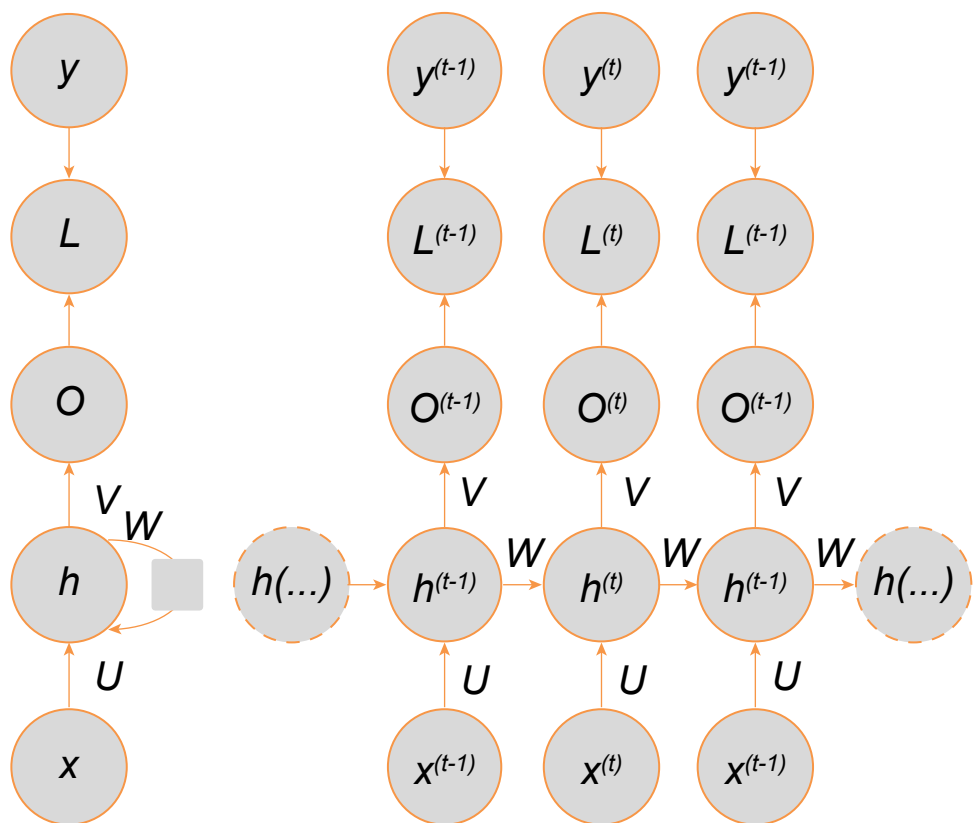


Fig 1. The computational graph to compute the training loss of the recurrent network.

<https://doi.org/10.1371/journal.pone.0273073.g001>

model loss and  $f$  is the nonlinear activation function.

$$h_t = f(Uh_t + Wx_t + b) \tag{1}$$

$$y^{(t)} = Vh_t + c \tag{2}$$

$$L = \theta(y - y^\wedge) \tag{3}$$

The long-term and short-term memory network, based on the simple cyclic neural network, specializes in cyclic information transmission by introducing a new internal state  $c_t \in \mathbb{R}^D$ . At the same time, it outputs information nonlinearly to the external state  $h_t \in \mathbb{R}^D$  of the hidden layer, which solves the problem of gradient explosion or disappearance. The internal state  $c_t$  is calculated by the following Formulas (4)–(5).

$$c_t = f_t \odot c_{t-1} + i_t \odot c_t^\wedge \tag{4}$$

$$h_t = o_t \odot \tanh(c_t) \tag{5}$$

Among them, the forget gate  $f_t \in [0,1]^D$ , the input gate  $i_t \in [0,1]^D$ , and the output gate  $o_t \in [0,1]^D$  are three gates to control the information transmission path.  $\odot$  is the product of vector elements, the internal state  $c_{t-1}$  is the memory unit at the previous moment, and the candidate state  $c_t^\wedge \in \mathbb{R}^D$  is the candidate state obtained through a nonlinear function:

$$c_t^\wedge = \tanh(W_c x_t + U_c h_{t-1} + b_c) \tag{6}$$

The bidirectional recurrent neural network (Bi-RNN) combines the RNN that moves forward and the RNN that moves backward in the time series. The basic Bi-RNN training loss expansion calculation diagram is shown in Fig 2. The hidden state  $h$  recursively propagates to the left in the time sequence, the hidden state  $g$  recursively propagates to the right in the time sequence, and their inputs are the same.

The hidden state at  $t$  is defined as  $h_t^{(1)}$  and  $g_t^{(2)}$  (see calculation expressions (7)–(9)), where  $o_t$  is the vector concatenation of  $h_t^{(1)}$  and  $g_t^{(2)}$ .

