A Review of the Research on the Prediction Model of Extreme Learning Machine

Li Yibo\textsuperscript{1,a} Liu Fang\textsuperscript{2,b} Cao Qi\textsuperscript{3,c}

Professor
\textsuperscript{1}Shenyang Aerospace University
13998825712, Shenyang, China

student
\textsuperscript{2}Shenyang Aerospace University
13940073081, Shenyang, China

student
\textsuperscript{3}Shenyang Aerospace University
18740095117, Shenyang, China

e-mail: liyibo_sau@163.com \hspace{1em} e-mail: 982373802@qq.com
\hspace{1em} e-mail: 851434033@qq.com

ABSTRACT: With the wide application of prediction, the research of prediction algorithm and theory has made a great progress. In recent years, extreme learning machines have been used in the field of prediction, such as stock price prediction. The prediction algorithms of ELM are reviewed in this paper, which includes the single extreme learning machine prediction algorithm and the combined prediction algorithm. At the same time, the existing problems and research directions are pointed out.

1. INTRODUCTION
The continuous development of artificial intelligence provides more strategies for scientific prediction, especially neural networks which has been used in prediction by more and more scholars as for the characteristics of adaptive learning and good robustness. The feedforward neural network, support vector machine and other methods advanced later have achieved a wider application on the field of prediction. But the traditional feedforward neural network and support vector machine have some shortcomings to some extent. On this basis, Huang and other scholars proposed the algorithm of extreme learning machine, which avoids the disadvantage that traditional neural network easily fall into the local optimal. Extreme learning machine developed from the neural network with single hidden layer, which has some remarkable characteristics, such as the easily realized algorithm, the relatively faster execution speed, the more stronger generalization performance, a better robustness, etc. Prediction is a kind of behavior that describes the future development trend according to the regularity of the development in the past certain period of time, and it was produced and developed according to the social development and social management. In recent years, due to the unprecedented rapid development of science and technology of the global economy, uncertainties in social life have increased significantly, such as the financial crisis, economic decision-making, energy demand, etc.
These uncertainties increase the urgency and necessity for people to understand and grasp the future development, and promote the further development of prediction. Therefore, it has important research value to improve the extreme learning machine and apply it to the prediction field.

The application of the improved extreme learning machine in prediction is not only of great practical significance, but also of great economic benefit and broad application prospect. In this paper, the prediction algorithm of ELM is divided into single ELM prediction algorithm and combined prediction algorithm according to the structure of prediction model. At the same time, the article reviewed the research results and present situation of the prediction algorithm of ELM, and pointed out the research direction.

2. Single Extreme Learning Machine Prediction Algorithms

2.1 Prediction of Extreme Learning Machine

Extreme Learning Machine (ELM), as a new single hidden layer feedforward neural network (SLFN), is a fast learning algorithm proposed by Professor Huang Guangbin of Nanyang University of Technology in Singapore in 2004[1]. The characteristic of it can be stated that in the process of determining network parameters, the input weights and the bias of hidden layer can be selected randomly without any adjustment, which improves the training speed and reduces the adjustment time of parameters.

For \( N \) different samples \((x_i, t_i)\), with regard to: \( x_i = [x_{i,1}, x_{i,2}, \ldots, x_{i,n}]^T \in \mathbb{R}^n \), \( \hat{b}_i \), the number of hidden layer nodes is \( N \), the model of single hidden layer feed forward neural network of excitation function \( g(x) \) is:

\[
\sum_{i=1}^{N} \beta_i g(x_i) = \sum_{i=1}^{N} \beta_i g(\omega_i x_i + b_i) = o_j, j = 1, \ldots, N
\]

In the formula, \( \omega_i \) is the weight of connecting the first hidden layer node and the input node, \( \beta_i \) is the weight of connecting the second hidden layer node and the output node, \( b_i \) is the threshold of the first hidden layer node.

