Research on Fertility Policy Under the Background of Three-Child Policy

Yingxiang Ma[1,#], Jia Zhang[2,#]
[1]Beijing Forestry University, School of materials science and technology, Beijing, 100091 China
[2]Beijing Forestry University, School of landscape architecture, Beijing, 100091 China

*Corresponding author: vj162774@163.com

#These authors contributed equally.

Abstract: Reasonable analysis and formulate relevant policy to control population stable growth is of great significance to social development and economic development in this paper, combining with the relevant data search conducted fertility policy mathematical modeling and analysis, introduces both short-term and long-term memory neural networks with time characteristics (LSTM), considering the population forecast may be is more sensitive to high-dimensional feature, The multi-layer stacked long and short-term memory neural network (DLSTM) was used as the prediction model, and the multi-layer stacked structure significantly improved the robustness and accuracy of prediction.

Keywords: Birth policy, long and short-term memory neural network, artificial neural network

1. Introduction

The number and structure of population are important factors affecting social and economic development.China has experienced the "family planning", "universal two-child policy", and then the implementation of the "three-child policy", which are the adjustment of China's population development and change trend.Therefore, it is of great significance for social and economic development to rationally analyze and formulate relevant fertility policies to control stable population growth.Therefore, our team conducted mathematical modeling and analysis of the birth policy based on searching relevant data. In this paper, a mathematical model was established based on the age structure of China's population to predict China's population situation in the next 10 years after the implementation of the three-child policy, and the impact of the implementation of the "double reduction" policy on China's new born population was analyzed[1].

2. Demographic data analysis

The number of births in China fell by 580,000 to 14.65 million in 2019, after falling by 2 million in 2018.Since the founding of The People's Republic of China, China has witnessed three rounds of baby boom, with an average annual average of 21 million from 1950 to 1958, 26.28 million from 1962 to 1975, 22.46 million from 1981 to 1994, and then gradually declining to about 16 million from 2003 to 2012, with 16.35 million in 2012.China's total fertility rate fell from about six before the 1970s to about two in 1990 and then to about 1.5 after 2010.A fourth baby boom was supposed to have occurred after 2010, but it disappeared because of a long period of strict family planning.Against the above background, the one-child policy was finally eased. At the end of 2012, the central Government decided to implement the one-child policy, but the effect was not as expected. From 2013 to 2015, the birth population was 16.4 million, 16.87 million and 16.55 million respectively.At the end of 2015, the central government decided to allow all children to have two children. In 2016, the number of births reached 17.86 million, the highest since 2000.However, it dropped to 17.25 million in 2017, another 2 million to 15.23 million in 2018, and 14.65 million in 2019. Further, based on national public data[2], factors influencing the future population structure can be obtained, as shown in Table 1.

Table 1 Characteristics of population impact

The population structure	Objective factors			Policy of public opinion		The economic development	
Factor importance	population	New number	Sex ratio	Policy initiatives	Public opinion	Birth allowance	Economic level
symbol	X_1	X_2	X_3	X_4	X_5	X_6	X_7

We have obtained a total of 7 influencing factors of population structure above, and the modeling of LSTM is given below.

3. Establishment of population prediction model

3.1 Basic structure of LSMT model

LSTM is an improvement of recurrent neural network RNN, which can extract long and short term effects of sequences (especially time series) [3].It has a variety of internal control gates, including input gate, forgetting gate and output gate. Its neuron structure is shown in Figure 1[4].

