

Research Article

A Novel Improved ELM Algorithm for a Real Industrial Application

Hai-Gang Zhang, Sen Zhang, and Yi-Xin Yin

School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China

Correspondence should be addressed to Sen Zhang; zhangsen@ustb.edu.cn

Received 2 December 2013; Accepted 29 January 2014; Published 16 April 2014

Academic Editor: Ramachandran Raja

Copyright © 2014 Hai-Gang Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

It is well known that the feedforward neural networks meet numbers of difficulties in the applications because of its slow learning speed. The extreme learning machine (ELM) is a new single hidden layer feedforward neural network method aiming at improving the training speed. Nowadays ELM algorithm has received wide application with its good generalization performance under fast learning speed. However, there are still several problems needed to be solved in ELM. In this paper, a new improved ELM algorithm named R-ELM is proposed to handle the multicollinear problem appearing in calculation of the ELM algorithm. The proposed algorithm is employed in bearing fault detection using stator current monitoring. Simulative results show that R-ELM algorithm has better stability and generalization performance compared with the original ELM and the other neural network methods.

1. Introduction

In the last few years, feedforward neural networks have received very wide range of applications and development. However, its applications encountered a lot of restrictions because of the slow training speed [1–3]. The most widely used method in feedforward neural network is gradient-based learning algorithm. Huang et al. pointed out that all the parameters of networks are tuned iteratively by using gradient-based learning algorithm, which is the main reason to lead slow training speed. On the other hand, feedforward neural network has a very high chance to fall into local minima [4]. In order to deal with these two shortcomings, Huang et al. proposed extreme learning machine (ELM) algorithm. ELM is a novel single hidden layer feedforward neural network where the input weights and the bias of the hidden nodes are generated randomly without tuning and the output weight is determined analytically. Compared with traditional feedforward network learning algorithms such as back-propagation (BP) algorithm, ELM has the following advantages: (1) the fast training speed, (2) good generalization performance, (3) partially overcoming the problem of

local minima, and (4) not needing manual intervention like setting a stop criterion or learning rate. In summary, ELM algorithm overcomes several shortcomings of feedforward network. Unfortunately, the multicollinear problem in ELM algorithm becomes a restriction in many complex industrial processes, which deteriorate its generalization performance. Although in [5, 6], Huang et al. proposed the multicollinear problem appearing in calculation and gave a simple method based on ridge regression, it lacks the rigorous mathematical proof and a more rational approach requires research, which motivates us.

Induction motors play a pivotal role in industrial production. The occurrence of motor faults often cause a lot of damage to property, even threat to the lives of workers [7, 8]. Figure 1 presents the occurrence frequencies of different motor faults, showing that bearing faults are the most prone to appearing ones occupying about 40%–50% of motor damages [9]. Bearing fault detection is a challenging and great significant problem, which attracted many researchers' attention and thus lots of variable results are proposed recently [10–14]. Detection of bearing fault using vibration signals is a very popular way [10, 11]. However, sound

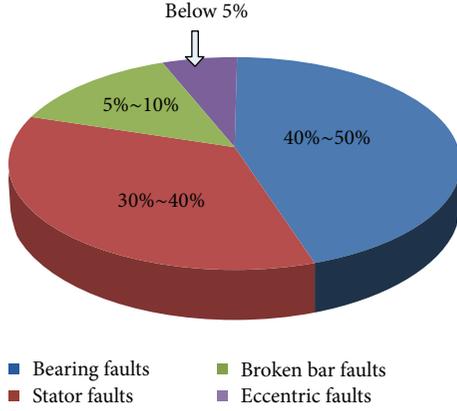


FIGURE 1: The probability map of different types of faults in induction motors.

senor installation will change the motor's mechanical body, increasing the difficulty of on-line implementation. Until the late 20th century, artificial neural networks (ANNs) were employed in bearing fault detection [12]. The shortcomings, like being time-consuming and easy to converge to local minima, limited its promotion. Recently, many researchers have realized that bearing faults will affect the stator current spectrum [13, 14]. In this paper, bearing fault detection using stator current monitoring is employed to verify the effective and robust performance of our proposed ELM algorithm.