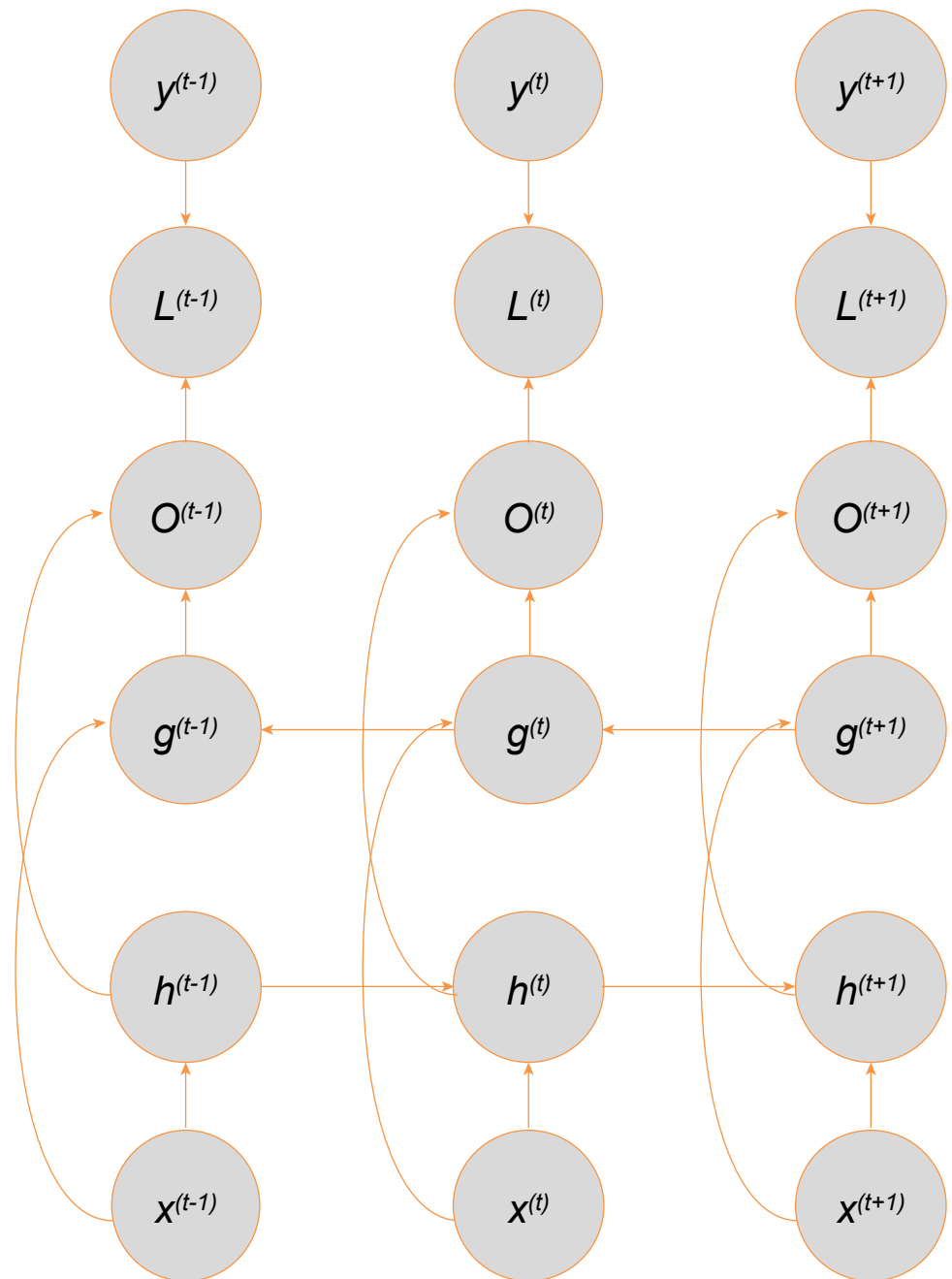
$$h_t^{(1)} = f(U^1 h_{(t-1)}^{(1)} + W^{(1)} x_t + b^{(1)}) \tag{7}$$

$$g_t^{(2)} = f(U^2 g_{(t+1)}^{(1)} + W^{(2)} x_t + b^{(2)}) \tag{8}$$

$$o_t = h_t^{(1)} \oplus g_t^{(2)} \tag{9}$$

### Construction of the tailings pond monitoring and early warning system

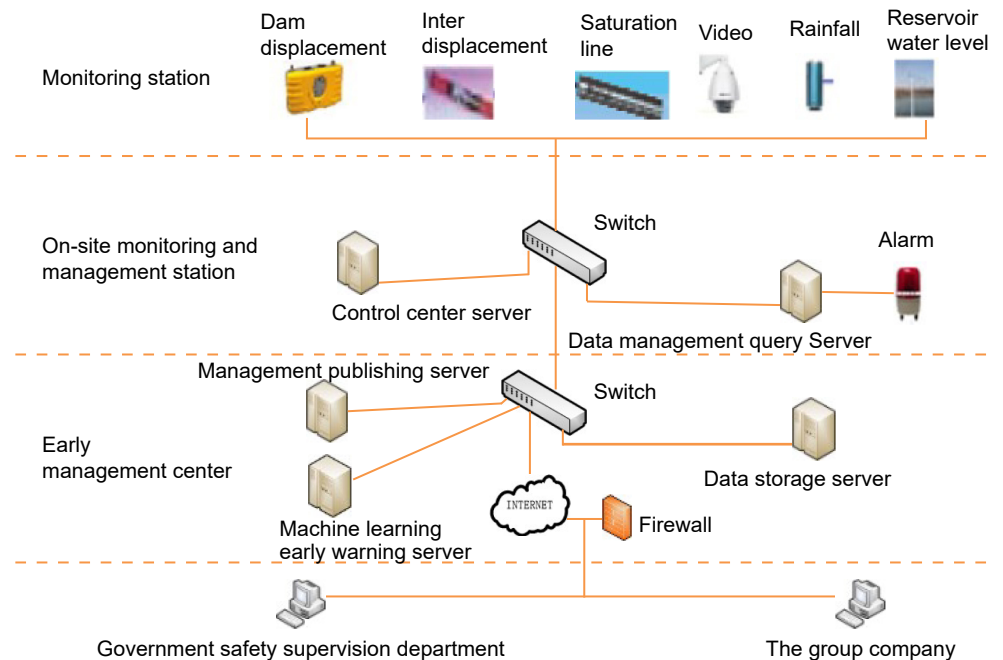
**System structure.** The system structure is shown in Fig 3, which adopts a three-level structure that includes a monitoring station, a monitoring management station and a prediction management center. As the first level of the overall architecture, the monitoring station is used to obtain the data of monitoring points in real time. The on-site monitoring and management station is the second level of the overall architecture, which is used for data collection, display, query, fault alarm, etc. The prediction management center station is the third level of the overall architecture and is used for 3D display, collection, storage, management, analysis, early warning and remote network release of the data.