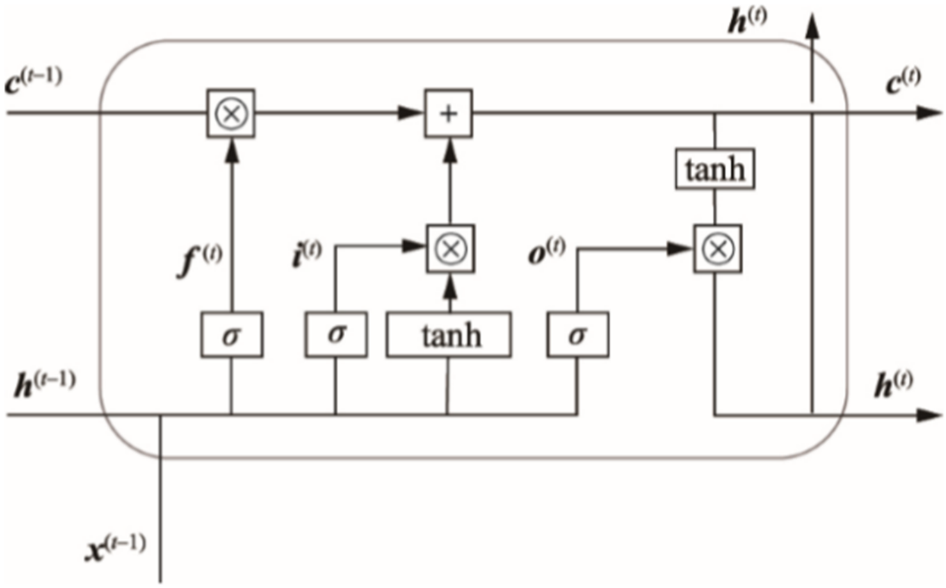

Figure 1 Neuronal structure

T is the timestamp of the sequence to be predicted, where I (T) is the input gate, equivalent to the input layer of the neural network, where the feature data to be processed is stored.F (t) is the forgetting gate[5], which is used to simulate the neural network's forgetting of transactions at a certain stage in the past (represented here as characteristic information).O (t) is the output gate, which is similar to the output layer of neural network[6].

Due to the introduction of these three gating mechanisms, the LSTM has more real neural network characteristics, so the fitting effect is better.In addition, LSTM can learn the dependency relationship with long span, and it has very good robustness to the information that needs to be trained to predict the population number across 10 years. Meanwhile, it does not have the gradient disappearance and gradient explosion that occur in general networks[7].

The structure of DLSTM neural network is shown in Figure 2.In order to further improve the accuracy of model prediction and the stability and robustness of small perturbations to variable features[8], DLSTM neural network stacked with three-layer LSTM network is used for modeling, which can mine the sequence information of deeper and more complex action system, and the three-layer network interaction has higher robustness.Figure X is the structure of DLSTM neural network.It can be seen from the figure that each hidden layer is the input layer of the next layer, let the number of layers be L, then l, update mode of layer DLSTM can be expressed as[9]:

$$i_l^{(t)} = \sigma(W_{i,l}h_{l-1}^{(t)} + V_{i,l}h_l^{(t-1)} + b_{i,l}); \tag{1}$$

$$f_l^{(t)} = \sigma(W_{f,l}h_{l-1}^{(t)} + V_{f,l}h_l^{(t-1)} + b_{f,l}); \tag{2}$$

$$o_l^{(t)} = \sigma(W_{o,l}h_{l-1}^{(t)} + V_{o,l}h_l^{(t-1)} + b_{o,l}); \tag{3}$$

$$c_l^{(t)} = f_l^{(t)} \otimes c_l^{(t-1)} + i_l^{(t)} \otimes \tanh(W_{c,l}h_{l-1}^{(t)} + V_{c,l}h_l^{(t-1)} + b_{c,l}); \tag{4}$$