In this paper, a new improved ELM algorithm is proposed in order to overcome the multicollinear problem. Compared with the traditional ELM algorithm, it has less covariance (V) and mean square error (MSE). In order to test and verify its effectiveness, the improved ELM algorithm is employed in bearing fault detection using stator current monitoring. Furthermore, our improved ELM algorithm shows the outstanding fault identification ability compared with other common classification algorithms, such as BP and SVM.

The following sections are organized as follows: the basic theory of the proposed improved ELM algorithm is presented in Section 2. Section 3 describes the current research of bearing faults. Simulative experiments are given in Section 4. Section 5 summarizes the conclusion of this paper.

2. The Improved Extreme Learning Machine Algorithm

Extreme learning machine (ELM) is an efficient training algorithm to deal with single-hidden-layer feedforward neural networks (SLFNs). ELM randomly generates the input weights and the bias of hidden nodes while determining the output weights according to the theory of least square (LS). It can quickly reach the training results without maneuvering intervention and get rid of the puzzle of local minimum simultaneously [1, 4]. So far, ELM algorithm has received wide application like directing multicategory classification problems in the cancer diagnosis area [15], classifying mental tasks from different subjects [16] and fault diagnosis [17]. According to the theory of Huang et al., ELM algorithm

can also be applied in human action recognition, location positioning system, human computer interface, security and data privacy, and so on. As a result of the theory of least squares and the random generation of input weights together with the bias of hidden nodes, ELM algorithm suffers the multicollinear problems which shows that the results have large variance and mean square error. This section proposes an improved ELM algorithm named R-ELM to overcome the multicollinear problems.

2.1. The Review of Ordinary ELM Algorithm. Suppose there are n samples x_i, t_i , where $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in R^n$ denotes an n -dimensional features of the i th sample and $t_i = [t_{i1}, t_{i2}, \dots, t_{im}] \in R^m$ denotes the target vector. The mathematical model of SLFNs with \tilde{N} hidden nodes is as follows:

$$\sum_{i=1}^{\tilde{N}} \beta_i g_i(x_k) = \sum_{i=1}^{\tilde{N}} \beta_i g(w_i \cdot x_k + b_i) = o_k, \quad k = 1, 2, \dots, m, \quad (1)$$

where $W = [w_1, w_2, \dots, w_{\tilde{N}}]^T$ is $\tilde{N} \times n$ input weight matrix connecting the hidden and the input nodes, $\beta = [\beta_1, \beta_2, \dots, \beta_{\tilde{N}}]^T$ is $\tilde{N} \times m$ output weight matrix connecting the hidden and the output nodes, and $B = [b_1, b_2, \dots, b_{\tilde{N}}]^T$ is $\tilde{N} \times 1$ bias of hidden layer nodes. And $w_i \cdot x_i$ denotes the inner product of w_i and x_i and $O = [o_1, o_2, \dots, o_m]^T$ is $m \times 1$ output vector. $g(x)$ represents the activation function here.

ELM algorithm aims at finding an optimal solution β which has the minimum mean square error (MSE) such that

$$\sum_{i=1}^{\tilde{N}} \beta_i g(w_i \cdot x_k + b_i) = t_k, \quad k = 1, 2, \dots, N,$$

$$\text{s.t. } \text{MSE}(\hat{\beta}) = \frac{1}{n} \sum_{i=1}^n (\hat{\beta}_i - \beta)^2 = \frac{1}{N} \sum_{k=1}^N \|o_k - t_k\| \rightarrow \min, \quad (2)$$

where $\hat{\beta}$ is the estimation of β .

The above N equations can be written in matrix form as

$$H\beta = T, \quad (3)$$

where

$$H(W, B, X) = \begin{pmatrix} g(w_1 \cdot x_1 + b_1) & \cdots & g(w_{\tilde{N}} \cdot x_1 + b_{\tilde{N}}) \\ \vdots & \ddots & \vdots \\ g(w_1 \cdot x_N + b_1) & \cdots & g(w_{\tilde{N}} \cdot x_N + b_{\tilde{N}}) \end{pmatrix}_{N \times \tilde{N}} \quad (4)$$

According to the theory of least square, the output weight β can be estimated as

$$\hat{\beta} = H^+ T, \quad (5)$$

where H^+ is the Moore-Penrose generalized inverse of H .

