Efficient Hidden Danger Prediction for Safety Supervision System: An Advanced Neural Network Learning Method

Zhigang Zhao, Yongfeng Wei, Xinyan Wang, Ruixin Li and Jing Deng

1 Introduction

Safety production of enterprises is an important issue for sustainable development. Hidden danger prediction plays an important role in ensuring the safety and efficient development of enterprises [1]. By providing effective prediction method for hidden danger data, enterprises can know their future safety production situation and possible changes in advance, so as to take measures to control [2]. And safety supervision department can evaluate the hidden danger management ability of enterprises by predicting the trend of hidden danger situation. Aiming at the influence of the relevant management index on the trend of hidden danger, this paper proposes a new learning algorithm based on management index for the number of hidden danger.

Generally, the prediction methods of the number of hidden danger can be divided into two categories. One is specialist analytical prediction according to the artificial experience. Such methods can play a positive role in some special occasions, but the efficiency will decreases rapidly when the number of hidden danger index becomes larger. Another is automatic prediction based on model driven or data driven methods. This kind of methods is less used in practical situations

Z. Zhao · Y. Wei · X. Wang
Communication and Information Center, State Administration
of Work Safety, Beijing 100013, China

R. Li · J. Deng (✉)
School of Computer and Communication Engineering,
University of Science and Technology Beijing (USTB),
30 Xueyuan Road, Haidian District, Beijing 100083, China
e-mail: informationscience@126.com

R. Li · J. Deng
Beijing Key Laboratory of Knowledge Engineering for Materials Science,
30 Xueyuan Road, Haidian District, Beijing 100083, China

Z. Deng (ed.), *Proceedings of 2017 Chinese Intelligent Automation Conference*,
Lecture Notes in Electrical Engineering 458,
https://doi.org/10.1007/978-981-10-6445-6_51

because of their complex prediction models. Aiming at hidden danger data source, the automatic prediction based on data driven method can be widely used because it carries out optimal learning through some artificial intelligence methods without accurate mathematical model [3].

Different artificial intelligence algorithms can be used flexibly in data driven prediction methods [4, 5]. Recently, the prediction learning method based on neural network (NN) has witnessed a growing interest [6, 7]. This paper mainly implements hidden danger prediction using an advanced NN learning algorithm. The main contributions are as follows.

(1) At present, hidden danger prediction is mainly applied to coal chemical industry and other special industries. However, in some small-medium cities, the hidden danger prediction accuracy of medical treatment and other tertiary industries does not require as high as coal chemical industry, so there may exist some inaccurate data. Data driven hidden danger prediction algorithms for such industries is still scarce. This paper proposes an advanced algorithm to predict hidden danger for these tertiary industries.
(2) In the hidden danger prediction applications, the existing NN prediction algorithm is mainly implemented on the basis of Back-propagation (BP) algorithm [3]. In practice, there exist a lot of problems, such as over fitting and low learning precision. Recently, researchers have proposed various ways to improve the efficiency, e.g., extreme learning machine (ELM) [8], and achieved good results. Motivated by it, this paper discusses the application for the above areas using ELM.

2 Backgrounds

2.1 Theoretical Basis of Hidden Danger Prediction

Safety supervision is to avoid the occurrence of accidents and the damage caused by the accident. Only by fully understanding the causes and rules of the accident, can we make the production reach a safe state. Accident-causing theory is a kind of theory to grasp the law of accident occurrence [9]. By analyzing the original data, we can find out the potential impact of the hidden danger and the law of its development, so as to provide a favorable protection for safety production.

Accident-causing theory is of great significance to the study of the mechanism of accident, analysis and elimination of hidden danger, accident prevention and so on. At present, there are more than ten kinds of theories related to accident-causing theories, and the typical accident-causing theories include the accident causation sequence theory, the human error model, the orbit intersecting theory, the energy transfer theory and so forth [10]. Modern accident-causing theory is the model of action mechanism extracted from many typical accidents [10]. It reveals a certain rule of the accident, provides an effective and scientific reference basis to perfect and improve the work of safety management, and is helpful to enhance the ability of enterprise's hidden danger investigation and governance.

In this paper, the application of ELM based hidden danger prediction method can be regarded as a novel practice of accident-causing theory.

2.2 ELM

ELM is a learning algorithm for single-hidden layer feedforward NN, and its schematic diagram is shown in Fig. 1 [8].

