International Journal of Engineering Research And Management (IJERM)
ISSN: 2349- 2058, Volume-11, Issue-05, May 2024

Research on Neural Network Optimization Method Based on Genetic Algorithm

Yaxin Lu, Han Wu

***Abstract*—The design of neural network models relies on the expertise and experience of experts, which makes it difficult to obtain excellent neural network models, and manual adjustment of model parameters is time-consuming and complex. LSTM neural network shows good performance in time series prediction. The parameters and structure of the LSTM neural network will affect the fitting ability of the model, and have a great influence on the prediction performance of the model. As a global optimization algorithm, genetic algorithm has the advantages of strong searching ability and good robustness, which is suitable for the optimization of neural networks. Therefore, this paper will study the optimization method of LSTM neural network based on genetic algorithm. This paper first introduces the basic principle of genetic algorithm and LSTM neural network, then expounds the flow and method of genetic algorithm optimization of LSTM neural network model, and verifies the optimization effect through experiments.**

***Index Terms*—genetic algorithm, neural networks, optimization, LSTM**

I. INTRODUCTION

With the advent of the era of big data, machine learning is more and more widely used in various fields. As an important machine learning method, neural network has made remarkable achievements in image recognition, speech recognition, natural language processing and other fields. However, parameters greatly affect the performance of neural networks [1]. The current practice of designing neural networks relies heavily on subjective judgment, often determined by the level of expertise possessed by the architectural designer. Therefore, how to optimize neural network parameters automatically and efficiently has become an urgent problem to be solved [2]. As a global optimization algorithm, genetic algorithm has strong searching ability and robustness. When genetic algorithm is applied to the optimization of neural network, the optimal network parameters can be found automatically and the performance of neural network can be improved.

Genetic algorithm (GA) is a global search algorithm based on natural principles proposed by John Holland of University of Michigan. Based on the basic ideas of biological evolution, basic concepts such as population, selection, mutation, crossover and fitness are introduced [3]. The basic elements of genetic algorithm are chromosome representation, fitness selection and biological heuristic operators [4]. It simulates selection, crossover and mutation in biological evolution to find the optimal solution. In genetic algorithms, each possible solution to a problem is encoded into a chromosome or individual. Multiple individuals form a population, the individuals in the population are evaluated by fitness function, and the excellent individuals have a greater chance to be selected to reproduce the next generation. Through continuous iteration, the individuals in the population gradually evolve and eventually find the optimal solution to the problem.

Long short-term memory (LSTM) was invented by Hochreiter and Schimdhuber in 1997, it has gained popularity as an Recurrent Neural networks (RNN) architecture in various applications in recent years [5]. RNN have recurrent links in their Network structure, and the relationships between samples can be considered in the learning process, so they are particularly suitable for processing time series signals. But if there is a long-term dependence between samples, RNN will suffer from gradient disappearance and gradient explosion. The LSTM network is an improved method to solve this problem [6]. In recent years, LSTM has been applied more and more in soil moisture prediction, air pollution prediction and power load prediction [7,8,9].

The network structure and the setting of hyperparameters of LSTM model have important influence on the performance of LSTM. Some recent studies have applied genetic algorithms to optimize LSTM structures and hyperparameters for time series prediction tasks. Santra et al. [10] proposed a genetic algorithm to optimize the weight of LSTM and applied the proposed method to power load forecasting. The results show that this method produces a small average absolute percentage error on the test data. Paweena et al [11]. used LSTM model and Bi-LSTM model to predict soil moisture, and the results showed that The RMSE errors of LSTM model and Bi-LSTM model were 0.72% and 0.76%, respectively.

In this paper, the method of optimizing LSTM by genetic algorithm is introduced in detail, and its effectiveness is verified by experiments.