There are several methods to calculate the Moore-Penrose generalized inverse. ELM algorithm makes use of singular value decomposition (SVD), where $H^+ = (H^T H)^{-1} H^T$. So

$$\hat{\beta} = (H^T H)^{-1} H^T T. \quad (6)$$

The ELM algorithm steps can be summarized as follows.

Step 1. Assign arbitrary input weights W and bias of hidden layer nodes B .

Step 2. Calculate the hidden layer output matrix H .

Step 3. Calculate the output weights: $\hat{\beta} = (H^T H)^{-1} H^T T$.

The above is the theory of ordinary ELM algorithm. Some properties of the solution will be discussed [18, 19]. Considering the noisy environment in real industrial process when obtaining the data, the model (3) should be modified:

$$T = H\beta + \varepsilon, \quad (7)$$

where $\varepsilon \in N(0, \sigma^2)$ represents the model uncertainty or noise disturbance.

The modified solution is

$$\hat{\beta} = (H^T H)^{-1} H^T T = \frac{\sum_{i=1}^n H_i^T (H_i \beta_i + \varepsilon_i)}{\sum_{i=1}^{\tilde{N}} H_i^2} = \beta + \frac{\sum_{i=1}^{\tilde{N}} H_i^T \varepsilon_i}{\sum_{i=1}^{\tilde{N}} H_i^2}, \quad (8)$$

- (1) $E(\hat{\beta}) = \beta$,
- (2) $V(\hat{\beta}) = E(\hat{\beta}^2) - E(\hat{\beta})^2 = \sigma^2 / \sum_{i=1}^{\tilde{N}} H_i^2 = \sigma^2 \sum_{i=1}^{\tilde{N}} (1/\lambda_i)$
where λ_i is the i th eigenvalue of $H^T H$,
- (3) $\text{MSE}(\hat{\beta}) = (1/\tilde{N})E[(\hat{\beta} - \beta)^T(\hat{\beta} - \beta)] = (1/\tilde{N})(\sigma^2 / \sum_{i=1}^{\tilde{N}} H_i^2) = (1/\tilde{N})\sigma^2 \sum_{i=1}^{\tilde{N}} (1/\lambda_i)$.

Suppose $H^T H$ may not always be nonsingular; that is to say when the H matrix is multicollinear, some eigenvalues will tend to zero, while $V(\hat{\beta})$ and $\text{MSE}(\hat{\beta})$ will become larger, which will affect its stability and generalization.

Generally speaking, the data obtained from the field tend to have the presence of multicollinear problems. In the next subsection, one of the main results in our paper will be presented. Like the theory of Ridge Regression to overcome the multicollinear problem in least square method, we call our improved algorithm R-ELM.

2.2. The Improved R-ELM Algorithm. As discussed above, the ELM algorithm encounters severe instability and bad performance while meeting the multicollinear data. In our proposed R-ELM algorithm, LDL^T decomposition is applied against the symmetric matrix $H^T H$. LDL^T decomposition (also called LDL^T factorization) factors a symmetric matrix as the product of a lower triangular matrix (L), a diagonal matrix (D), and the transpose matrix of the first lower triangular matrix (L^T). During decomposition, we set the minimum threshold to some singular elements of D matrix, which gets rid of the puzzle of multicollinear problem.

The LDL^T decomposition of a symmetric matrix ($H^T H$) is as follows:

$$H^T H = LDL^T, \quad (9)$$

where

$$L = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ l_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{\tilde{N}1} & l_{\tilde{N}2} & \cdots & 1 \end{bmatrix}, \quad D = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_{\tilde{N}} \end{bmatrix}. \quad (10)$$

Each element of L and D are sequentially calculated using an iterative approach; that is,

$$d_k = a_{kk} - \sum_{m=1}^{k-1} u_{km} l_{km},$$

$$u_{jk} = a_{jk} - \sum_{m=1}^{k-1} u_{jm} l_{km}, \quad j = k+1, k+2, \dots, \tilde{N}, \quad (11)$$

$$l_{jk} = \frac{u_{jk}}{d_k}, \quad k = 1, 2, \dots, \tilde{N},$$

where a_{jk} is the elements of the ordinary matrix $H^T H$.