$$h_l^{(t)} = o_l^{(t)} \otimes \tanh(c_l^{(t)}). \tag{5}$$

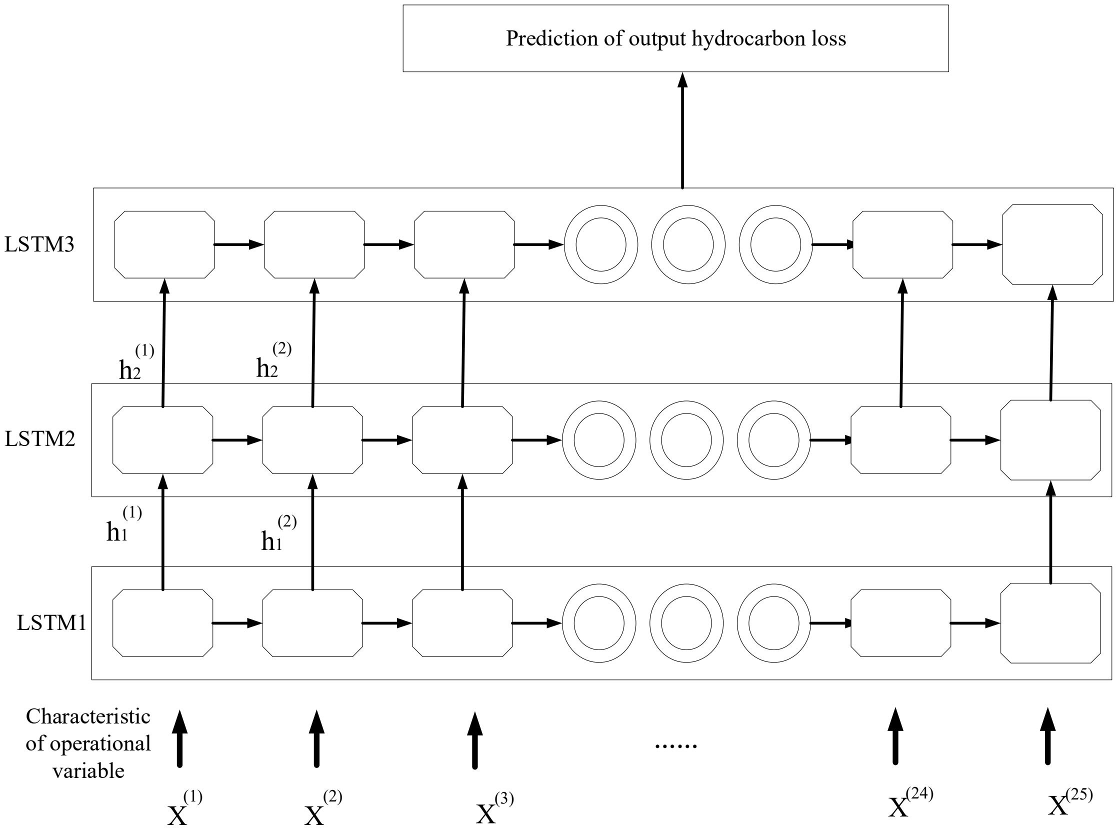

Figure 2 DLSTM neural network structure

3.2 Prediction model building based on DLSTM neural network

The complete DLSTM has an embedding layer, three LSTM layers and a logistic regression layer stacked on top of each other.In the model, DLSTM neural network is set as 3 layers, and the number of neurons is 7, 50 and 1 respectively.Where s is the number of iterations, s is the set total number of iterations, and E(s) is the value of softmax regression function for the s-th time.The population prediction model based on DLSTM neural network is shown in Figure 3[10].