In reference [2], the extreme learning machine is applied for stock price prediction. Firstly, the relevant data affecting stock price are collected and processed, and the learning model is established by selecting the main factors. Then the prediction model is obtained by using the extreme learning machine on the basis of data set. The simulation results are shown in Figure 1. The experimental results show that the extreme learning machine has obvious advantages in parameter selection and learning speed compared with support vector machine and BP neural network.

| Predictive model | Correlation coefficient | RRMSE |
|------------------|-------------------------|-------|
| ELM              | 0.993 2                 | 1.31  |
| SVM              | 0.994 1                 | 1.28  |
| BP               | 0.969 1                 | 8.77  |
Figure 1 Comparison of stock price true value and ELM prediction result

Reference [3] proposes a method of estimating the cost of power line construction based on extreme learning machine, establishes a neural network model for estimating the factors affecting the cost of power line construction and the cost of power construction, trains the network model with extreme learning machine, and determines the corresponding parameters of the network. The experiment shows that ELM is not only fast in estimating the cost of power construction, but also has good results. Generalization performance is also better. Reference [4] proposes an improved Regularized ELM (RELM) for off-line modeling and prediction of network traffic, which improves the generalization ability. Reference [5] proposes another improved online sequential extreme learning machine (OS-ELM) to model and predict online network traffic. Reference [6] proposes a hybrid prediction model of coal mine gas concentration, W-ELM based on wavelet transform, which combines wavelet transform with extreme learning machine. The ELM model was established to predict the components of the gas concentration time series, and then the prediction values were superimposed to obtain the prediction values of the original series. Experiments show that the model has good application prospects and can be predicted one or more steps in advance.

2.2 Prediction of Kernel Extreme Learning Machine

Huang Guangbinl. proposed a kernel function based on ELM in 2010 [7]. KELM algorithm has more relaxed constraints, fewer parameters, more concise computational complexity and better generalization. The performance of KELM is mainly determined by the selected kernel functions and parameters [8-9]. The output and output weights of the nuclear extreme learning machine are:

\[
f(x) = \left[ k(x, x_1) \quad \cdots \quad k(x, x_N) \right]^T \left( \frac{I}{C} + \Omega_{ELM} \right)^{-1} T
\]

\[
\beta = \left( \frac{I}{C} + \Omega_{ELM} \right)^{-1} T
\]

In the equation, \( I \) is a diagonal matrix, \( C \) is a penalty coefficient, \( T \) is an output target vector, \( \Omega_{ELM} \) is a kernel matrix, and \( k(x, x_i) \) is a kernel function.

In reference [10], a new combined KELM is proposed by using the weighted polynomial kernel of Gauss kernel. The parameters of the combined KELM are optimized by chaotic particle swarm optimization. Finally, the model is applied to network traffic prediction area. The simulation results are shown in Table 2. By comparing the mean square error of the prediction, it is found that the prediction accuracy of the algorithm is better than other models, which provides a basis of reasonable allocation and decision-making of reliable transmission network resources.
Table 2 Comparison of MSE predicted by four algorithms

| Number of runs | SVM       | CPSO-KELM | CPSO-MKELM | EMD-CPSO-MKELM |
|----------------|-----------|-----------|------------|----------------|
| 1              | 0.0088    | 0.0060    | 0.0056     | 0.0040         |
| 2              | 0.0072    | 0.0055    | 0.0051     | 0.0046         |
| 3              | 0.0068    | 0.0050    | 0.0049     | 0.0040         |
| 4              | 0.0084    | 0.0064    | 0.0051     | 0.0042         |
| 5              | 0.0065    | 0.0056    | 0.0047     | 0.0038         |
| 6              | 0.0074    | 0.0061    | 0.0058     | 0.0043         |
| 7              | 0.0076    | 0.0056    | 0.0046     | 0.0035         |
| 8              | 0.0064    | 0.0046    | 0.0043     | 0.0042         |
| 9              | 0.0072    | 0.0076    | 0.0054     | 0.0040         |
| 10             | 0.0067    | 0.0051    | 0.0048     | 0.0047         |
| Mean           | 0.0072    | 0.0058    | 0.0052     | 0.0041         |

Reference [11] presents a prediction model of transformer top oil temperature interval based on KELM and Bootstrap method. Firstly, through Bootstrap sampling, L groups of training samples are obtained, and L nuclear extreme learning machine models are trained to fit regression prediction of top oil temperature, then a nuclear extreme learning machine model is trained to estimate the variance of top oil temperature observation noise. Finally, the prediction interval of top oil temperature at a certain confidence level is estimated based on the results of a nuclear limit learning machine model. It is clear that the computing speed of kernel extreme learning machine is significantly higher than that of support vector machine.

In order to improve the accuracy of short-term traffic flow prediction, reference [12] proposed a hybrid model based on singular spectrum analysis (ssa) and KELM is proposed. The KELM model is trained with SSA filtered traffic flow data, and the optimized KELM model is applied to traffic flow prediction of an Expressway in Xiamen. The results of experiment shows that the accuracy and robustness of the model are better than those of several commonly used prediction models such as support vector machine, extreme learning machine and single KELM model.