II. GENETIC ALGORITHM TO OPTIMIZE NEURAL NETWORK

A. Genetic algorithm

Genetic algorithm is an optimization algorithm based on natural selection and genetic mechanism. It simulates selection, crossover and mutation in the process of biological evolution to search for the optimal solution. In genetic algorithms, each possible solution to a problem is encoded into a chromosome, or individual. Multiple individuals form a population, and the individuals in the population are evaluated by the fitness function, and the excellent individuals

Manuscript received May 09, 2024
Yaxin Lu, School of Software, Tiangong University, Tianjin, China
Han Wu, School of Software, Tiangong University, Tianjin, China

have a greater chance of being selected for breeding the next generation. Through continuous iteration, the individuals in the population gradually evolve and eventually find the optimal solution to the problem. The flow chart of the genetic algorithm model is shown in Figure 1.

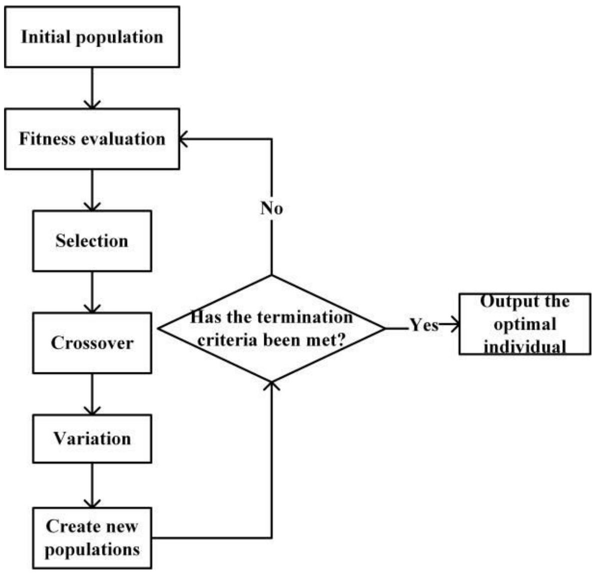

Figure 1 Genetic algorithm flow chart

B. LSTM

The core of the LSTM model is the Memory Cell, which is responsible for storing and transmitting information. The state of the memory cell is updated by gating mechanisms, including Input gates, Forget gates, and Output gates.

Input gates are used to control the input of new information. When the input gate is activated, new information is added to the memory cell. This process consists of the sigmoid function, which converts the input to a value between 0 and 1, and a dot product operation that determines what information can pass through.

Forgetting gates are used to control the retention or forgetting of information in memory cells. The forgetting door decides what information to discard from the memory cell based on the hidden state of the previous moment and the input of the current moment. The forgetting gate also uses the sigmoid function to make decisions, and realizes the forgetting of information by multiplying the state of the memory cell by the dot multiplication operation.

The output gate is used to control the output of information. The output gate determines which information to output based on the current moment of input and the state of the memory cell. The output gate also uses the sigmoid function to make decisions, but its output is dotted with the tanh value of the memory cell to get the final output.

In this way, LSTM models are able to selectively retain and update information to better handle time series data with long-term dependencies. This makes LSTM widely used in speech recognition, natural language processing, time series prediction and other fields. The LSTM model structure diagram is shown in Figure 2.

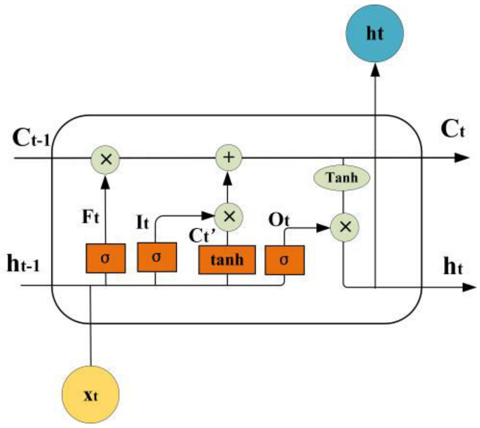

Figure 2 LSTM structure

III. GENETIC ALGORITHM TO OPTIMIZE LSTM MODEL CONSTRUCTION

The main task of LSTM optimization method based on genetic algorithm is to use genetic algorithm to search LSTM model automatically. The 1/ MSE loss of the LSTM network model is used as the fitness function of the genetic algorithm, so that the genetic algorithm and the LSTM model are combined, and the LSTM optimization method based on genetic algorithm (GA-LSTM) is constructed.