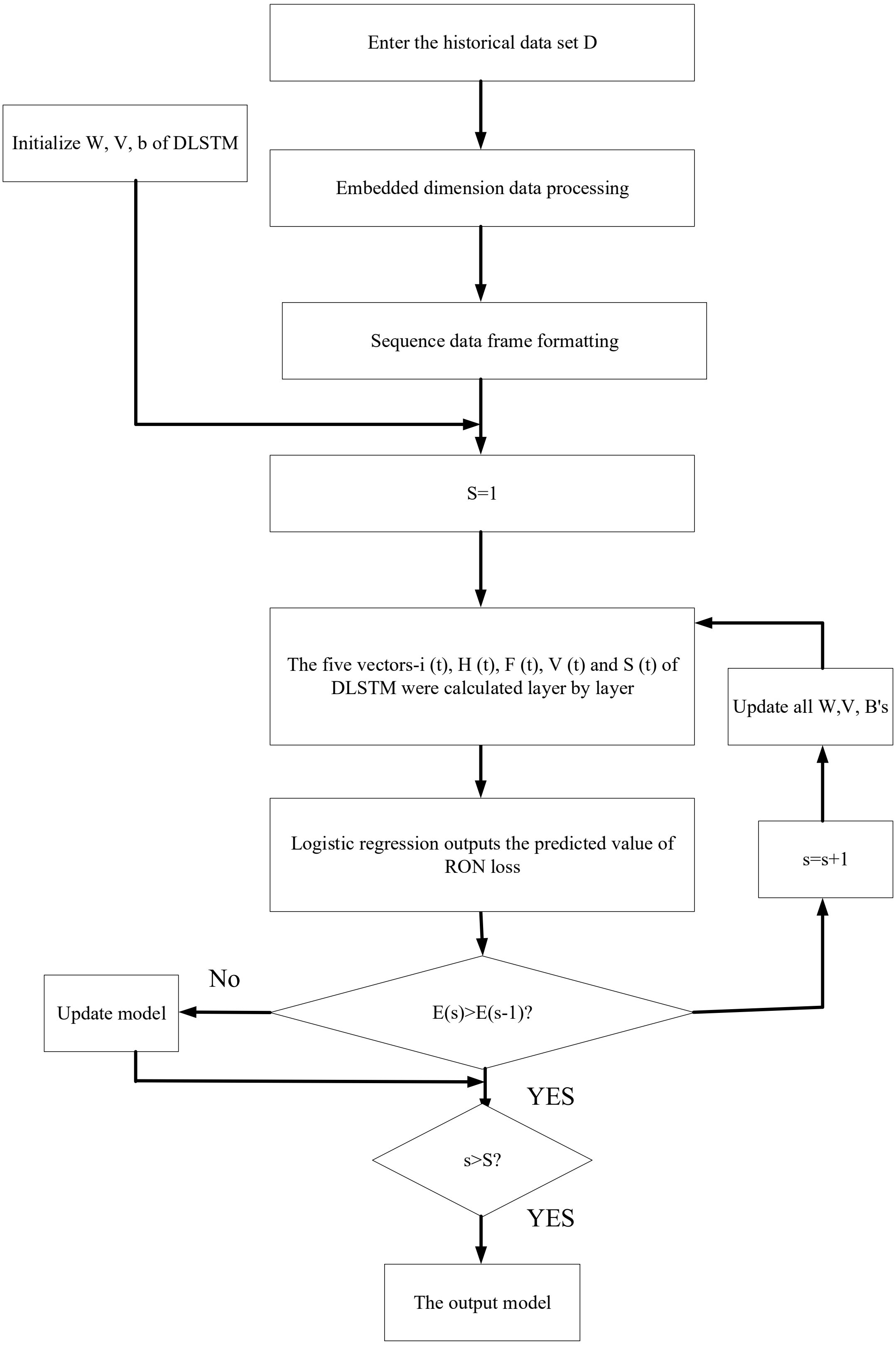

Figure 3 DLSTM prediction flow chart

3.3 DlSTM-based model training

Due to the use of deep learning, the training cycle of the model is longer, and more data need to be used for training until convergence can be used for population prediction.The data set division is explained in detail here: The training set needs to be repeatedly trained on DLSTM model until the model converges, which requires a large amount of data, so the first 80% of the data sorted by the timestamp is used for training.In 80% of the data, 70% of them are used as DLSTM learning sets, 15%

as cross-checking data sets, and 15% as verification sets to verify the accuracy of the network.The test set is used as extrapolation data to test the prediction ability and robustness of the model for real situations.Select the last 20% of data here as the test set.Matlab neural network toolbox was used to train the established DLSTM, and the number of iterations was set to 1000.The training state diagram of DLSTM neural network is shown in Figure 4.

