Early warning modeling and analysis based on analytic hierarchy process integrated extreme learning machine (AHP-ELM): Application to food safety

ZhiQiang Genga, b, ShanShan Zhaoa, b, GuangCan Taaoc, YongMing Han a, b, *

a College of Information Science & Technology, Beijing University of Chemical Technology, Beijing 100029, China
b Engineering Research Center of Intelligent PSE, Ministry of Education in China, Beijing 100029, China
c Guizhou Academy of Testing and Analysis, Guiyang, Guizhou 550002, China

A R T I C L E   I N F O

Article history:
Received 27 October 2016
Received in revised form
25 January 2017
Accepted 19 February 2017
Available online 20 February 2017

Keywords:
Food safety
Extreme learning machine
Analytic hierarchy process
Artificial neural network
Early warning modeling

A B S T R A C T

Since the actual food safety monitoring data have characteristics of high-dimension, complexity, discreteness and nonlinear properties, it is difficult to accurately predict the risk of actual food inspection process. Therefore, this paper proposes a predictive modeling approach based on analytic hierarchy process (AHP) integrated extreme learning machine (ELM) (AHP-ELM). The proposed approach utilizes the AHP model to obtain the effective process characteristic information (PCIs). Compared with the analytic hierarchy process (AHP) integrated traditional artificial neural network (ANN) approach, the AHP-ELM prediction model is effectively verified by executing a linear comparison between all PCIs and the effective PCIs through daily inspection data source from the supervision and inspection department repository of China quality supervision system. Finally, the PCIs and the prediction value are obtained to provide more reliable food information and identification of potentially emerging food safety issues. The proposed method is applied to the food safety early warning and monitoring system in China. The result shows that the proposed model is effective and feasible in processing the complex food inspection data. Meanwhile, it can help to improve the quality of food products, ensure food safety and reduce the risk of food safety.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

In recent years, a series of food safety incidents not only endanger people's health and life safety directly (Xue & Zhang, 2013), but also undermine the reputation of the food processing industry in China (Lu & Wu, 2014; Pei et al., 2011). The frequent occurrences of major food safety incidents expose the imperfection and the low efficiency of Chinese food safety supervision system (Liu, Xie, Zhang, Cao, & Pei, 2013). The risk analysis-based food safety supervision system is commonly considered as the most effective approach to ensure food safety. Developed countries and regions, like Europe, the United States, Japan and South Korea have widely and commonly established food safety regulations or warning systems (Mensah & Julien, 2011; Neltner et al., 2011). The successful experience of developed countries is very useful for China to perfect our food safety regulation systems. In order to ensure food safety and protect public health, the Standing Committee of the 12th National People's Congress approved the Food Safety Law of the People's Republic of China on October 1, 2015. In the second chapter: the food safety risk monitoring and evaluation, the 17th article provides that the country builds the food safety risk assessment system. The system analyzes the risks of biological, chemical and physical hazards in food, food additives and food-related products based on the food safety risk monitoring data, relevant information and scientific method. The establishment of an effective food safety supervision and early warning system of potential food safety issues can avoid or reduce losses caused by the food safety hazards.