**Fig 2. The computational graph to compute the training loss of the bidirectional recurrent network.**

<https://doi.org/10.1371/journal.pone.0273073.g002>

Global navigation satellite system (GNSS) ground receiving sensors are used for dam displacement monitoring. According to the requirements of the code, a monitoring section is set with a section spacing of 100~300 m, and multiple monitoring sections are set in the dam body. There are multiple monitoring points in each section, a monitoring reference point is set in the stable area near the duty room, and a reference point is set in the stable area at the dam tail. GNSS data processing software is deployed in the control center server, and all monitoring points and reference points of the tailings pond transmit GNSS signals received in real time with the control center through the network. The GNSS data processing software denoises and



**Fig 3. Monitoring and early warning system architecture.**

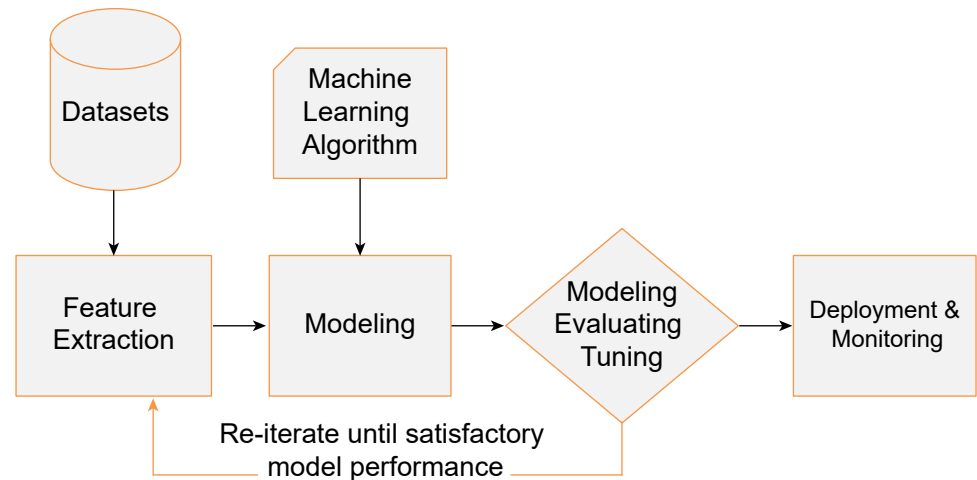
<https://doi.org/10.1371/journal.pone.0273073.g003>

solves the position information of each monitoring point in real time according to the model and analyzes it by associating the initial coordinates, thus obtaining the displacement variation of each monitoring point. The main functions of the GNSS data solution software are the remote management of GNSS receivers at monitoring points and control points, real-time and on-time GNSS raw data analysis and processing, independent ring network adjustment and data management.

Internal displacement monitoring and internal displacement monitoring points should be deployed in combination with surface displacement, and monitoring sections should be set at the dam crest at the initial stage of the tailings pond. Each section is provided with a plurality of monitoring vertical lines, and each vertical line is provided with a plurality of sensors to monitor the displacement in the downstream direction of the dam body axis.

*Infiltration line monitoring.* The infiltration line is monitored by using the intelligent vibrating string osmometer built in the piezometer. The osmometer sensor transmits data to the data acquisition unit through special hydraulic cables, converges them with other data through communication cables, and finally transmits all data to the monitoring center server through an industrial network. For reservoir water level monitoring, a radar level gauge is adopted, and monitoring points are set on the drainage wells in the reservoir area. The radar level gauge moves to monitor the change in the reservoir water level. For rainfall monitoring, a grid rain gauge with a heating module is adopted, and monitoring points are set at the duty room on site. Video surveillance deployment, using high-definition infrared network cameras, sets up monitoring points near the duty room of the reservoir area, the top of the initial dam, the drainage well, etc., to meet the night vision and key area monitoring requirements.

**Early warning management center.** The early warning management center of the system is deployed in the mine management center, centrally manages the monitoring of the tailings pond, sets up the storage server, machine learning early warning system and data management publishing server, and is responsible for and realizes the inquiry, analysis and early warning of the monitoring system. According to the monitoring items, it has good stability and



**Fig 4. Machine learning process.**

<https://doi.org/10.1371/journal.pone.0273073.g004>

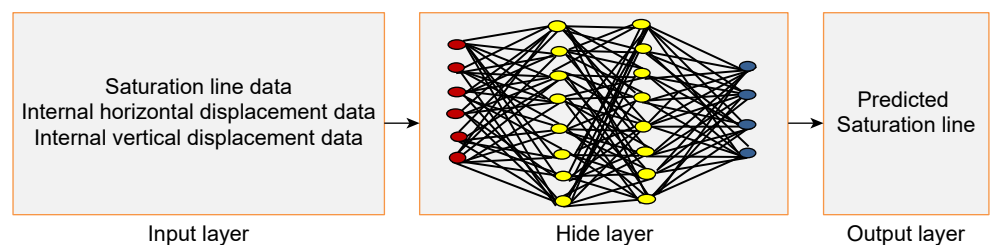
expandability, meets the needs of mine sites and remote management, and can realize networking with relevant government supervision departments.