After LDL^T decomposition, whether the original matrix has multicollinear problem is determined by the values of matrix D . If values of some elements are close to zero, the ordinal matrix is multicollinear. In order to obtain robust matrix, a modified algorithm to calculate the values of L and D is given as follows:

$$d_k = \begin{cases} a_{kk} - \sum_{m=1}^{k-1} u_{km} l_{km}, & \text{if } \left| a_{kk} - \sum_{m=1}^{k-1} u_{km} l_{km} \right| > \varepsilon_0, \\ \text{sgn} \left[a_{kk} - \sum_{m=1}^{k-1} u_{km} l_{km} \right] \cdot \varepsilon_0, & \text{else,} \end{cases} \quad (12)$$

$$u_{jk} = a_{jk} - \sum_{m=1}^{k-1} u_{jm} l_{km}, \quad j = k+1, k+2, \dots, \tilde{N},$$

$$l_{jk} = \frac{u_{jk}}{d_k}, \quad k = 1, 2, \dots, \tilde{N},$$

where ε_0 is an appropriate positive number.

After decomposition, the output weights can be calculated in the following new way:

$$\hat{\beta} = (LDL^T)^{-1} H^T T = \lambda_0 \beta + \frac{\sum_{i=1}^{\tilde{N}} H_i^T \varepsilon_i}{\sum_{i=1}^{\tilde{N}} L_i D_i L_i^T}, \quad (13)$$

where $\lambda_0 = \sum_{i=1}^{\tilde{N}} H_i^T H_i / \sum_{i=1}^{\tilde{N}} L_i D_i L_i^T$, whose value is close to 1.

Assume that the matrix D can be decomposed into two parts. That is, $D = D^{1/2} \cdot (D^{1/2})^T$, and if some elements have negative values, complex decomposition is to be considered. Then recalculate performance indicators:

- (1) $E(\hat{\beta}) = \lambda_0 \beta$,
- (2) $V(\hat{\beta}) = E(\hat{\beta}^2) - E(\hat{\beta})^2 = \lambda_0 \sigma^2 / \sum_{i=1}^{\tilde{N}} L_i D_i^{1/2} (L_i D_i^{1/2})^T = \lambda_0 \sigma^2 \sum_{i=1}^{\tilde{N}} (1/\bar{\lambda}_i)$, where $\bar{\lambda}_i$ is the i th eigenvalue of LDL^T ,

$$(3) \text{MSE}(\hat{\beta}) = (1/\bar{N})E[(\hat{\beta} - \beta)^T(\hat{\beta} - \beta)] = (1/\bar{N})[(\lambda_0 - 1)\beta^2 + \lambda_0\sigma^2 \sum_{i=1}^{\bar{N}}(1/\bar{\lambda}_i)].$$

Setting a threshold to the values of every element in matrix D can make sure that $|V(\hat{\beta})| \leq \bar{N}\lambda_0\sigma^2/\varepsilon_0$ and $|\text{MSE}(\hat{\beta})| \leq 1/\bar{N}[|(\lambda_0 - 1)\beta^2| + \bar{N}\lambda_0\sigma^2/\varepsilon_0]$. Additionally, the estimation $\hat{\beta}$ is not unbiased as compensation ($\lambda_0 \neq 1$).

Remark 1. Unlike Huang's method based on Ridge Regression to add a threshold to every element of matrix $H^T H(\hat{\beta} = (H^T H + 1/C)^{-1} H^T T)$, the proposed R-ELM algorithm just set a proper threshold to some singular element to matrix D after LDL^T decomposition to $H^T H$. In contrast, our improved algorithm seems more reasonable.