The GA-LSTM model is divided into two parts, as shown in Figure 3. The right side is the GA algorithm flow, and the left part is the LSTM network training and prediction process.

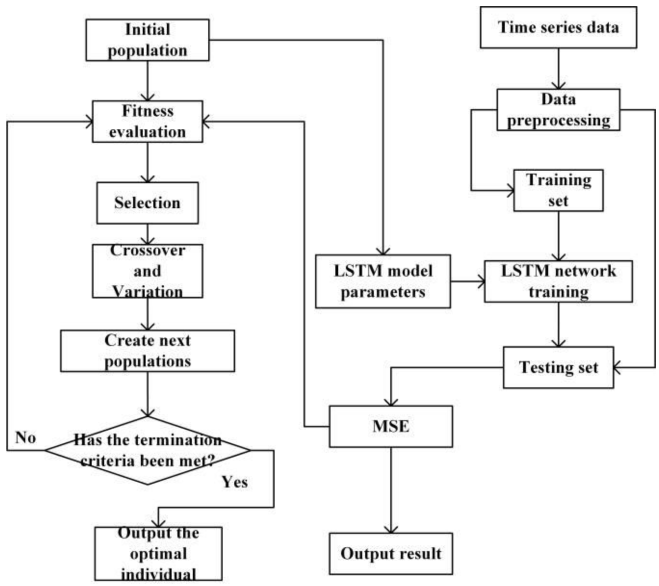

Figure 3 GA-LSTM flow chart

First, the time series data is preprocessed and the training set and test set are divided, initial the population, set evolutionary algebra, initial population size, crossover probability, and mutation probability. Secondly, the LSTM network is constructed using the initial population to obtain the MSE on the test set. The MSE obtained from the training set is returned to the GA and the fitness of the LSTM network is evaluated; Finally, the genetic operation (crossover, variation) is carried out until the termination condition is reached, and finally the predicted value of the most individual is output.

IV. EXPERIMENTS AND RESULTS ANALYSIS

A. Data set

In this paper, the baseline data set of PM2.5 concentration pollution prediction in the machine learning knowledge base of University of California Irvine (UCI) is selected. The data set contains the PM2.5 concentration value and the corresponding date and meteorological information. The data set is divided into training set, verification set and test set in a ratio of 6:2:2.

B. Evaluation indicators

In order to evaluate the prediction performance of LSTM model, mean square error (MSE), mean absolute error (MAE) and R-Square (R^2) are used as evaluation indexes. The formula is shown in equation (1),(2),(3).

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{1}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right| \tag{2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \overline{y})^2} \tag{3}$$

Where N is the number of samples, y_i and $\hat{y}_i$ are the ground truth and the predicted value of the i sample respectively, $\overline{y}$ is the average value.

MSE and MAE are used to measure the deviation between the predicted value and its corresponding true value, and the lower the value, the better. R^2 indicates the degree of model fit, the higher the better.

C. Experimental parameter setting

The population size is 20, the evolutionary algebra is 20 generations, the crossover probability is 0.7, and the mutation probability is 0.06. LSTM model optimization parameters include learning rate and epoch, where the value of learning rate [0.0001,0.001] and epoch range [10,100].

D. Experimental result

Table 1 shows the experimental results of the GA-LSTM method and the unoptimized LSTM model on the training set.

Table 1 Training set experiment results

	MSE	MAE	R^2
GA-LSTM	555.11	13.01	0.93
LSTM	575.14	13.20	0.93

Table 2 shows the experimental results of the GA-LSTM method and the unoptimized LSTM model on the testing set.