The goal of the machine learning server is to automate the decision-making of tasks. The learning process is shown in Fig 4. Feature extraction involves extracting important features or attributes from the original data or creating new features from existing features. Modeling involves providing data features to machine learning methods or algorithms and training them, aiming at the evaluation index of the loss function, reducing errors and summarizing expressions learned from the data. Model evaluation and adjustment involve evaluation and testing on the validation dataset and gradual optimization to obtain the optimal model. The basic structure of the model is shown in Fig 5. The input data are the time series data of the infiltration line and internal displacement, the output layer is the prediction data of the infiltration line, and the hidden layer performs deep learning according to the algorithm. For deployment and monitoring, the selected model is deployed in production and continuously monitored according to its prediction and results.

## Experiment

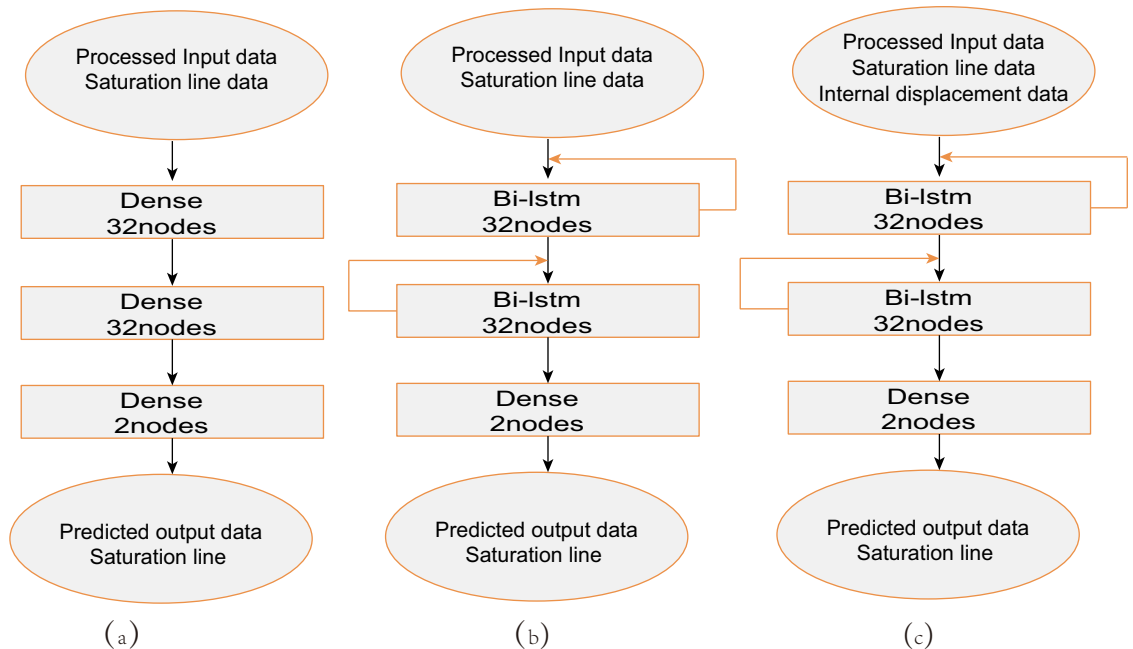
### Experimental environment

Model training machine hardware configuration: 8-core AMD Ryzen7 2.00 GHz processor and 16 GB RAM. Software: Python 3.8.5 and TensorFlow 2.3.0.



**Fig 5. Basic structure of the model.**

<https://doi.org/10.1371/journal.pone.0273073.g005>



**Fig 6.** Experimental Model Structure (a) Multi-Layer Perceptron Model, (b) Univariate Input Model, and (c) Multivariable Input Model.

<https://doi.org/10.1371/journal.pone.0273073.g006>

## Experimental design

The goal of the experiment is to answer two questions:

1. What is the loss difference of the model when different optimization algorithms are utilized?
2. How does the multilayer perceptron model compare with a model based on a bidirectional recurrent long short-term memory network?

To answer the first question, we use algorithms such as stochastic gradient descent (SGD), adaptive gradient (AdaGrad), root mean square prop (RMSprop), and adaptive moment estimation (Adam). For the second question, the multilayer perceptron model and the model based on the bidirectional recurrent long-short memory network are compared, the model structure diagram is shown in Fig 6. This model contains three layers. The multilayer perceptron model consists of all fully connected layers, with 32 nodes in the first layer, 32 nodes in the second layer, and 2 nodes in the third layer. Based on the bidirectional cyclic long-short memory network model, the first and second layers are bidirectional cyclic long-short memory network layers, each layer with 32 nodes, and the third layer is a fully connected layer with 2 nodes.

## Model evaluation index

The root mean square error (RMSE) was chosen as the evaluation index of model performance. The RMSE is expressed as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (10)$$



## Data preparation

The research data come from the monitoring system database of a metal mine tailings pond in western China. The monitoring equipment for the saturation line of the metal mine tailings pond adopts an intelligent vibrating wire sensor, and the internal displacement monitoring equipment adopts an intelligent inclinometer. According to the monitoring design of the tailings pond, the monitoring system of the whole tailings pond is composed of seven cross-sections, each of which has 3–4 monitoring points of the saturation line and one internal displacement monitoring point, and each monitoring point has different buried depths of the intelligent vibrating wire sensors and intelligent inclinometers.