3. Stator Current Feature Extraction of Bearing Fault

The majority of electrical machines employ ball or rolling bearings, which consist of two rings—one inner and the other outer. Even under normal operating conditions of balanced load and good alignment, fatigue failure can still take place from the location below the raceway and rolling elements to the surface, which may lead to increased vibration and noise levels. With the increase of continued stressing, flaking or spalling of bearing might occur, while fragments of the material tend to break loose [7, 9, 10].

Many researchers obtain vibration signals from the motor and then extract spectrum information through FFT or wavelet transform to identify the bearing faults. Considering the inconvenience of collection and analysis of vibration, this paper employs the stator current signals. The relation between vibration signal and current signal is as follows:

$$f_{bng} = |f_e \pm m \cdot f_v|, \quad (14)$$

where $m = 1, 2, \dots, f_e$ is the electrical supply frequency, f_v is one of the characteristic vibration frequencies, and f_{bng} is the characteristic frequencies reflected in stator current spectrum.

There are many kinds of bearing faults. Here we consider two main defects:

- (1) for outer bearing race defects

$$f_{vo} = \frac{n}{2}f_r \left[1 - \frac{b_d}{p_d} \cos(\beta) \right], \quad (15)$$

- (2) for inner bearing race defects

$$f_{vi} = \frac{n}{2}f_r \left[1 + \frac{b_d}{p_d} \cos(\beta) \right], \quad (16)$$

where b_d and p_d are the ball diameter and the bearing pitch diameter, respectively, f_r is the rotational speed in Hertz, n is the number of balls in a bear, and β is the contact angle of the balls on the races. Figure 2 is a legend to operating parameters of ball bearing.

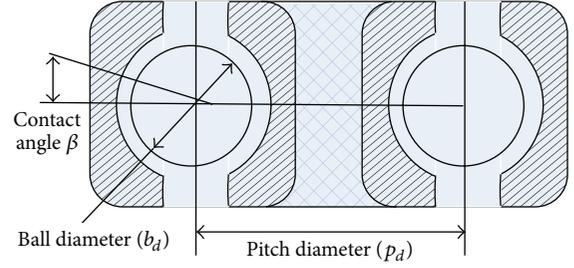


FIGURE 2: Ball bearing dimensions.

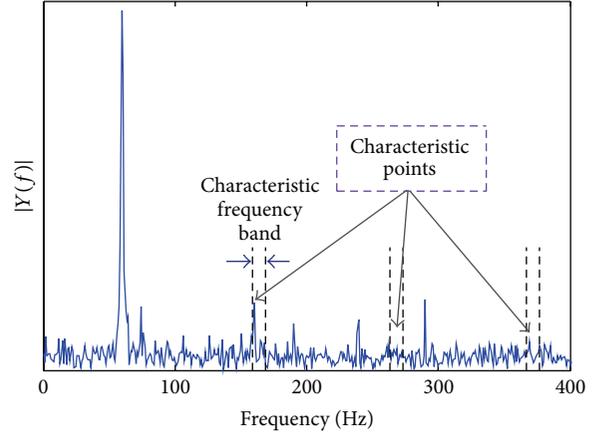


FIGURE 3: The stator current spectrums ($\varepsilon = 5$): healthy operation.

Here an experiment is given to prove that bearing fault can be checked through the monitoring of stator current. A four-pole test motor with nine balls ($n = 9$) in the bearing is employed. Other parameters about the motor can be found from the data sheet [13]: the pitch diameter is about 60 mm ($p_d = 60$) and the ball diameter equals 12 mm ($b_d = 12$) approximately. Assume that the experimental motor runs at the rated shafted speed of 1735 rpm ($f_r = 28.9$ Hz) and has a contact angle ($\beta = 1$). So $f_{vo} = 104$ Hz and $f_{vi} = 156$ Hz based on (15) and (16). For broken bearing, we drill a hole through the outer race and the inner broken races are not considered here. Figures 3 and 4 show the experimental results. As shown in Figures 3 and 4, the amplitudes of three characteristic frequency points above the fundamental frequency of stator current are extracted as the inputs of R-ELM algorithm. In order to get rid of the interface of noise, we set the characteristic points who have the highest spectrum amplitude during characteristic frequency band [$f_{bng} - \varepsilon, f_{bng} + \varepsilon$], where ε represents the bandwidth.