Reference [13] proposes a particle swarm optimization based kernel extreme learning machine (PSO-KELM) model for wind power interval prediction. Fast interval prediction of wind power is realized. Experiments show that PSO-KELM model has higher prediction accuracy, faster speed and stronger generalization ability, and can realize fast interval prediction of wind power.

2.3 Prediction of Online Sequence Extreme Learning Machine

The training set of ELM algorithm is fixed in the training process. However, in practical application, the training set cannot be given totally at first while new samples will be generated during the process. When more and more samples exist, it will complicate the calculation of H generalized inverse matrix, and the calculation time will be greatly increased. In view of this situation, researchers proposed an online Sequence ELM algorithm, based on the improved algorithm ELM, The idea of OS-ELM algorithm is to solidify the sample data in the output matrix of hidden layer nodes. After the new sample is added in the training set, the network only needs to update the current generated new samples, and does not need to update the sample set before repeated updates, so the calculation time and workload are greatly reduced, and the real-time requirement of travel time is effectively solved. Assuming that k samples enters the model, the output and output weights of the online sequence extreme learning machine algorithm are as follows:

\[
\begin{align*}
P_{k+1} &= P_k - P_k H_{k+1}^T (I + H_{k+1}P_k H_{k+1}^T)^{-1} H_{k+1}P_k \\
\beta^{k+1} &= \beta^k + K_{k+1}^H H_{k+1} (I - \beta^k) 
\end{align*}
\]
Reference [14] proposed a prediction model of online sequence-extreme learning machine based on online renewal theory. When the model was updated, the prediction factor data were gradually imported into the model for training according to different batches. In addition, the standard precipitation index was selected as the drought evaluation index to predict the study area. The simulation results are shown in Figure 2. The results show that the prediction accuracy of online sequence-extreme learning machine is greatly improved compared with that of extreme learning machine.

![Figure 2 Drought prediction results of online sequence extreme learning machine](image)

Reference [15] proposes a travel time prediction algorithm based on online sequence extreme learning machine. In order to improve the prediction accuracy of link travel time, an incremental learning framework method is introduced. In the online sequence learning stage, the algorithm updates the data of the input samples and gets the parameters of the algorithm until all the data in the input set are input, and then predicts the test set. The experimental results show that the prediction algorithm has better adaptability, real-time and accuracy.

Reference [16] presents a time series prediction method based on the combination of online sequence extreme learning machine and adaptive forgetting factor and bootstrap. Experiments show that B-AFF-OS-ELM improves the prediction accuracy and stability.

Reference [17], an online sequence extreme learning machine based on M estimator is proposed to predict chaotic time series with outliers. The traditional least squares cost function is replaced by robust M estimation cost function, which enhances the robustness of the model to outliers. By minimizing the cost function based on M estimator, the possible outliers are prevented from entering the output weight update scheme of the model. Experiments show that the M-OSELM algorithm has good anti-outlier performance and guaranteeing all the advantages of online sequence method.

### 2.4 Prediction of Regular Extreme Learning Machine

Compared with the traditional BP neural network, the extreme learning machine has the characteristics of simple structure, fast learning speed and strong generalization ability. However, the least squares loss function is usually used to construct ELM model only considering the experience risk without the structural risk, which easily leads to over-fitting situation. In order to solve this problem, a regular extreme learning machine is constructed by introducing regular coefficients to improve the generalization ability of the traditional extreme learning machine.

The regular extreme learning machine regression model is:

\[
\hat{y} = \sum_{i=1}^{L} \beta_i g(\alpha_i x + b_i)
\]

In the formula, \(L\) represents the number of input hidden layer nodes; \(\alpha_i\) indicates the weight of the input layer node, \(b_i\) represents the threshold of the hidden layer, \(\beta_i\) represents the weight of the hidden layer, \(g(x)\) is the excitation function. Since the above formula consists of many equations, it can be converted into \(H \beta = T\), \(H\) is the hidden layer matrix. For a nonlinear problem, the regular extreme learning machine can transform it into the following form:

\[
\arg\min \left( \frac{1}{2} \| \beta \|^2 + \gamma \frac{1}{2} \| F \|^2 \right)
\]
In the formula, $\|\beta\|$ is the structural risk and $\gamma$ is the adjustment parameter.