In this study, the state of the infiltration line over the next six hours is predicted, mainly using the databases of four monitoring points of the infiltration line in cross-section 1 and the databases of the horizontal internal displacement and vertical internal displacement of adjacent internal displacement monitoring points, which are named DataSetI, DataSetII, DataSetIII, DataSetIV, DataSetV and DataSetVI. The monitoring data are collected every three hours from January 2019 to June 2020. There are 8 monitoring data points every day and 3,850 records in each dataset. The curve of the data collected by the monitoring points with time is shown in Fig 7, and some original data are shown in Table 1.

To improve the convergence speed and accuracy of the model, the collected data are normalized. The normalization method adopted is the MinMaxScalar method, and the calculation formula is as follows:

$$X_{\text{std}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \quad (11)$$

$$X_{\text{scaled}} = X_{\text{std}} * (\text{max} - \text{min}) + \text{min} \quad (12)$$

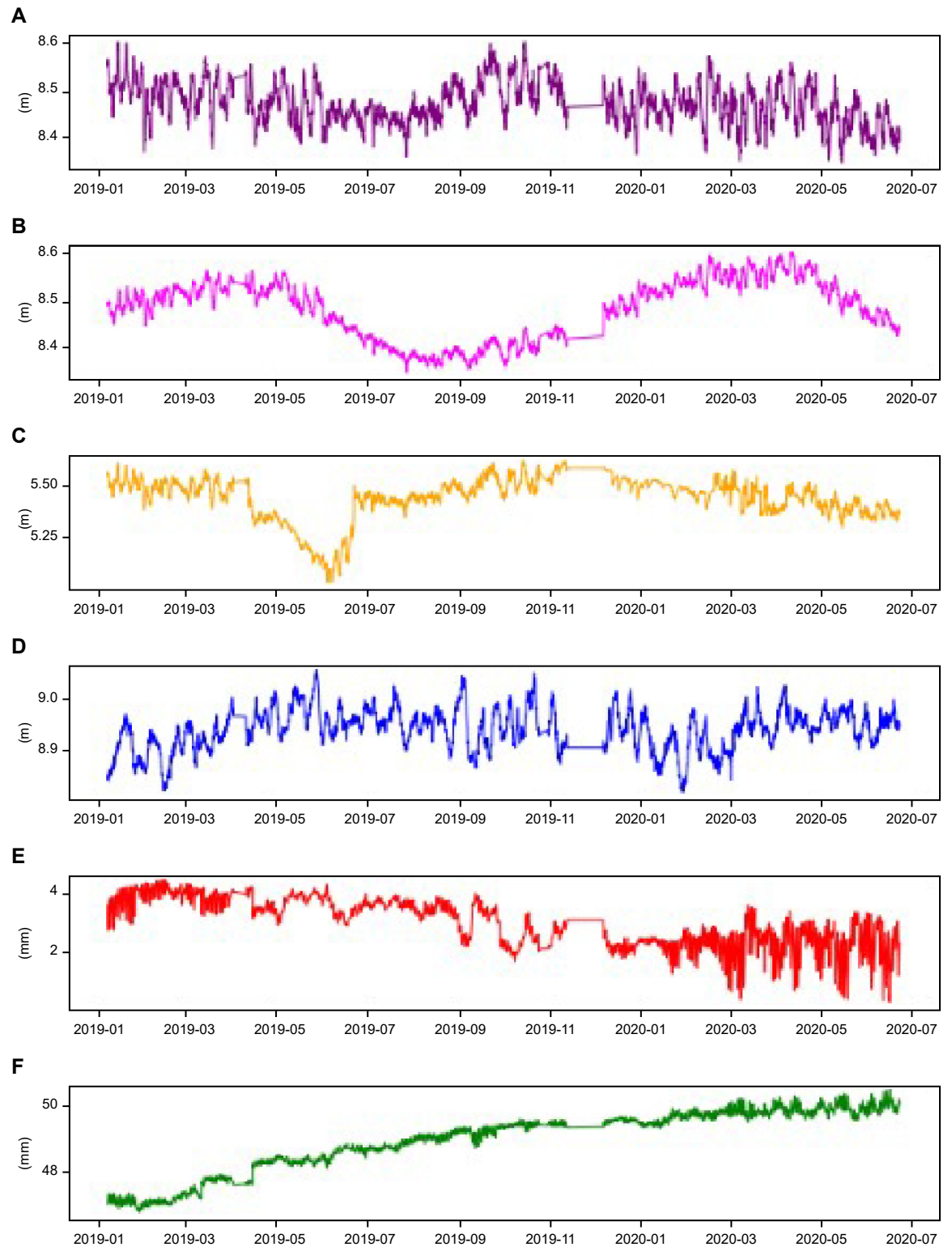
where  $X$  is the original saturation line data or internal displacement data;  $x_{\text{min}}$  and  $x_{\text{max}}$  are the maximum and minimum, respectively;  $\text{max}$  and  $\text{min}$  are the normalized characteristic ranges, which are 1 and 0 by default; and  $X_{\text{scaled}}$  is the normalized data.

## Model training

The model parameters depend on the input layer, hidden layer and output layer of the model. The dimension of the input layer depends on the feature dimension of the training set. In this study, there are one-dimensional and three-dimensional time series data, which correspond to the infiltration line, internal horizontal displacement and internal vertical displacement time series data. There are many neurons in the hidden layer. These neurons transform the input of the previous layer and then use the activation function to activate it and output it to the next layer. The output layer is the predicted target result.