The output function of a single hidden layer feedforward NN with L hidden layer neurons is:

$$f_L(x) = \sum_{i=1}^{L} \beta_i G(a_i, b_i, x) \tag{1}$$

where a_i and b_i are hidden node parameters, β_i is the output weight vector between the i-th hidden layer node and the output nodes, $G(a_i, b_i, x)$ is a nonlinear output function of the i-th hidden layer node corresponding to the sample x. For the additive type hidden layer nodes, $G(a_i, b_{i,} x) = g(a_i x + b)$, where g: $\mathbf{R} \rightarrow \mathbf{R}$ denotes activation function, $a_i\, x$ denotes the inner product of the inner weight vector a_i and sample x in $\mathbf{R}^n$.

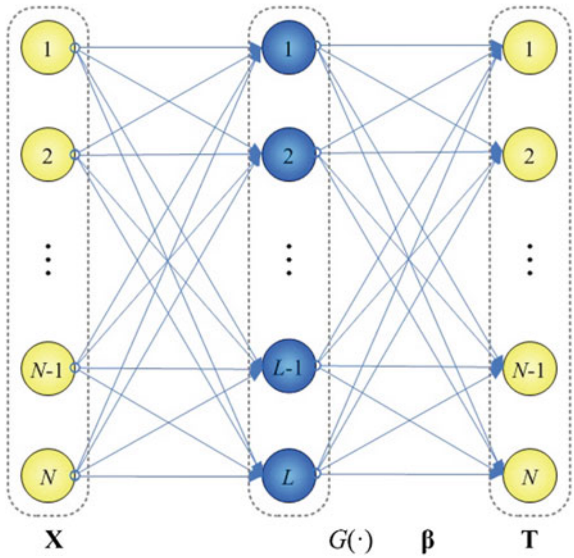

Fig. 1 The schematic diagram of ELM for a single hidden layer feedforward NN

For radial basis function (RBF) NN type hidden layer nodes, $G(a_i, b_i, x)$ can be expressed as $G(a_i, b_i, x) = g(b_i\|x - a_i\|)$, where a_i and b_i ($b_i > 0$) represent the center and influence factors of the i-th RBF node, respectively.

Considering N different data samples $\{(x_i, t_i)\}_{i=1}^{N} \subset \mathbf{R}^n \times \mathbf{R}^m$, if there is a single hidden layer NN with L hidden layer neurons, which can approximate the N different data samples with zero error, i.e., there exist a_i, b_i, β_i ($i = 1,2,\ldots, L$), making

$$f_L(x_j) = \sum_{i=1}^{L} \beta_i G(a_i, b_i, x_j) = t_j, \quad j = 1, \ldots, N \tag{2}$$

Here, (2) can be simplified as $\mathbf{H}\boldsymbol{\beta} = \mathbf{T}$, where

$$\mathbf{H}(a_1, \ldots, a_L, b_1, \ldots, b_L, x_1 \ldots, x_N) = \begin{bmatrix} G(a_1, b_1, x_1) & \cdots & G(a_L, b_L, x_1) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, x_N) & \cdots & G(a_L, b_L, x_N) \end{bmatrix}_{N \times L} \tag{3}$$

$$\boldsymbol{\beta} = \left(\left[\beta_1^{\mathrm{T}}, \ldots, \beta_L^{\mathrm{T}} \right]^{\mathrm{T}} \right)_{L \times m} \text{ and } \quad \mathbf{T} = \left(\left[t_1^{\mathrm{T}}, \ldots, t_N^{\mathrm{T}} \right]^{\mathrm{T}} \right)_{N \times m} \tag{4}$$

Here, $\mathbf{H}$ is called the hidden layer output matrix, the i-th column of $\mathbf{H}$ represents the output of the i-th hidden layer neuron corresponding to the input sample x_1, x_2, …, x_N, and the j-th row indicates the output of all hidden layer neurons corresponding to the input sample x_j.

But in most cases, the number of hidden layer neurons is far less than the number of training samples ($L \ll N$). It is difficult to achieve the objective of making the single hidden layer NN with L hidden layer neurons approximate the N different data samples with zero error. In this case, (2) can be further rewritten as $\mathbf{H}\boldsymbol{\beta} = \mathbf{T} + \mathbf{E}$, where $\mathbf{E} = \left([e_1^{\mathrm{T}}, \ldots, e_N^{\mathrm{T}}]^{\mathrm{T}} \right)_{N \times m}$.

Define the square loss function

$$J = \sum_{i=1}^{L} \left(\beta_i G(a_i, b_i, x_j) - t_j \right) \tag{5}$$

Then, (5) can be rewritten as $J = (\mathbf{H}\boldsymbol{\beta} - \mathbf{T})^{\mathrm{T}}(\mathbf{H}\boldsymbol{\beta} - \mathbf{T})$.