Reference [18] proposes a network traffic prediction model based on phase space reconstruction and regular extreme learning machine. Firstly, the network traffic is pretreated, then the optimal delay time and embedding dimension are determined according to chaos theory, and the network traffic learning samples are reconstructed. Finally, the network traffic model is established by using regular extreme learning machine. The simulation diagram is shown in Figure 3. The experimental results show that the proposed model can be applied to the network traffic prediction. The improved model predicts the time series of network traffic with different time scales, and the predicted value is in good agreement with the actual value, which can accurately reflect the non-linear change law of network traffic.

![Figure 3 Curved predicted value of network traffic with a sampling time of 1 hour and the actual value](image)

A new short-term wind speed prediction method, which is based on weighted regular extreme learning machine (WRELM), is proposed in reference [19]. Firstly, the maximum correlation minimum redundancy (MCMR) criterion is used to select the feature set which is the most relevant factor to the wind speed series as the prediction input, and then the training set and test set of the prediction network are determined to construct the WRELM prediction model. Finally, the WRELM model is used to predict the short-term wind speed. Experiments show that the method is effective and can be used in short-term wind speed forecasting practice.

In reference [20], a sequential regularized extreme learning machine (SRELM) and a prediction method of feature parameters based on SRELM are proposed. Firstly, the SRELM model is trained by the characteristic parameters of the hydraulic pump, and the latest characteristic parameters are used to update the prediction model iteratively. Then, the trained prediction model is used to predict the future characteristic parameters. Experiments show that this method has better performance in prediction accuracy, computation and generalization performance.

Reference [21] predicts the average value of PM2.5 using the atmospheric data of NO2 and PM 10 in Changsha in 2017, and uses BIC criterion to select features. On the basis of Extreme Learning Machine (ELM), the regularization term is introduced to control the complexity of the model, and the prediction model of PM2.5 is established. Experiments show that the model has better prediction performance, and the mean square error, mean absolute error and mean absolute error are reduced.

3. Combination Prediction

Combined forecasting is to make full use of the model advantages provided by various single forecasting methods and adopt appropriate combination methods for fusion, thereby expanding the application range of prediction methods and improving the accuracy of prediction results.

Gray extreme learning machine prediction model is mainly composed of gray prediction model and extreme learning machine prediction model. It is suitable for small sample data. It has good prediction accuracy for fitting historical data of exponential law but low prediction accuracy for medium and long term. Extreme Learning Machine (ELM) has strong ability to process non-linear information and high accuracy for medium and long term prediction of prediction objects, but it needs a large number of training samples. Moreover, because its input weights are randomly generated, the
same historical data may produce different prediction results with different prediction accuracy. The flow chart of the algorithm is as follow:

In reference [22], a method combining gray model with extreme learning machine is proposed to predict rolling bearings, and the data of future bearing operation state are analyzed to achieve the purpose of fault prediction. The simulation results are shown in Figure 4. The experimental results show that the gray ELM prediction model has better prediction accuracy than the gray prediction model, and it is closer to the fitting curve of the actual value.

Reference [23] combines gray model and extreme learning machine algorithm to get a new combination forecasting model, and applies the combination forecasting model to price forecasting of supply chain management. This algorithm combines the advantages of high prediction accuracy, stable prediction results of gray prediction on small sample data with high prediction accuracy of limit learning machine on non-linear data. It is a good prediction algorithm with good prediction accuracy and prediction speed.

4. Conclusion
At present, the prediction algorithm of extreme learning machine is an important issue in the field of artificial intelligence. Through the research of many experts, the research of extreme learning machine prediction has achieved much results. This paper summarizes some research results and current situation of extreme learning machine prediction in recent years. The above analysis shows that each extreme learning machine prediction algorithm has its own advantages and disadvantages. Combination forecasting method is better than single forecasting method on the forecasting accuracy. In the future, the study of extreme learning machine forecasting algorithm can use appropriate
combination method to fuse different single forecasting methods after improving the single extreme learning machine forecasting model, so as to expand the application scope of forecasting methods and improve the accuracy of forecasting results.