Although deep learning is very powerful, it requires different optimization methods, such as stochastic gradient descent (SGD), adaptive gradient (AdaGrad), root mean square prop (RMSprop), and adaptive moment estimation (Adam). When the input sequence  $x_{1:T} = (x_1, \dots, x_T)$  of length  $T$  and the label sequence  $y_{1:T} = (y_1, \dots, y_T)$  constitute the training sample  $(x, y)$ , there is label supervision information  $y_t$  at time  $t$ , and the loss function is defined as:

$$L_t = L(y_t, g(h_t)) \quad (13)$$



**Fig 7. Saturation line height and internal displacement over time.**

<https://doi.org/10.1371/journal.pone.0273073.g007>

**Table 1. Part of the original saturated line and internal displacement data.**

Acquisition time	DataSetI (m)	DataSetII (m)	DataSetIII (m)	DataSetIV (m)	DataSetV (mm)	DataSetVI (mm)
2019-01-08 01:06:00	8.5534	7.6418	5.5572	8.8431	3.5777	47.0169
2019-01-08 04:06:00	8.5633	7.6418	5.5549	8.8508	3.2718	47.0885
2019-01-08 07:06:00	8.5668	7.6443	5.5450	8.8525	3.0608	47.1234
2019-01-08 10:06:00	8.5590	7.6350	5.5519	8.8525	2.8494	47.2027

<https://doi.org/10.1371/journal.pone.0273073.t001>

L is the differentiable loss function,  $g(h_t)$  is the output at time t, and the loss function of the entire sequence is:

$$L = \sum_{t=1}^t L_T \tag{14}$$

The mean squared error (MSE) loss function can be expressed mathematically as:

$$L_{mse} = (y^t - g(h_t))^2 \tag{15}$$

The gradient of the loss function L of the entire sequence with respect to the weight parameter U is the sum of the partial derivatives of the loss  $L_t$  with respect to the parameter U at each moment. The expression is as follows:

$$\frac{\partial L}{\partial U} = \sum_{t=1}^T \frac{\partial L_t}{\partial U} \tag{16}$$

Calculate the partial derivative  $\frac{\partial L_t}{\partial U}$  because the weight parameter U and the net input of the hidden layer at each time  $k$  ( $1 \leq k \leq t$ ) are:

$$z_k = U h_{k-1} + W x_k + b \tag{17}$$

Therefore, the gradient of the loss function  $L_t$  with respect to the parameter  $u_{ij}$  at time t is:

$$\frac{\partial L_t}{\partial u_{ij}} = \sum_{k=1}^t \frac{\partial^+ z_k}{\partial u_{ij}} \frac{\partial L_t}{\partial z_k} \tag{18}$$

where  $\frac{\partial^+ z_k}{\partial u_{ij}} = [0, \dots, [h_{k-1}], \dots, 0]$  and the error term  $\delta_{t,k} = \frac{\partial L_t}{\partial z_k}$  is defined as the derivative of the loss at time t with respect to the net input  $z_k$  of the hidden neural layer at time k. Then, when  $1 \leq k \leq t$ ,

$$\delta_{t,k} = \frac{\partial L_t}{\partial z_k} = \frac{\partial h_k}{\partial z_k} = \frac{\partial z_{k+1}}{\partial h_k} \frac{\partial L_t}{\partial z_{k+1}} \tag{19}$$

Substituting Formulas (18) and (17) into Formula (16) yields a matrix of the form:

$$\frac{\partial L_t}{\partial U} = \sum_{k=1}^t \delta_{t,k} h_{k-1}^T \tag{20}$$

The process of model training is to apply an optimization algorithm, iterate the model parameters, gradually improve the loss of the model and index, minimize the loss function, and terminate the model iteration to obtain the optimal parameters learned from the model training. The model training is time consuming, sometimes requiring hours or even weeks. The efficiency of model training is related to the advantages and disadvantages of the optimization algorithm. Understanding the optimization algorithm is beneficial to targeted model parameter adjustment, which makes the model perform better. The training models in this

experiment include a multivariate input infiltration line prediction model and a univariate input infiltration line prediction model.

**Algorithm 1. Back-propagation algorithm stochastic gradient descent optimization training.**

Input: training set  $D = (x^{(n)}y^{(n)})_{n=1}^N$ , validation set  $V$ , learning rate  $\alpha$ , regularization system  $\lambda$ , number of network layers  $L$ , number of neurons  $M_l$ ,  $1 \leq l \leq L$ .  
 1: Random initialization  $W, b$ ;  
 2: Repeat  
 3: Randomly reorder the samples in the training set;  
 4: For  $n = 1 \dots N$  do  
 5: Select samples  $(x(n), y(n))$  from training set  $D$ ;  
 6: Feed-forward calculate the net input  $z(l)$  and activation value  $a(l)$  of each layer until reaching the last layer;  
 7: Inversely calculate the error  $\delta(l)$  of each layer;  
 8:  $\forall l, \frac{\partial L(y^{(n)}, y^{\wedge(l)})}{\partial w^{(l)}} = \delta^{(l)}(\alpha^{(l-1)})^T$ ;  
 9:  $\forall l, \frac{\partial L(y^{(n)}, y^{\wedge(l)})}{\partial b^{(l)}} = \sigma^{(l)}$ ;  
 10:  $W^{(l)} \leftarrow W^{(l)} - \alpha(\sigma^{(l)}(a^{(l-1)})^T + \lambda W^{(l)})$ ;  
 11:  $b^{(l)} \leftarrow b^{(l)} - \alpha \delta^{(l)}$ ;  
 12: End  
 13: Until the error rate of the deep Bi-LSTM network model on verification set  $V$  no longer decreases;  
 Output:  $W, b$

The parameter update difference  $\Delta\theta_t$  of the adaptive moment estimation algorithm (Adam) is calculated using the Formulas (21)–(23), where  $\alpha$  is the learning rate,  $\beta_1$  and  $\beta_2$  are the decay rates,  $\epsilon$  is a very small constant to maintain numerical stability, and  $M_t^\wedge$  is the modified first-order moment bias, and  $G_t^\wedge$  is the corrected second moment bias.

$$M_t^\wedge = \frac{M_t}{1 - \beta_1^t} \tag{21}$$

$$G_t^\wedge = \frac{G_t}{1 - \beta_2^t} \tag{22}$$

$$\Delta\theta_t = -\frac{\alpha}{\sqrt{G_t^\wedge + \epsilon}} M_t^\wedge \tag{23}$$

The univariate input infiltration line prediction model uses DataSetII for training the input data. For example, Algorithm 1 depicts SGD random gradient descent optimization training. Train the model 10,000 times.