Currently, the early warning methods of food safety are mainly divided into the qualitative analysis method and the quantitative data prediction method. The qualitative analysis method is mainly used to draw warning conclusions from policy and theory analysis. This method is very common in food control research, analysis of...
the participants’ attitudes towards food safety risk, and analysis of the participants’ awareness of food safety issues (Clayton, Smith, Neff, Pollack, & Esmeninger, 2015; Manning & Soon, 2013). However, the data are mainly obtained from the questionnaire, rather than the laboratory testing, so they cannot be used to accurately build the early warning model (Chow & Mullan, 2010). The quantitative data prediction method is the early warning method based on data analysis. The researchers analyzed the laboratory testing data including the results of time series and regression analysis in the daily monitoring network to get the early warning results of the food safety (De Dominicis, Commissati, Griffi, Catellani, & Suman, 2015; Den Besten & Zwiertering, 2012; Qi, Zhong, & Liu, 2014). However, these statistical analysis-based methods need to know the structure of systems or the clear relationship between variables, which are often unrealistic in practice. Many regression analysis models deviate from reality, because the structure of the model itself is not clear (Granato, Calado, & Jarvis, 2014; Pan, Sun, Cheng, & Pu, 2016). The existing research method can make warning only if the current statistics or variations exceeded the standard or has network (RBF) (Wong., Seng., & Chen, 2011; Ding, Su, Pu, 2016). The existing research method can make warning only if the current statistics or variations exceeded the standard or has network (RBF) (Wong., Seng., & Chen, 2011; Ding, Su, Pu, 2016). The existing research method can make warning only if the current statistics or variations exceeded the standard or has network (RBF) (Wong., Seng., & Ang, 2011; Yu, Xie, Pasczynski, & Wilamowski, 2011) and so on. However, some problems are exposed in the traditional BP and RBF, such as the slow convergence usually requires thousands of iterations, the computational complexity increases rapidly when the network has many layers and nodes. In addition, the BP algorithm is very easy to fall into local optima so that the efficiency of the network training process cannot be guaranteed (Xie., Yu., & Wilamowski, 2011). In 2006, Huang et al. proposed a learning algorithm for a single hidden layer feed-forward neural network, called extreme learning machine (ELM) (Huang, Zhou, Ding, & Zhang, 2012, 2015). Under the premise of random selection of input layer weight, the algorithm calculates the output layer weight by using the Moore-Penrose generalized inverse matrix, and the generalization ability of the network is guaranteed. Compared with the BP and the RBF, the ELM has faster learning speed, less adjustable parameters and strong generalization ability (Wang, Cao, & Yuan, 2014; Huang, 2014; Pratama., Lu., & Zhang, 2015). Moreover, the ELM has been successfully used in speech recognition (Lu et al., 2016), face recognition (Peng, Wang, Long, & Lu, 2015), biological information processing (Li, Deng, He, & Du, 2010) and fault diagnosis (Yang, Wang, & Zhong, 2016).

However, since data produced by the inspection process have the characteristics of high-dimensional attributes, complexity, discreteness and nonlinear properties, the food safety inspection process model based on the ELM is inaccurate. To deal with the high dimensional data modeling problem, scholars explored many high dimensional data feature extraction methods, including principal component analysis (PCA) (Geng., Chen., & Han, 2016), input training neural network (Chen, Jiang, Li, Jia, & Ghamisi, 2016) and analytic hierarchy process (AHP). Among them, the AHP is a hierarchical and multi-criteria analysis method that combines qualitative analysis and quantitative analysis (Bernardon et al., 2011; Kou & Lin, 2014; Li, Anavatti, & Ray, 2014; Vigit, Isik, & Ince, 2014) to obtain the process characteristic information through data compression and data filtering. By combining the AHP and the ELM, this paper firstly extracted and preprocessed the raw data to fuse the classified reduction indicators. Then we got three kinds of early warning indicators as the output of the ELM: the risk of heavy metal pollution, the risk of chemical contaminant pollution and the risk of pathogenic bacteria pollution. Finally, we established a food safety warning model based on the AHP-ELM. The proposed method has been applied to the food safety early warning and monitoring system in one province in China. The result shows that the proposed model is effective and feasible in dealing with complex food inspection data. Meanwhile, it provides a new idea for the early warning of food safety and can help to improve the quality of food products, ensure food safety and reduce the risk of food safety.

2. Preprocessing of food inspection data

2.1. Data source and preprocessing

In this paper, from 2010 to 2014, a total of 18657 lines of the daily inspection data of food safety have been obtained from the Analysis and Testing Institute of one province in China. A part of the inspection data are shown in Table 1.

Table 1 shows the characteristics of food safety inspection data:

(1) High dimensional attribute. Each food inspection information consists of different attributes, such as manufacturers information, origin information, product information, production date, inspection items, inspection results and so on. The attributes are independent of each other and represent different dimensions, so we need a method to incorporate all of these items into a model or a system.

(2) Complexity. The description of inspection items are very complicated and their values contain many descriptive language. For example, because the inspection process includes sensory inspection, the table shows color is milky white as a descriptive result. These descriptive information are unquantifiable, leading to the increasing of analysis difficulty.

(3) Discreteness. The entire table has about 50 kinds of inspection indicators. The dimension of one indicator is different from the other. And the national standard requirements for each indicator are also different. This means that these indicators have different discrete domains, and the inspection results are disordered in the discrete domain. In addition, some inspection results are boundary descriptions and cannot determine their specific values, for example, the number of Escherichia coli does not exceed 100.