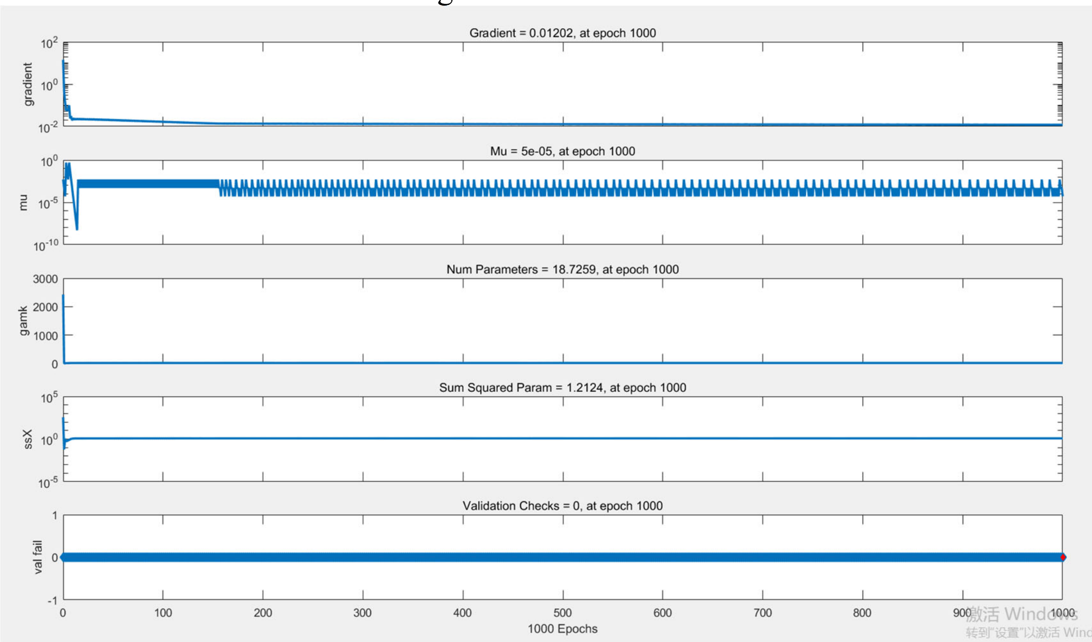

Figure 4 Training state diagram of DLSTM neural network

3.4 Artificial neural network (ANN) prediction model comparison

ANN is one of the most common neural networks, which has some commonalities of other neural networks.The purpose of using ANN modeling and prediction here is to compare with the modeling results of DLSTM neural network and detect the improved performance of our model compared with ordinary neural network.

The ANN structure is composed of input layer, hidden layer and output layer, in which the input layer has 25 features. The ANN constructed by our team has 50 hidden layers, and one output variable is the population to be predicted.

Between each connected neuron there is a specific weight applied to the signal transmitted by the neuron, and these weight values are transmitted backward after each layer acting on the neuron, and finally linearly coupled in the output layer.When input signals form output signals through the dip network model, the relationship between them is shown in the following figure[11]:

$$y_i = g(v) = g(\sum_{j=1}^{n} w_{ij} x_j + \theta_i) \tag{6}$$

Where, Yi is the output result of ANN, Xi is the input result, and WIJ is the weight imposed by the layer I neural network on the JTH input signal.G is a nonlinear activation function and plays a role of nonlinear mapping.There is also a threshold θ I for paranoid treatment of variable xi.

Using ANN also requires training data to make the model converge, and the ANN convergence error we constructed is set to 10^{-12}It can achieve good experimental results.Moreover, we find that when the training cycle is set to 4000 times, the convergence speed of ANN training will be significantly slower and no longer has training value.As ANN is used as a comparative neural network and limited by space, the building code and training process of ANN will not be shown.

4. Model solving and index testing

According to the need of statistical error analysis and actual operation prediction, This paper mainly uses mean Absolute error (MAE), root mean Square error (RMSE) and mean absolute percentage error (MEAN absolute Percentage error) Error,MAPE) are used as evaluation indicators of the prediction results of the subsequent model, and their symbols are respectively, and. E_{RMS} E_{MA} E_{MAP} Experimental indicators are defined as follows:

$$E_{RMS}=\sqrt{\frac{1}{N}\sum_{p=1}^{N}(\widehat{y}_p-y_p)^2} \tag{7}$$

$$E_{MA}=\frac{1}{N}\sum_{p=1}^{N}\left|\widehat{y}_p-y_p\right| \tag{8}$$

$$E_{MAP}=\frac{1}{N}\sum_{p=1}^{N}\left|\frac{\widehat{y}_p-y_p}{y_p}\right| \tag{9}$$

Where, is the number of years to be predicted; N $\widehat{y}_p$ Is the predicted population; y_p Is the actual population.According to the above definition, the smaller the value of, and, the more consistent the predicted value is with the real value, and the more perfect the model is. E_{RMS} E_{MA} E_{MAP} On the contrary, the greater the difference between the predicted value and the actual value, the worse the prediction effect.According to the above test set Settings, we tested the last 60 data, and some predicted results are shown in Figure 2.

Table 2 DLSTM model predicts the population in the next 10 years

2021	2022	2023	2024	2025	2026
141628.7253	142286.6528	142950.7281	143618.8134	144291.9411	144973.3973
2027	2028	2029	2030	2031	2032
145657.5457	146343.8174	147031.6612	147718.2872	148402.0122	149081.8829

The comparison between DLSTM neural network and ANN neural network is shown in Table 3.DLSTM performed better than ANN in all indicators, especially the average accuracy. The average accuracy of DLSTM neural network with multiple training iterations increased by 5 percentage points to 92.9%, while the average accuracy of ANN was only 87.75%.