The training input data of the multivariate input model adopted include DataSetII, DataSetV, and DataSetVI (i.e., the input data are the saturation line, internal horizontal displacement and internal vertical displacement). The univariate input infiltration line prediction model and the multivariate input infiltration line prediction model adopted are the Adam optimization method and the same evaluation standard. The model was trained 10,000 times.

**Results and discussion**

The univariate input infiltration line prediction model of the deep learning bidirectional cyclic long and short memory network is used to perform prediction on DataSetI, DataSetII, DataSetIII and DataSetV, as shown in Table 2, where the prediction root mean square error (RMSE) is used to compare performance. The Adam optimization algorithm has the lowest RMSE

**Table 2. Prediction RMSE of different optimized methods.**

Dataset	SGD	AdaGrad	RMSprop	Adam
I	0.04711	0.04738	0.04743	0.04688
II	0.04726	0.04844	0.04849	0.04599
III	0.07602	0.10724	0.10762	0.06145
IV	0.10902	0.11966	0.11983	0.09966

<https://doi.org/10.1371/journal.pone.0273073.t002>

(approximately 0.046) among these optimization methods, while the RMSprop optimization algorithm has the highest prediction RMSE (approximately 0.119).

As shown in Table 3, the multivariate input infiltration line prediction model and multi-layer perceptron model carry out predictions on DataSetI, DataSetII, DataSetIII and DataSetV. Comparing the loss between the multilayer perceptron model, multivariate input model and univariate input model, the multivariate input infiltration line model is slight worse. The RMSE of prediction is basically satisfactory and can provide some decision support.

This article fuses deep learning technology in the construction of a tailings pond early warning system. At present, many tailings pond early warning systems compare with the threshold after real-time data collection for early warning, which is somewhat different from the data-driven early warning method. The early warning indicators of tailings ponds include dam body displacement, internal displacement, infiltration line, reservoir water level, rainfall, and infiltration line. The prediction index of the early warning model fused with deep learning in this paper is the infiltration line. The experimental data adopted are the same cross-sectional infiltration line monitoring point data and adjacent internal displacement data. Although there are certain limitations, the fusion of multisource data of the infiltration line and internal displacement is realized, and the trained model can be migrated to related cross-sections to improve the efficiency of model training as a whole.

## Conclusions

In this paper, a method for constructing the monitoring and early warning system of tailings reservoirs that includes the infiltration line, dam displacement, internal displacement, reservoir water level, rainfall, video, etc., is introduced, and an infiltration line prediction model of a bidirectional recurrent long and short memory network is proposed, which provides technical support for the design and daily management of monitoring and early warning systems of tailings reservoirs.

The tailings pond monitoring and early warning system offers more real-time response and intelligence, and the data-driven tailings pond risk early warning method has certain applicability. In the early warning of issues related to the tailings reservoir infiltration line, comparing the multilayer perceptron model, the univariate input model, and the multivariate input model, their RMSEs are 0.10611, 0.09966, and 0.11955, respectively. The data-driven early warning of tailings pond risk, integrating monitoring indicators such as dam body displacement, internal displacement, the infiltration line, the reservoir water level, and rainfall, is conducive to further risk evaluation.

**Table 3. Prediction RMSE of different models.**

Model	DataSetI	DataSetII	DataSetIII	DataSetIV
Multilayer perceptron model	0.04688	0.04658	0.07486	0.10611
Univariate input model	0.04688	0.04599	0.06145	0.09966
Multivariable input model	0.04692	0.04871	0.10855	0.11955

<https://doi.org/10.1371/journal.pone.0273073.t003>

## Supporting information

### S1 Dataset.

(CSV)

### S1 File.

(PDF)