(4) Nonlinear properties. Each inspection information includes numerical information and description information. The numerical information contains a lot of null values and the description information contains some discrete values. Therefore, in a complete inspection information, the distribution of information between different attributes is asymmetrical and unbalanced. In other words, the emergence frequencies of two kinds of attributes are very different.

We have the following preliminary treatments of the raw data (Samplers, Toyofuku, Luning, Uyttendaele, & Jacxsen, 2012):

(1) Remove useless information. Each sample includes 31 descriptive fields. For analysts, it is necessary to simplify the information to emphasize the safety features. We remove food sensory quality items that are not closely related to food safety, such as the taste, the color, the shape and so on.

(2) Remove extra symbols. For example, the inspection result of total arsenic in a sample is “<0.011”. While the standard limit value of total arsenic is less than or equal to 0.25, we remove the “<” symbol in the result and record the inspection value as “0.011”.

### Table 1: Characteristics of Food Safety Inspection Data

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>High dimensional attribute</td>
<td>Each food inspection information consists of different attributes.</td>
</tr>
<tr>
<td>Complexity</td>
<td>The description of inspection items are very complicated.</td>
</tr>
<tr>
<td>Discreteness</td>
<td>The entire table has about 50 kinds of inspection indicators.</td>
</tr>
<tr>
<td>Nonlinear properties</td>
<td>Each inspection information includes numerical information and description information.</td>
</tr>
</tbody>
</table>
Table 1

<table>
<thead>
<tr>
<th>Bar code</th>
<th>Commodity name</th>
<th>Specifications</th>
<th>Date of production</th>
<th>Date of inspection</th>
<th>Inspection item name</th>
<th>Unit</th>
<th>Technical indicator</th>
<th>Inspection result</th>
<th>Individual evaluation</th>
<th>Inspection standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>6923644270940</td>
<td>Deluxe</td>
<td>organic milk</td>
<td>8/11/2013</td>
<td>00:00:00</td>
<td>color</td>
<td>–</td>
<td>Milky white</td>
<td>Qualified</td>
<td>GB 25190-2010</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>organic milk</td>
<td>8/11/2013</td>
<td>00:00:00</td>
<td>Chromium</td>
<td>mg/kg</td>
<td>≤0.3</td>
<td>0.021</td>
<td>GB/T 5009.123</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>organic milk</td>
<td>5/3/2013</td>
<td>00:00:00</td>
<td>protein</td>
<td>g/l</td>
<td>≥2.9</td>
<td>3.19</td>
<td>GB/T 5009.123</td>
<td></td>
</tr>
<tr>
<td>6917878020112</td>
<td>Nestle</td>
<td>milk</td>
<td>4/2/2014</td>
<td>00:00:00</td>
<td>iron</td>
<td>mg/g</td>
<td>5.00–10.01</td>
<td>6.34</td>
<td>GB/T 5009.123</td>
<td></td>
</tr>
<tr>
<td>6907992100272</td>
<td>Pure milk</td>
<td>900 g per tin</td>
<td>5/3/2013</td>
<td>00:00:00</td>
<td>Melamine</td>
<td>mg/kg</td>
<td>≤2.5</td>
<td>Negative</td>
<td>GB/T 5009.123</td>
<td></td>
</tr>
<tr>
<td>6907992507064</td>
<td>Low lactose</td>
<td>250 ml per pack</td>
<td>11/21/2013</td>
<td>00:00:00</td>
<td>AFT M1</td>
<td>µg/kg</td>
<td>≤0.5</td>
<td>Negative</td>
<td>GB/T 5009.123</td>
<td></td>
</tr>
</tbody>
</table>

(3) Use a minimum value instead of zero. For example, the national standard of China requires to prohibit the addition of melamine in dairy products and the actual inspection result of the sample was not detected. In order to ensure the generated information matrix is a positive definite matrix, we use a very small positive number 0.001 instead of zero.

(4) If the inspection result is positive and does not contain special characters, we add the value to the information matrix directly. For example, the inspection result of vitamin D is 9.70, so we add 9.70 to the information matrix.