Table 3 Comparison of prediction indexes between DLSTM neural network and ANN neural network

Prediction model	Ema/g	Emap/%	Erms/g	Average accuracy /%
DLSTM neural network	0.089141077	0.743965439	0.1054698	0.928954270
ANN neural network	0.153708543	0.128702738	0.1956244	0.877500526

5. Conclusion

The number and structure of population are important factors affecting social and economic development.This paper mainly establishes a population prediction model, considering that the predicted value is a time series spanning 10 years, which has a certain time correlation. However, the predicted characteristics are only 25 dimensions according to the relevant data of the country we have investigated, so it is not suitable to use the time series method for modeling.So we introduced both short-term and long-term memory neural networks with time characteristics (LSTM), at the same time, considering the population forecast may be is more sensitive to high-dimensional feature, use stacked both short-term and long-term memory neural network (DLSTM, significant short-term memory neural network) as a predictive model, multilayer superposition prediction, significantly improve the robustness and accuracy, The results of DLSTM model are compared with the traditional artificial neural network prediction model to prove the feasibility of DLSTM model, and the analysis of related parameters is completed.

References

[1] Alsharef Ahmad, Aggarwal Karan,Sonia Deepika,Koundal Hashem,Alyami Darine,Ameyed Carmen. An Automated Toxicity Classification on Social Media Using LSTM and Word Embedding[J]. Computational Intelligence and Neuroscience,2022,2022.

[2] Nadda Wanchaloem,Boonchieng Waraporn,Boonchieng Ekkarat. Influenza, dengue and common cold detection using LSTM with fully connected neural network and keywords selection[J]. BioData Mining,2022,15(1).

[3] Vukovic Darko B.,Romanyuk Kirill,Ivashchenko Sergey,Grigorieva Elena M.. Are CDS spreads predictable during the Covid-19 pandemic? Forecasting based on SVM, GMDH, LSTM and Markov switching autoregression[J]. Expert Systems With Applications,2022,194.

[4] Chi Dianwei. Research on electricity consumption forecasting model based on wavelet transform and multi-layer LSTM model[J]. Energy Reports,2022,8(S4).

[5] Jiang Jingfei,Xiao Tao,Xu Jinwei,Wen Dong,Gao Lei,Dou Yong. A low-latency LSTM accelerator using balanced sparsity based on FPGA[J]. Microprocessors and Microsystems,2022,89.

[6] Tsan YuTse,Chen DerYuan,Liu PoYu,Kristiani Endah,Nguyen Kieu Lan Phuong,Yang ChaoTung. The Prediction of Influenza-like Illness and Respiratory Disease Using LSTM and ARIMA[J]. International Journal of Environmental Research and Public Health,2022,19(3).

[7] Singh, Kuldeep,Malhotra, Jyoteesh. Two-layer LSTM network-based prediction of epileptic seizures using EEG spectral features[J]. Complex & Intelligent Systems,2022(prepublish).

[8] Richard Jacqueline A.,Sa'don Norazzlina M.,Karim Abdul Razak Abdul. Artificial Neural Network (ANN) Model for Shear Strength of Soil Prediction[J]. Defect and Diffusion Forum,2021,6277.

[9] Kamaraj Logesh,Hariharasakthisudhan P,Arul Marcel Moshi A. Optimizing the ultrasonication effect in stir-casting process of aluminum hybrid composite using desirability function approach and artificial neural network[J]. Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications,2021,235(9).

[10] Millán Ocampo, D. E.,Parrales Bahena, A.,de Lourdes Llovera Hernández, Ma.,Silva Martínez, S.,Porcayo Calderón, J.,Hernández, J. A.. Modeling of electrochemical removal of cadmium under galvanostatic mode using an artificial neural network[J]. International Journal of Environmental Science and Technology,2021(prepublish).

[11] Aklilu, Ermias Girma,Waday, Yasin Ahmed. Optimizing the process parameters to maximize biogas yield from anaerobic co-digestion of alkali-treated corn stover and poultry manure using artificial neural network and response surface methodology[J]. Biomass Conversion and Biorefinery,2021(prepublish).