The training problem of network parameters is transformed into the problem of minimizing the square loss function. It is to find the least squares solution, making

$$\left\| \mathbf{H}\hat{\boldsymbol{\beta}} - \mathbf{T} \right\| = \min_{\boldsymbol{\beta}} \| \mathbf{H}\boldsymbol{\beta} - \mathbf{T} \| \tag{6}$$

If the output matrix of the hidden layer has full column rank, by utilizing Moore-Penrose generalized inverse, we can obtain

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \|\mathbf{H}\boldsymbol{\beta} - \mathbf{T}\| = \mathbf{H}^{\dagger}\mathbf{T} \tag{7}$$

where $\mathbf{H}^{\dagger} = (\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}$[8].

If the output matrix of the hidden layer does not have full column rank, the weight vector β_i can be obtained via singular value decomposition (SVD) method.

In the parameter training process of ELM, the parameters of hidden layer nodes are determined randomly. In practice, the samples should be normalized, and the hidden layer node parameter values are selected randomly in the interval $[-1, 1]$.

3 Hidden Danger Prediction Using ELM

According to the above analysis, the ELM algorithm framework for the prediction of special industry hidden danger can be described in Table 1 [8]. Here, the activation function of additive type hidden layer nodes in single hidden layer NN can be chosen as an arbitrary bounded nonlinear piecewise continuous function. But for the single hidden layer NN with RBF type hidden layer nodes, the activation function can be chosen as an arbitrary piecewise continuous integrable function.

4 Experimental Results and Discussions

4.1 Dataset

From nationwide enterprise hidden danger data provided by the State Administration of Work Safety, we find that the hidden danger investigation in some small-medium cities mainly aims at tertiary industries, and there may exist some inaccurate data because of the lack of high accuracy requirements. For example, Fig. 2 gives a summary of hidden danger data of a city in August 2014.

Table 1 The ELM algorithm framework for the prediction of special industry hidden danger

Input	Training set $\{(x_i, t_i)\}_{i=1}^{N} \subset \mathbf{R}^n \times \mathbf{R}^m$, activation function $G(x)$, the number of hidden layer nodes L
Output	Optimal the outer weight vector $\boldsymbol{\beta}$
(1)	Randomly generate input weight vector a_i and bias b_i
(2)	Calculate the output matrix $\mathbf{H}$ of hidden layer
(3)	Compute the optimal the output weight vector $\boldsymbol{\beta}$:$\boldsymbol{\beta} = \mathbf{H}^{\dagger}\mathbf{T}$

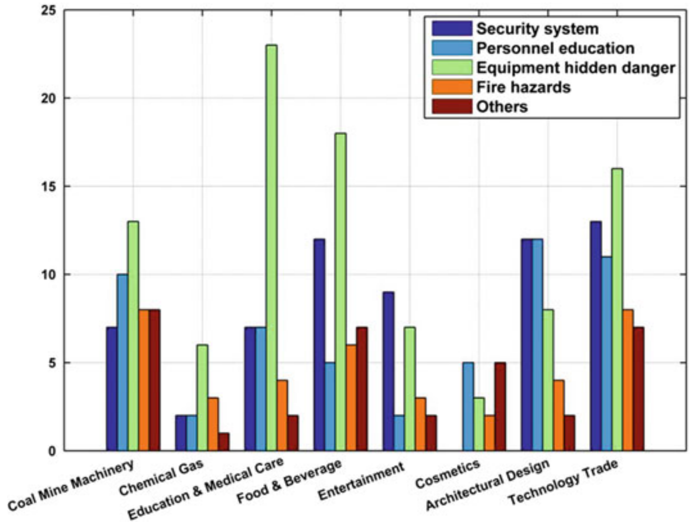

Fig. 2 The summary of hidden danger data of a city in August 2014

The experimental data is the spot check result of hidden danger of catering industry in a city from January to September 2014. We predict the total number of hidden danger in the following month according to historical data. Data obtained from the first 8 months and the 9th month constructs the training set and test set, respectively. The input terms are the inspection items of different locations, including offices, workshops, kitchens, and warehouses; the inspection items of hidden danger are security standard, personnel education, equipment hidden danger, and fire hidden danger. Thus, the input of training set is a high-dimensional numerical matrix.

4.2 Case 1

In the absence of some data, Fig. 3 gives a comparison of the prediction result between BP algorithm and ELM after running 5 times. It is obvious that ELM has higher prediction accuracy than BP. Moreover, in practice, ELM spends less time than BP. It illustrates the feasibility of ELM algorithm in practice.

4.3 Case 2

In order to further verify the robustness of algorithm, the experimental data are processed by partial perturbation.