A part of pre-processed inspection data are shown in Table 2:

2.2. Classification of inspection indicators

There are many kinds of food inspection indicators (Jacxsens et al., 2010), including pesticides, veterinary drug, biological toxins, food additives, non-food additives (Huang, Min, Duan, Wu, & Li, 2014), prohibited ingredients, persistent organic pollutants, processed products, pathogenic bacteria (Gentschew & Ferguson, 2012; National food safety standards of China, 2013) and so on. For existing data table, there are about 50 kinds of inspection indicators. In this paper, we selected 11 kinds of important inspection indicators and then categorized them (Van Boekel et al., 2010; Dong et al., 2015). This 11 kinds of indicators can directly determine whether the product is safe (National food safety standards of China, 2012; National food safety standards of China, 2010) Table 3.

3. The AHP-ELM method

3.1. Analytic hierarchy process (AHP)

The traditional AHP model constructs the judgment matrix by establishing a hierarchical structure model and combining with expertise. Firstly, the process is affected by people’s subjective consciousness, thus it is prone to errors in judgment. The rough comparison and calculation process of the traditional AHP method does not meet the requirements of high accuracy in the actual inspection process. Secondly, the data of food inspection process are complex and the data items are various, so it is not realistic for people to build up a precise mathematical model. Using the data driven technology, we only need to capture the input and output signals without knowing the internal principle of the production process. It’s obviously impossible to predict all items one by one as there are too many inspection indicators and results are complicated. Therefore, this paper uses the AHP algorithm based on correlation function differential driver to fuse different indicators in the inspection sample, and ultimately constructs the output of food safety inspection data.

Suppose the matrix $Y$ is the initial sample data and its attributes:

**Definition 1.** Suppose the $j$ parameters of the upside correlation function (Han, Geng, & Liu, 2014) of the sample is:

\[
c_{ij}(y) = \begin{cases} 
0 & y \notin \{y_{i1}, y_{i2}, \ldots, y_{in}\} \\
\frac{y_{ij} - y_{ij}(1)}{y_{ij}(2) - y_{ij}(1)} & y \in \{y_{i1}, y_{i2}, \ldots, y_{in}\} \\
1 & y \notin \{y_{i1}, y_{i2}, \ldots, y_{in}\} \\
\frac{y_{ij}(4) - y_{ij}}{y_{ij}(3) - y_{ij}(4)} & y \notin \{y_{i1}, y_{i2}, \ldots, y_{in}\}
\end{cases}
\]

(1)

where $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$.

The preprocessed information matrix is as follows:

\[
C_{n \times m} = \begin{bmatrix} 
C_{11} & C_{12} & \cdots & C_{1m} \\
C_{21} & C_{22} & \cdots & C_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
C_{n1} & C_{n2} & \cdots & C_{nm}
\end{bmatrix}
\]

(2)

Center normalization:

\[
\begin{align*}
C_{ij} &= \left(C_{ij} - \bar{C}_j\right) / S_j \\
\bar{C}_j &= \frac{1}{n} \sum_{i=1}^{n} C_{ij} \\
S_j &= \left(\frac{1}{n} \sum_{i=1}^{n} (C_{ij} - \bar{C}_j)^2\right)^{0.5}
\end{align*}
\]

(3)

where $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$, then move the negative to zero (the zero is positive, represented by zero plus a positive decimal $\varepsilon$)

\[
d_{ij} = c_{ij} - t_j + \varepsilon, i = 1, 2, \ldots, n, j = 1, 2, \ldots, m.
\]

where $t_j = \min(a_{ij}) < 0, j = 1, 2, \ldots, m$. And get the positive matrix $D_{n \times m}$:

\[
D_{n \times m} = \begin{bmatrix} 
d_{11} & d_{12} & \cdots & d_{1m} \\
d_{21} & d_{22} & \cdots & d_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
d_{n1} & d_{n2} & \cdots & d_{nm}
\end{bmatrix}
\]

(4)

Use $D_{n \times m}$ to get the n-dimensional matrix:

\[
COR = DD^T = \begin{bmatrix} 
0_{11} & 0_{12} & \cdots & 0_{1n} \\
0_{21} & 0_{22} & \cdots & 0_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
0_{n1} & 0_{n2} & \cdots & 0_{nn}
\end{bmatrix}
\]

(5)
Table 2
Food inspection pre-processed data.