Table 1 Training set experiment results

	MSE	MAE	R^2
GA-LSTM	426.40	11.46	0.95
LSTM	449.93	11.77	0.94

The prediction diagram of GA-LSTM model is shown in Figure 4.

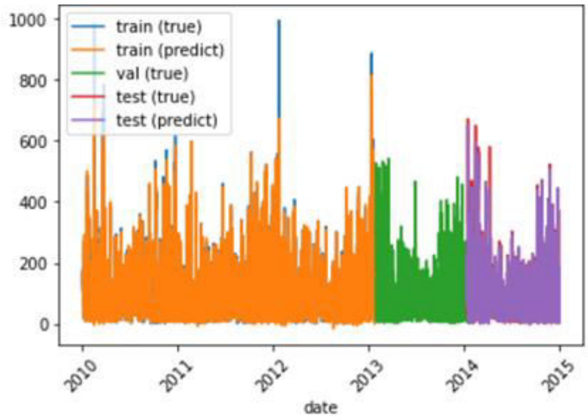

Figure 4 GA-LSTM model prediction results

As can be seen from Table 1 and Table 2, the prediction results of GA-LSTM model are better than those of LSTM model in the three indexes. For example, the MSE of the LSTM model on the test set is 426.40, while the MSE of the LSTM model is 449.93.

V. CONCLUSION

Genetic algorithm optimization neural network is an effective technique, which can automatically find the optimal parameters of neural network and improve the performance of neural network. Through experimental verification, we prove the superiority of this method in dealing with complex problems. However, genetic algorithms still face some challenges, such as the selection of coding mode, the definition of fitness function and the convergence speed of the algorithm. In the future, we will continue to study these problems in depth, explore more efficient optimization methods, and make greater contributions to the development of artificial intelligence.

REFERENCES

[1] Fatyanosa, Tirana Noor, and Masayoshi Aritsugi. "An automatic convolutional neural network optimization using a diversity-guided genetic algorithm." IEEE Access 9 (2021): 91410-91426.

[2] Pijackova, Kristyna, et al. "Genetic algorithm designed for optimization of neural network architectures for intracranial EEG recordings analysis." Journal of neural engineering 20.3 (2023): 036034.

[3] H. VIKHAR P A. Evolutionary algorithms: A critical review and its future prospects.2016 International conference on global trends in signal processing, information computing and communication (ICGTSPICC). IEEE, 2016: 261-265.

[4] Katoch, Sourabh, Sumit Singh Chauhan, and Vijay Kumar. "A review on genetic algorithm: past, present, and future." Multimedia tools and applications 80 (2021): 8091-8126.

[5] Pascanu, R.; Mikolov, T.; Bengio, Y. On the difficulty of training Recurrent Neural Networks. In Proceedings of the 30th International Conference on Machine Learning, Atlanta, GA, USA, 16–21 June 2013; pp. 1310–1318.

[6] Lei P, Ma F, Zhu C, Li T. LSTM Short-Term Wind Power Prediction Method Based on Data Preprocessing and Variational Modal Decomposition for Soft Sensors. Sensors. 2024; 24(8):2521.

[7] Suebsombut, Paweena, et al. "Field Data Forecasting Using LSTM and Bi-LSTM Approaches." Applied Sciences (2021).

[8] Zhang, Jiaxuan, and S. Li . "Air quality index forecast in Beijing based on CNN-LSTM multi-model." Chemosphere 308.Pt 1(2022):136180..

[9] Tang, Jiajie, et al. "Bus Load Forecasting Method of Power System Based on VMD and Bi-LSTM." Sustainability 13(2021).

[10] Santra AS, Lin J-L. Integrating Long Short-Term Memory and Genetic Algorithm for Short-Term Load Forecasting. Energies. 2019; 12(11):2040.

[11] Suebsombut P, Sekhari A, Sureephong P, Belhi A, Bouras A. Field Data Forecasting Using LSTM and Bi-LSTM Approaches. Applied Sciences. 2021; 11(24):11820.