## Author Contributions

**Conceptualization:** Zhanjie Jing, Xiaohong Gao.

**Data curation:** Zhanjie Jing, Xiaohong Gao.

**Formal analysis:** Zhanjie Jing.

**Funding acquisition:** Xiaohong Gao.

**Investigation:** Zhanjie Jing.

**Methodology:** Zhanjie Jing, Xiaohong Gao.

**Project administration:** Xiaohong Gao.

**Resources:** Xiaohong Gao.

**Software:** Zhanjie Jing, Xiaohong Gao.

**Visualization:** Xiaohong Gao.

**Writing – original draft:** Zhanjie Jing, Xiaohong Gao.

**Writing – review & editing:** Zhanjie Jing, Xiaohong Gao.

## References

1. Yan H, He F, Xu T. Study of double-cable-truss controlling system for large section coal roadway of deep mine and its practice. *Chin J Rock Mech Eng.* 2012; 31: 2248–2257.
2. Bowker L.N., Chambers D.M. The risk, public liability, and economics of tailings storage facility failures. *Earthwork Act.* 2015: 1–56.
3. Dong L, Sun D, Li X. Theoretical and case studies of interval nonprobabilistic reliability for tailing dam stability. *Geofluids.* 2017; 2017: 1–11.
4. Li W, Ye Y, Hu N, Wang X, Wang Q. Real-time warning and risk assessment of tailings dam disaster status based on dynamic hierarchy-grey relation analysis. *Complexity.* 2019; 2019: 1–14.
5. Li S, Yuan L, Yang H, An H, Wang G. Tailings dam safety monitoring and early warning based on spatial evolution process of mud-sand flow. *Saf Sci.* 2020; 124: 104579.
6. Wang L, Yang X, He M. Research on safety monitoring system of tailings dam based on internet of things. *IOP Conf Ser Mater Sci Eng.* 2018; 322: 052007.
7. Shakirov VV, Solovyeva KP, Dunin-Barkowski WL. Review of state-of-the-art in deep learning artificial intelligence. *Opt Mem Neural Netw.* 2018; 27: 65–80.
8. Lyu Y, Chen J, Song Z. Image-based process monitoring using deep learning framework. *Chemom Intell Lab Syst.* 2019; 189: 8–17.
9. Yang X, Zhou P, Wang M. Person reidentification via structural deep metric learning. *IEEE Trans Neural Netw Learn Syst.* 2019; 30: 2987–2998. <https://doi.org/10.1109/TNNLS.2018.2861991> PMID: 32175851
10. Morchid M. Parsimonious memory unit for recurrent neural networks with application to natural language processing. *Neurocomputing.* 2018; 314: 48–64.
11. Liu ZT, Xie Q, Wu M, Cao WH, Mei Y, Mao JW. Speech emotion recognition based on an improved brain emotion learning model. *Neurocomputing.* 2018; 309: 145–156.
12. Zhang R, Chen Z, Chen S, Zheng J, Büyükköztürk O, Sun H. Deep long short-term memory networks for nonlinear structural seismic response prediction. *Comput Struct.* 2019; 220: 55–68.

13. Wang KX, Huang QH, Wu SH. Application of long short-term memory neural network in geoelectric field data processing. *Chin J Geophys*. 2020; 63: 3015–3024.
14. Xi X, Huang JQ. Location and imaging of scatterers in seismic migration profiles based on convolution neural network. *Chin J Geophys*. 2020; 63: 687–714.
15. Cortez B, Carrera B, Kim YJ, Jung JY. An architecture for emergency event prediction using LSTM recurrent neural networks. *Expert Syst Appl*. 2018; 97: 315–324.
16. Maragatham G, Devi S. LSTM model for prediction of heart failure in big data. *J Med Syst*. 2019; 43: 111. <https://doi.org/10.1007/s10916-019-1243-3> PMID: 30888519
17. Riordon J, Sovilj D, Sanner S, Sinton D, Young EWK. Deep learning with microfluidics for biotechnology. *Trends Biotechnol*. 2019; 37: 310–324. <https://doi.org/10.1016/j.tibtech.2018.08.005> PMID: 30301571
18. Reddy BK, Delen D. Predicting hospital readmission for lupus patients: an RNN-LSTM-based deep-learning methodology. *Comput Biol Med*. 2018; 101: 199–209. <https://doi.org/10.1016/j.compbiomed.2018.08.029> PMID: 30195164
19. Wang X, Wu J, Liu C. Exploring LSTM based recurrent neural network for failure time series prediction. *J Beijing Univ Aeronaut Astronaut*. 2018; 44: 772–784.
20. Jia F, Lei Y, Guo L, Lin J, Xing S. A neural network constructed by deep learning technique and its application to intelligent fault diagnosis of machines. *Neurocomputing*. 2017; 272: 619–628.
21. Khan S, Yairi T. A review on the application of deep learning in system health management. *Mech Syst Signal Process*. 2018; 107: 241–265.
22. Petersen NC, Rodrigues F, Pereira FC. Multi-output bus travel time prediction with convolutional LSTM neural network. *Expert Syst Appl*. 2019; 120: 426–435.
23. Yang B, Sun S, Li J, Lin X, Tian Y. Traffic flow prediction using LSTM with feature enhancement. *Neurocomputing*. 2019; 332: 320–327.
24. Geng Z, Wang H, Fan M, Lu Y, Nie Z, Ding Y, et al. Predicting seismic-based risk of lost circulation using machine learning. *J Pet Sci Eng*. 2019; 176: 679–688.
25. Torres JM, Aguilar RM, Zuñiga-Meneses KV. Deep learning to predict the generation of a wind farm. *J Renew Sustain Energy*. 2018; 10: 013305.
26. Wei YZ, Xu XN. Ultra-short-term wind speed prediction model using LSTM networks. *J Electron Meas Instrum*. 2019; 33: 64–71.
27. Liu C, Li Q, Wang K. State-of-charge estimation and remaining useful life prediction of supercapacitors [J]. *Renewable and Sustainable Energy Reviews*, 2021, 150: 111408.
28. Che D, Liang A, Li X, Ma B. Remote sensing assessment of safety risk of iron tailings pond based on runoff coefficient. *Sensors (Basel, Switzerland)*. 2018; 18: 4373. <https://doi.org/10.3390/s18124373> PMID: 30544894
29. Li J, Chen H, Zhou T, Li X. Tailings pond risk prediction using long short-term memory networks. *IEEE Access*. 2019; 7: 182527–182537.
30. Seyedashraf O, Rezaei A, Akhtari AA. Dam break flow solution using artificial neural network. *Ocean Eng*. 2017; 142: 125–132.
31. Hooshyaripor F, Tahershamsi A, Golian S. Application of copula method and neural networks for predicting peak outflow from breached embankments. *J Hydro Environ Res*. 2014; 8: 292–303.