<table>
<thead>
<tr>
<th>id</th>
<th>coliform group</th>
<th>total numbers of colony</th>
<th>Escherichia coli</th>
<th>lead</th>
<th>copper</th>
<th>arsenic</th>
<th>acephate</th>
<th>hexachlorocyclohexane</th>
<th>fenitrothion</th>
<th>cypermethrin</th>
<th>dichlorodiphenyltrichloroethane</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.001</td>
<td>47</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>47</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.001</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>15</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.001</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0.001</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>0.001</td>
<td>34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3
Initial inspect indicators and its category.

<table>
<thead>
<tr>
<th>ID</th>
<th>inspect indicators</th>
<th>category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>coliform group</td>
<td>pathogenic bacteria</td>
</tr>
<tr>
<td>2</td>
<td>total numbers of colony</td>
<td>pathogenic bacteria</td>
</tr>
<tr>
<td>3</td>
<td>Escherichia coli</td>
<td>pathogenic bacteria</td>
</tr>
<tr>
<td>4</td>
<td>lead</td>
<td>heavy metal</td>
</tr>
<tr>
<td>5</td>
<td>copper</td>
<td>heavy metal</td>
</tr>
<tr>
<td>6</td>
<td>arsenic</td>
<td>heavy metal</td>
</tr>
<tr>
<td>7</td>
<td>acephate</td>
<td>chemical contaminant</td>
</tr>
<tr>
<td>8</td>
<td>hexachlorocyclohexane</td>
<td>chemical contaminant</td>
</tr>
<tr>
<td>9</td>
<td>fenitrothion</td>
<td>chemical contaminant</td>
</tr>
<tr>
<td>10</td>
<td>cypermethrin</td>
<td>chemical contaminant</td>
</tr>
<tr>
<td>11</td>
<td>dichlorodiphenyltrichloroethane</td>
<td>chemical contaminant</td>
</tr>
</tbody>
</table>

The AHP can filter the redundant information in the high-dimensional input, extract the feature of components and effectively remove the noise.

3.2. Extreme learning machine

The ELM neural network is composed of input layer, hidden layer and output layer, and the neurons of the two adjacent layers are all connected. The weights between the input layer and the hidden layer can be randomly generated, and the weights between the hidden layer and the output layer are obtained by using the least square method. Because the parameters are not adjusted, so the ELM has an extremely fast training speed. The network structure is shown in Fig. 1:

Suppose there are p arbitrary training samples, which the set is represented as: \( P = \{(x_p, y_p) | p = 1, 2, \ldots, p; x_p \in \mathbb{R}^P; y_p \in \mathbb{R}^W\} \)

The ELM network model with L hidden layer nodes is defined as follows:

\[
y_p = \Phi(x_p) = \sum_{l=1}^{L} \beta_l g(y_l, x_p + b_l) \tag{8}
\]

where \( \omega_l = [\omega_{l1}, \omega_{l2}, \ldots, \omega_{ln}]^T \) are the weight vector between the \( lth \) input layer node and the \( lth \) hidden layer node, \( b_l \) is the threshold of the \( lth \) hidden layer neuron, \( \beta_l = [\beta_{l1}, \beta_{l2}, \ldots, \beta_{lW}]^T \) are the weight connecting the \( lth \) hidden layer neurons to the \( lth \) output layer neurons, \( y_l, x_p \) is the inner product between \( \omega_l \) and \( x_p \), \( g(\cdot) \) is a linear activation function.

In the ELM theory, \( \beta_l, \omega_l, b_l \) are existed for the \( lth \) hidden layer neurons. And a activation function has been defined, which makes the single hidden layer feed forward neural network can approximate the sample expectation of \( y_p \) with the error of zero. Therefore, the Eq. (8) can be rewritten as the following equation:

\[
G \beta = y \tag{9}
\]

\[
G = \begin{bmatrix}
g(\omega_1 \cdot x_1 + b_1) & \ldots & g(\omega_L \cdot x_1 + b_L) \\
g(\omega_1 \cdot x_p + b_1) & \ldots & g(\omega_L \cdot x_p + b_L) 
\end{bmatrix}_{p \times L} \tag{10}
\]

\[
\beta = \begin{bmatrix}
\beta_1^T \\
\beta_2^T \\
\vdots \\
\beta_L^T
\end{bmatrix}_{1 \times L} \tag{11}
\]

\[
y = \begin{bmatrix}
y_1^T \\
y_2^T \\
\vdots \\
y_p^T
\end{bmatrix}_{1 \times L} \tag{12}
\]

where \( G \) is the output matrix of hidden layer, \( \beta \) is the vector of weights between the hidden layer neurons and the output layer neurons, and \( y \) is the outputs vector of the ELM model. The output weights \( \beta \) can be calculated by the least square solution of minimal 2- norm of Eq. (9):
\[ \bar{y} = G^+ y \]  