REFERENCES

[1] Guangbin Huang, Qinyu Zhu, SIEW Cheekheong. Externling learning machine: theory and applications[J]. Neuro computing, 2006 (70): 489-501.
[2] Liao Hongyi, Wang Xin. Stock Price Forecast Based on Extreme Learning Machine[J]. Computer and Modernization, 2014, 12: 20-22.
[3] Shu Wei, Gan Lei. Application of Extreme Learning Machine Method in Power Line Construction Cost Estimation[J]. Modern Electric Power, 2011, 8 (28): 79-83.
[4] Liu Nan, Wang Han. Ensemble based extreme learning machine[J]. IEEE Signal Processing Letters, 2010, 17 (8): 754 — 757.
[5] Lan Yuan, Soh Yengchai, Huang Guanghui. Constructive hidden nodes selection of extreme learning machine for regression[J]. Neurocomputing, 2010, 73 (16): 3191-3199.
[6] Wu Xiang, Qian Jian-sheng, Huang Cheng-hua, Zhang Li, Marcelo J. Colaço. Short-Term Coalmine Gas Concentration Prediction Based on Wavelet Transform and Extreme Learning Machine[J]. Mathematical Problems in Engineering, 2014, Vol. 2014.
[7] G. B. Huang, X. Ding, H. Zhou. Optimization method based extreme learning machine for classification[J]. Neurocomputing, 2010, 74: 155-163.
[8] Huang G B, Wang D H, Lan Y. Extreme learning machines: A survey[C]. International Journal of Machine Learning and Cybernetics, 2011, v2, n2, p 107-122.
[9] Sheng L, Lin F, Yao X, Hui W. Robust activation function and its application: Semi-supervised kernel extreme learning method[J]. Neurocomputing, 2014, 144.
[10] Liu Yue, Wang Fang. Network Traffic Prediction Based on Optimized Combined Kernel Extreme Learning Machine[J]. Computer Technology and Development, 2016, 8 (26): 74-77.
[11] Xiaowu, Li Kejun, Yu Xiaowei, Zhang Zhengfa, Yan Jie. Prediction of top oil temperature range of transformer based on nuclear extreme learning machine and Bootstrap method[J]. Chinese Journal of Electrical Engineering, 2017, 10 (37): 5822-5828.
[12] Qiang Shang, Ciyun Lin, Zhaosheng Yang et al. A Hybrid Short-Term Traffic Flow Prediction Model Based on Singular Spectrum Analysis and Kernel Extreme Learning Machine[J]. PLOS ONE, 2016, 11 (8).
[13] Yang Xiyun, Guan Wenyuan, Liu Yuqi, Xiao Yunqi. Wind power power power interval prediction method based on kernel group learning machine model based on particle swarm optimization[J]. Journal of China Electrical Engineering, 2015, 9 (35): 146-153.
[14] Liu Zhennan, Zhou Jingnan. Drought prediction based on online sequence-extreme learning machine[J]. Pearl River, 2018, 8 (39): 84-87.
[15] Yang Minhui. Research on travel time prediction algorithm based on online sequence extreme learning machine[D]. Nanhua University, 2017. Jingzhong Liu. Adaptive forgetting factor OS-ELM and bootstrap for time series prediction[J]. International Journal of Modeling, Simulation, and Scientific Computing, 2017, Vol. 08 (03), pp. 19.
[16] Wei Guo, Tao Xu, Keming Tang. M-estimator-based online sequential extreme learning machine for predicting chaotic time series with outliers[J]. Neural Computing and Applications, 2017, Vol. 28 (12), pp. 4093-4110.
[17] Lu Zhenli. Phase space reconstruction and network traffic prediction of regular extreme learning machine[J]. Laser Journal, 2015, 36 (1): 109-112.
[18] YUAN Wei, YAN Jia-jin, WANG Wen-xia, HUANG Nan-tian. Short-term wind speed prediction based on weighted regular extreme learning machine[J]. Hydropower Energy Science, 2017, 5 (35): 209-211.
[19] Zhang, Xian, Wang, Hongli. Time series prediction based on sequential regularized extreme learning machine and its application[J]. Hangkong Xuebao/Acta Aeronautica et Astronautica Sinica, 2011, 7(32): 1302-1308.

[20] Futian Weng, Tianle Zhang, Muzhou Hou, Jianshu Luo. PM2.5 Prediction Based on Genetic Algorithm and Regularized Extreme Learning Machine[J]. Computer Science and Application, 2018, Vol. 08 (08), pp. 1207-1216.

[21] Xu Yao. Application of Grey Extreme Learning Machine in Rolling Bearing Fault Prediction[J]. Computer Measurement and Control, 2017, 25(7): 63-69.

[22] Yan Jinwei. Research on Grey Extreme Learning Machine Prediction Algorithm[D]. Harbin Engineering University, 2014.

[23] Fan Chao, Cao Peige, Guo Yafei, Yang Tiejun. Food Production Forecast Based on Gray Limit Learning Machine[J]. Jiangsu Agricultural Sciences, 2018, 46 (05): 212-214.