Safety supervision personnel may make mistakes in statistics because of the huge amount of data and the complex relationship between hidden danger items and

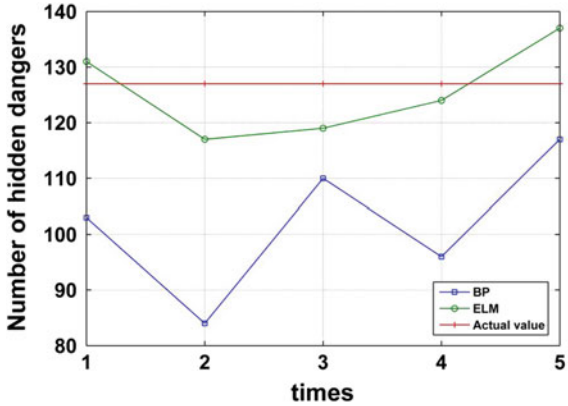

Fig. 3 The prediction results of ELM and BP neural network

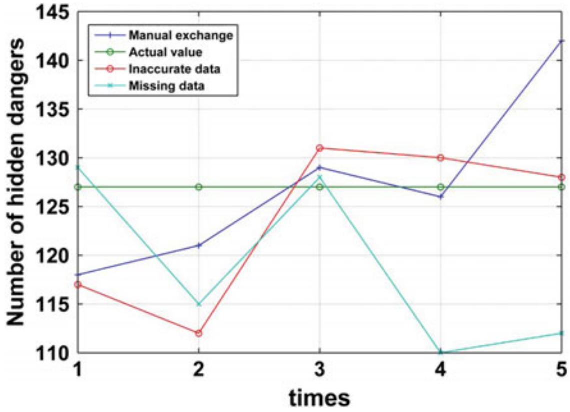

Fig. 4 The prediction results of ELM under different artificial intervention

locations in the actual production. In view of the fact that the data is not accurate in the actual production, 5% of the original input data is added or subtracted in a certain range to simulate the inaccurate statistics. For data missing case, 5% of the original input data is changed to 0 to simulate the situation of missing data.

In this case, we also conduct experiments 5 times, and the data are processed randomly before each experiment. As can be seen from Fig. 4, ELM still has a good prediction result in the case of data error, data missing and inaccurate.

5 Conclusion

Aiming at the application of safety supervision hidden danger prediction in education, medical treatment, restaurant, and other tertiary industries, this paper combines ELM algorithm to achieve application of hidden danger prediction.

Compared with the traditional BP learning algorithm for NN, the single hidden layer feedforward NN oriented ELM has the advantages of fast learning speed and high generalization performance [8, 11]. The experimental results verify the advantages of the method used in this paper. The method developed here is a novel practice of modern accident-causing theory. In the future, we will combine more effective ELM algorithm to further improve the prediction performance.

Acknowledgements This work was supported by the National Key Technologies R&D Program of China (Grant No. 2015BAK38B01).

References

1. Wei LJ, Fang LH (2014) Study on safety supervision law enforcement and hidden trouble inspection system based on mobile internet and cloud service. J Saf Sci Technol Supp 136–140 (in Chinese)
2. Shi H (2008) The risk early-warning of hidden danger in coal mine based on rough set-neural network. Proc Int Conf Risk Manag Eng Manag 314–317
3. Xiao HF, Tian YL (2011) Prediction of mine coal layer spontaneous combustion danger based on genetic algorithm and BP neural networks. Procedia Eng 26:139–146
4. Luo X, Zhang DD, Yang LT et al (2016) A kernel machine-based secure data sensing and fusion scheme in wireless sensor networks for the cyber-physical systems. Future Gener Comput Syst 61:85–96
5. Luo X, Liu J, Zhang DD et al (2016) A large-scale web QoS prediction scheme for the industrial internet of things based on a kernel machine learning algorithm. Comput Netw 101:81–89
6. Luo X, Luo H, Chang XH (2015) Online optimization of collaborative web service QoS prediction based on approximate dynamic programming. Int J Distrib Sens Netw 452492
7. Xu Y, Luo X, Wang WP et al (2017) Efficient DV-HOP localization for wireless cyber-physical social sensing system: a correntropy-based neural network learning scheme. Sensors 17(1):135
8. Huang GB, Zhu QY, Siew CK et al (2006) Extreme learning machine: theory and applications. Neurocomputing 70(1–3):489–501
9. He QX (2000) Safety engineering. China University of Mining and Technology Press, Xuzhou (in Chinese)
10. Jin Z (2011) Study on support vector machine classification algorithm and its application in safety evaluation of human accident in coal mine. China University of Mining and Technology Press, Xuzhou (in Chinese)
11. Luo X, Chang XH, Ban XJ (2016) Regression and classification using extreme learning machine based on L1-norm and L2-norm. Neurocomputing 174:179–186