4. Simulative Results for Bearing Fault Detection

This section presents the simulation results referring to the proposed R-ELM algorithm applied in bearing fault detection. An experiment motor with bearing fault mentioned above section is employed. The process of simulation experiments is shown in Figure 5.

TABLE 1: The choice of active functions.

Performance indicators	sig	sin	hardlim	radbas
Training time (s)	0.023	0.028	0.098	0.054
Testing time (s)	0.003	0.015	0.077	0.032
Training accuracy (%)	98.2	98.1	63.5	78.1
Testing accuracy (%)	99.5	98.4	70.6	81.2

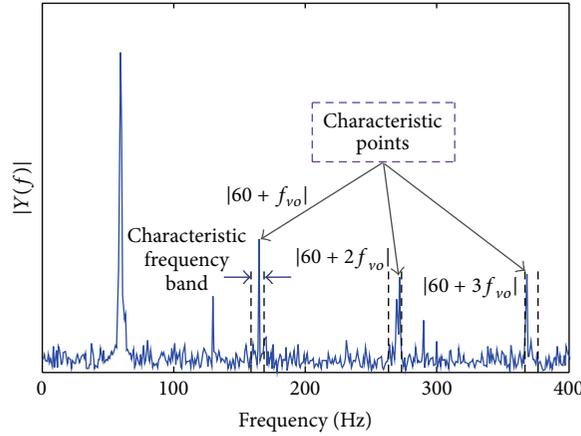


FIGURE 4: The stator current spectrums ($\epsilon = 5$): faulty operation.

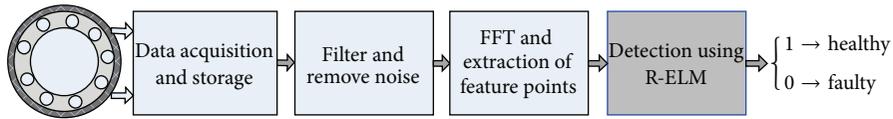


FIGURE 5: Schematic diagram of simulation experiments.

4.1. *Experimental Preparation.* All the simulations have been conducted in Matlab 7.8.0 running on a desktop PC with AMD Athlon II X2 250 processor, 3.00-GHz CPU, and 2 G RAM. Two operating conditions are considered here: normal operation and operation with current noise pollution. Under 1 KHz sampling, 900 samples (500 for training and 400 for testing) are extracted from the current spectrum of healthy and faulty bearings under each operating conditions. And all the data are normalized into $(-1, 1)$.

There are four performance indicators to measure the quality of ELM algorithm: training time, testing time, train accuracy, and test accuracy. In addition, the variance and mean square error are inversely proportional to train and test accuracy. In general, active functions play an important role in computing of neural networks. Table 1 lists the comparison results among four common active functions (sig, sin, hardlim, and radbas) under the four performance indicators using our sampling data, where we can see that the sig function has more outstanding performance than other ones. Then the proposed E-ELM algorithm applies sig function as active function.

4.2. *The Main Experimental Results.* Figure 6 presents the stator current signals under two operating conditions, where

dashed box indicates the faulty current signals. White noise is added in the current under normal operation through current transformer to test the robust performance of proposed algorithm. Figure 7 depicts the sampling data distribution of both healthy and faulty bearings under normal operating condition. As previously described, the stator current spectrum information from three characteristic frequency bands are sampled and extracted. Symbols with white shade represent the healthy operating condition, while the black symbols mean something is wrong with the experiment bearing. Following work is the classification of R-ELM. Table 2 shows the simulation results of R-ELM classifier in different operating situations, where we can see that R-ELM algorithm is able to achieve a reliable classification accuracy satisfying the need of on-line fault detection.

Table 3 lists the simulation results of comparing the proposed R-ELM algorithm with other common method include ELM under normal operation. R-ELM algorithm shows an outstanding classification ability with less time-consuming and high recognition rate than BP and SVM algorithms. Referring to the comparing results between ELM and R-ELM, we can see that R-ELM has less variance and mean square error (equals higher recognition rate) than ordinary ELM algorithm despite of a little more time-consuming.