(13)

where \( G^+ \) is the Moore-Penrose generalized inverse of \( G \). Therefore, if known the set of learning samples, the number of hidden layer neurons and the activation function of hidden layer, the ELM algorithm steps are:

1. Randomly generate the vector of weights \( \omega \) between the input layer neurons and the hidden layer neurons and the threshold \( b \) of the hidden neuron;
2. Calculate the output matrix \( G \) of the hidden layer;
3. Calculate the vector of weights \( \bar{\beta} \) between the hidden layer neurons and the output layer neurons by Eq. (13).

3.3. The framework of early warning model of food safety

Suppose there are \( K \) training samples with high dimensional input, which the set is represented as \( S = \{ (x_{ij}) | i = 1, 2, \ldots, K; j = 1, 2, \ldots, J \} \), the input of each training sample contains \( J \) attributes. For the training samples \( S \), the modeling process based on AHP-ELM is as follows:

Step 1: Construct the output. Using the method of inspection indicators classification in Section 2.2 to divide \( J \) input attributes of the set \( S \), record the classification group number \( M \), the attributes number of each group and the specific input attributes. The \( J \) group attributes are used as the input vectors of the AHP model. After feature extraction and fusion, a new set of samples \( S' = \{ (y_{it}) | i = 1, 2, \ldots, K; t = 1, 2, \ldots, M \} \) are obtained, where \( M \) group attributes are the output vectors of the AHP model.

Step 2: The establishment of ELM neural network. The training set \( S \) as the input of ELM neural network, the sample set \( S' \) obtained by step(1) as the expected output of the ELM neural network, so the network has \( J \) neurons in the input layer and \( M \) neurons in the output layer. The contribution of each feature is different due to the different description and the non-uniform dimension. In order to reduce the difficulty of training, each variable should provide the same importance and would vary in not too wide range. So we normalize the data set before training by using the following conversion formula:

\[
\begin{align*}
    x_{ij} & = \frac{x_{ij}^{max} - x_{ij}}{x_{ij}^{max} - x_{ij}^{min}} \\
    y_{it} & = \frac{y_{it}^{max} - y_{it}}{y_{it}^{max} - y_{it}^{min}}
\end{align*}
\]  

(14)

If \( x_{ij}^{max} = x_{ij}^{min} \), \( x_{ij} = -1 \), where in Eq. (14) \( x_{ij}^{max} = \max(x_{ij}, x_{2j}, \ldots, x_{Kj}) \), \( x_{ij}^{min} = \min(x_{ij}, x_{2j}, \ldots, x_{Kj}) \), \( i = 1, 2, \ldots, K; j = 1, 2, \ldots, J \). If \( y_{it}^{max} = y_{it}^{min} \), \( y_{it} = -1 \), where in Eq. (14) \( y_{it}^{max} = \max(y_{it}, y_{2t}, \ldots, y_{Kt}) \), \( y_{it}^{min} = \min(y_{it}, y_{2t}, \ldots, y_{Kt}) \), \( i = 1, 2, \ldots, K, t = 1, 2, \ldots, M \).

Step 3: Training process. Using the ELM neural network to train the normalized training sample set \( (S, S') \) = \( \{ (x_{ij}, y_{it}) | i = 1, 2, \ldots, K; j = 1, 2, \ldots, J; t = 1, 2, \ldots, M \} \), then record the vector of weights \( \omega \) between the input layer neurons and the hidden layer neurons, the threshold \( b \) of the hidden neuron and the vector of weights \( \bar{\beta} \) between the hidden layer neurons and the output layer neurons.