TABLE 2: Simulation results of classification using R-ELM for bearing fault detection.

Operation conditions	Training accuracy (%)			Testing accuracy (%)			Training time	Testing Time
	Healthy bearing	Faulty bearing	Total	Healthy bearing	Faulty bearing	Total		
Normal	96.15	96.15	96.15	97.50	94.83	96.17	0.0589	0.0312
Interferential	92.63	92.28	92.46	95.42	90.42	92.92	0.0612	0.0301

TABLE 3: Comparison of results of different classifiers.

Subjects	Training sections		Testing sections	
	Fault recognition rate (%)	Time (s)	Fault recognition rate (%)	Time (s)
R-ELM	96.15	0.0589	96.17	0.0312
BP	90.32	5.3280	90.44	2.1032
SVM	95.43	3.4320	95.30	1.3890
ELM	95.15	0.0156	94.83	0.0058

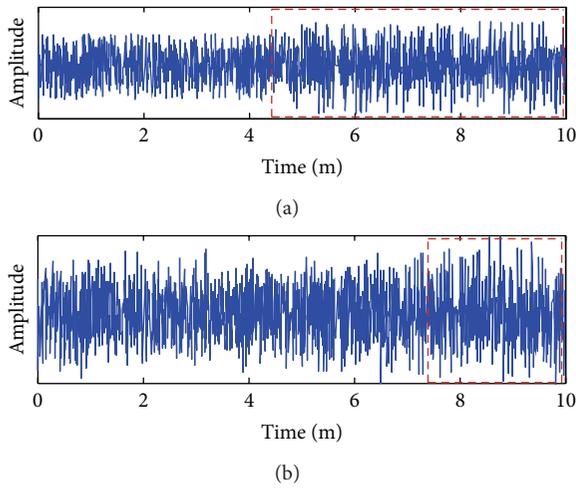


FIGURE 6: The stator current signals: (a) normal operation; (b) operation with noise interface.

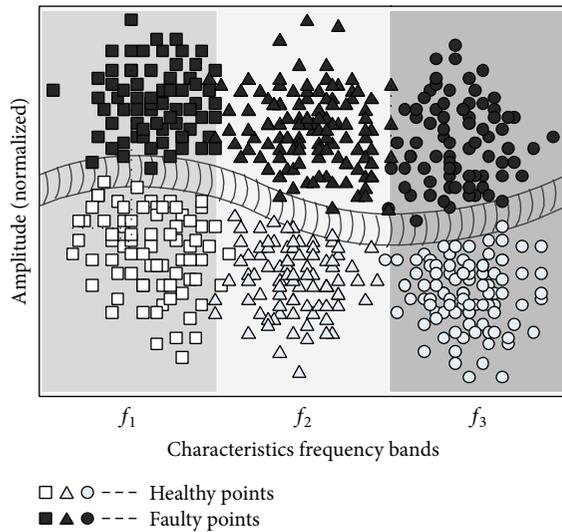


FIGURE 7: The distribution of sampling data under normal operation.

5. Conclusions

This paper proposes an improved ELM algorithm named R-ELM. Through the LDL^T decomposition of matrix $H^T H$ and setting limited values of the elements in matrix D , the proposed R-ELM algorithm can deal with the multicollinear problems in the application of ELM. We employ R-ELM algorithm in bearing fault detection and two operating conditions are considered. This algorithm shows better performance in fault identification. Compared with other neural network methods, R-ELM algorithm takes less time and has better performance. Compared to the original ELM algorithm, the proposed R-ELM improves the recognition rate (reduces the variance and mean square error) with trade of taking a little more time.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work has been supported by the National Natural Science Foundation of China (NSFC Grant no. 61333002) and Beijing Natural Science Foundation (Grant no. 4132065).