Step 4: Recall process. The AHP-ELM model has been obtained by step(3) training. We use the set \( S \) to calculate the recall output \( D \) = \( \{ y_{it}' | i = 1, 2, \ldots, K; t = 1, 2, \ldots, M \} \) of the model, then use Eq. (15) to anti-normalize the recall output:

\[
y_{it}' = y_{it}^{max} - y_{it}' \left( y_{it}^{max} - y_{it}^{min} \right)
\]  

(15)

where \( y_{it}^{max} = \max(y_{it}', y_{2t}', \ldots, y_{Kt}') \), \( y_{it}^{min} = \min(y_{it}', y_{2t}', \ldots, y_{Kt}') \), \( i = 1, 2, \ldots, K, t = 1, 2, \ldots, M \). We compare the recall output and the expected output, then calculate the recall relative error and the recall standard deviation.

Step 5: Generalization process. Using a set of data sample \( (S, S') = \{ (x_{ij}', y_{it}') | i = 1, 2, \ldots, K; j = 1, 2, \ldots, J; t = 1, 2, \ldots, M \} \) which is different from the training samples as a generalization sample of
the AHP-ELM network model. We compare the generalization output and the expected output $\hat{S}$, then calculate the generalization relative error and the generalization standard deviation.

The framework of early warning model of food safety is shown in Fig. 2.

Fig. 2. The framework of the AHP–ELM algorithm.

4. Early warning analysis of the food safety

We select 150 inspection data of the dairy products in 2012 to establish model from the Analysis and Testing Institute of one province in China. Firstly, we preprocess these raw data, and then select 11 inspection indicators: coliform group, total numbers of colony, Escherichia coli, lead, copper, arsenic, acephate,
hexachlorocyclohexane, fenitrothion, cypermethrin and dichlorodiphenyltrichloroethane. Above-mentioned indicators which can directly affect the safety of dairy products are selected as the input items of neural network. Then, we use the AHP model to extract the features of the input parameters and obtain pathogenic bacteria, heavy metals and chemical contaminants as the output items of neural network. A part of AHP fusion results are shown in Table 4.

In this paper, the 11 inspection indicators can be selected as the input variables of the neural network. And we set the AHP fusion results as the output variables of the network which refer to pathogenic bacteria, heavy metal and chemical contaminant. We randomly select 135 groups data as the training data, and the remaining 15 groups data as the generalization data. At the same time, we assign the initial hidden layer node as 20 by using the gradual increasing method, in order to shorten the training time (Gao, 1998; Han, Geng, & Zhu, 2016). So, the nodes number of input layer, hidden layer and output layer of ELM network are 11, 20 and 3.

In order to validate the effectiveness and robustness of the AHP-ELM model, we build the BP and RBF network model to compare and verify the neural network proposed in this paper. In the single-layer BP network, we set the learning factor as 0.1, the momentum factor as 0.9, the iteration number as 1000, the excitation function as sigmoid function, the nodes number of input layer, hidden layer and output layer are 11, 20, 3. In the RBF network, we set learning factor as 0.1, the iteration number as 1000, the nodes number of input layer, hidden layer and output layer are 11, 20, 3. In addition, the average relative generalization error (ARGE), root mean square error (RMSE) and generalization standard deviation (Std) are obtained by Eq. (16), Eq. (17) and Eq. (18), respectively.

Table 4
Results of the AHP fusion.