References

- [1] H. Zhang, Z. Wang, and D. Liu, "Global asymptotic stability of recurrent neural networks with multiple time-varying delays," *IEEE Transactions on Neural Networks*, vol. 19, no. 5, pp. 855–873, 2008.
- [2] H. Zhang, Z. Liu, G.-B. Huang, and Z. Wang, "Novel weighting-delay-based stability criteria for recurrent neural networks with time-varying delay," *IEEE Transactions on Neural Networks*, vol. 21, no. 1, pp. 91–106, 2010.
- [3] H. Zhang and Y. Wang, "Stability analysis of Markovian jumping stochastic Cohen-Grossberg neural networks with mixed

- time delays,” *IEEE Transactions on Neural Networks*, vol. 19, no. 2, pp. 366–370, 2008.
- [4] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme learning machine: theory and applications,” *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006.
- [5] G. B. Huang, H. M. Zhou, X. J. Ding, and R. Zhang, “Extreme learning machine for regression and multiclass classification,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 42, no. 2, pp. 513–529, 2012.
- [6] G.-B. Huang, D. H. Wang, and Y. Lan, “Extreme learning machines: a survey,” *International Journal of Machine Learning and Cybernetics*, vol. 2, no. 2, pp. 107–122, 2011.
- [7] H. M. Liu, J. Wang, and C. Lu, “Rolling bearing fault detection based on the teager energy operator and Elman neural network,” *Mathematical Problems in Engineering*, vol. 2013, Article ID 498385, 10 pages, 2013.
- [8] H. Zhang, D. Liu, Y. Luo, and D. Wang, *Adaptive Dynamic Programming for Control: Algorithms and Stability*, Communications and Control Engineering Series, Springer, London, UK, 2013.
- [9] S. Nandi, H. A. Toliyat, and X. Li, “Condition monitoring and fault diagnosis of electrical motors—a review,” *IEEE Transactions on Energy Conversion*, vol. 20, no. 4, pp. 719–729, 2005.
- [10] P. Konar and P. Chattopadhyay, “Bearing fault detection of induction motor using wavelet and Support Vector Machines (SVMs),” *Applied Soft Computing Journal*, vol. 11, no. 6, pp. 4203–4211, 2011.
- [11] J. R. Stack, T. G. Habetler, and R. G. Harley, “Effects of machine speed on the development and detection of rolling element bearing faults,” *IEEE Power Electronics Letters*, vol. 1, no. 1, pp. 19–21, 2003.
- [12] R. R. Schoen, B. K. Lin, T. G. Habetler, J. H. Schlag, and S. Farag, “Unsupervised, on-line system for induction motor fault detection using stator current monitoring,” *IEEE Transactions on Industry Applications*, vol. 31, no. 6, pp. 1280–1286, 1995.
- [13] R. R. Schoen, T. G. Habetler, F. Kamran, and R. G. Bartheld, “Motor bearing damage detection using stator current monitoring,” *IEEE Transactions on Industry Applications*, vol. 31, no. 6, pp. 1274–1279, 1995.
- [14] S. Chen and T. A. Lipo, “Bearing currents and shaft voltages of an induction motor under hard- and soft-S witching inverter excitation,” *IEEE Transactions on Industry Applications*, vol. 34, no. 5, pp. 1042–1048, 1998.
- [15] R. Zhang, G.-B. Huang, N. Sundararajan, and P. Saratchandran, “Multicategory classification using an extreme learning machine for microarray gene expression cancer diagnosis,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 4, no. 3, pp. 485–494, 2007.
- [16] N.-Y. Liang, P. Saratchandran, G.-B. Huang, and N. Sundararajan, “Classification of mental tasks from EEG signals using extreme learning machine,” *International Journal of Neural Systems*, vol. 16, no. 1, pp. 29–38, 2006.
- [17] X.-F. Hu, Z. Zhao, S. Wang, F.-L. Wang, D.-K. He, and S.-K. Wu, “Multi-stage extreme learning machine for fault diagnosis on hydraulic tube tester,” *Neural Computing and Applications*, vol. 17, no. 4, pp. 399–403, 2008.
- [18] R. Uemukai, “Small sample properties of a ridge regression estimator when there exist omitted variables,” *Statistical Papers*, vol. 52, no. 4, pp. 953–969, 2011.
- [19] A. E. Hoerl and R. W. Kennard, “Ridge regression: biased estimation for nonorthogonal problems,” *Technometrics*, vol. 42, no. 1, pp. 80–86, 2000.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