<table>
<thead>
<tr>
<th>id</th>
<th>Pathogenic bacteria</th>
<th>Heavy metal</th>
<th>Chemical contaminants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.9267</td>
<td>0.8614</td>
<td>1.8216</td>
</tr>
<tr>
<td>2</td>
<td>1.6467</td>
<td>1.7705</td>
<td>2.9061</td>
</tr>
<tr>
<td>3</td>
<td>1.2733</td>
<td>1.2356</td>
<td>1.9236</td>
</tr>
<tr>
<td>4</td>
<td>1.9733</td>
<td>1.0603</td>
<td>2.4848</td>
</tr>
<tr>
<td>5</td>
<td>2.5867</td>
<td>2.0144</td>
<td>2.8285</td>
</tr>
<tr>
<td>6</td>
<td>2.6067</td>
<td>1.5584</td>
<td>0.7871</td>
</tr>
<tr>
<td>7</td>
<td>1.9533</td>
<td>1.5276</td>
<td>0.7285</td>
</tr>
<tr>
<td>8</td>
<td>1.8933</td>
<td>1.9433</td>
<td>2.0106</td>
</tr>
<tr>
<td>9</td>
<td>3.0000</td>
<td>1.6423</td>
<td>2.9610</td>
</tr>
<tr>
<td>10</td>
<td>0.3333</td>
<td>1.4198</td>
<td>3.7891</td>
</tr>
<tr>
<td>11</td>
<td>1.5200</td>
<td>0.2873</td>
<td>3.2027</td>
</tr>
<tr>
<td>12</td>
<td>1.2000</td>
<td>1.7282</td>
<td>2.6069</td>
</tr>
<tr>
<td>13</td>
<td>1.9733</td>
<td>1.0791</td>
<td>2.1619</td>
</tr>
<tr>
<td>14</td>
<td>0.6333</td>
<td>0.3110</td>
<td>2.6516</td>
</tr>
<tr>
<td>15</td>
<td>2.0400</td>
<td>1.0388</td>
<td>2.9851</td>
</tr>
<tr>
<td>16</td>
<td>1.5600</td>
<td>2.0094</td>
<td>2.2373</td>
</tr>
<tr>
<td>17</td>
<td>1.4667</td>
<td>0.6462</td>
<td>2.2478</td>
</tr>
<tr>
<td>18</td>
<td>0.7733</td>
<td>1.1401</td>
<td>2.5283</td>
</tr>
<tr>
<td>19</td>
<td>1.3467</td>
<td>1.3603</td>
<td>1.9686</td>
</tr>
<tr>
<td>20</td>
<td>2.5467</td>
<td>1.6601</td>
<td>3.5147</td>
</tr>
</tbody>
</table>

Fig. 3. Comparison of different generalization results and actual results.

Table 5
Comparison of food inspection data set.

<table>
<thead>
<tr>
<th>Item</th>
<th>BP</th>
<th>RBF</th>
<th>RBF</th>
<th>ELM</th>
<th>ELM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hidden nodes number</td>
<td>20</td>
<td>20</td>
<td>30</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>Running time/ms</td>
<td>4526</td>
<td>3275</td>
<td>4665</td>
<td>386</td>
<td>468</td>
</tr>
<tr>
<td>Average relative generalization error/%</td>
<td>11.746</td>
<td>17.936</td>
<td>7.879</td>
<td>4.962</td>
<td>0.528</td>
</tr>
<tr>
<td>Root mean square error</td>
<td>0.2239</td>
<td>0.2861</td>
<td>0.1024</td>
<td>0.0719</td>
<td>0.0071</td>
</tr>
<tr>
<td>Generalization standard deviation</td>
<td>0.1270</td>
<td>0.2495</td>
<td>0.1265</td>
<td>0.0475</td>
<td>0.0041</td>
</tr>
</tbody>
</table>
where \( P_{out_{ji}} \) is the predictive output and \( E_{out_{ji}} \) is the expected output, \( i = 1, 2, \ldots, n; j = 1, 2, \ldots, k \).

The results of pathogenic bacteria actual value (PA), heavy metal actual value (HA), chemical contaminants actual value (CA), pathogenic bacteria predictive value of ELM (PE), heavy metal predictive value of ELM (HE), chemical contaminants predictive value of ELM (CE), pathogenic bacteria predictive value of BP (PB), heavy metal predictive value of BP (HB), chemical contaminants predictive value of BP (CB), pathogenic bacteria predictive value of RBF (PR), heavy metal predictive value of RBF (HR), chemical contaminants predictive value of RBF (CR) are shown in Fig. 3, and the prediction accuracy and running time of each model are shown in Table 5.

As we can see from Fig. 3, the generalization results of ELM are more close to the actual value than BP and RBF. It can be seen from Table 5 that both ARGE and RMSE of the AHP-ELM model are better than that of AHP-BP and AHP-RBF. When the number of hidden layer nodes is 20, the ARGE of ELM is 0.528%, however, the ARGE of BP and RBF are 11.746% and 17.936% respectively. Secondly, the standard deviation of the ELM is smaller than the other two networks. When the number of hidden layer nodes is 20, the standard deviation of ELM is 0.0041, however, the standard deviation of BP and RBF are 0.1270 and 0.2495 respectively. Thirdly, in terms of the time required for the entire training and generalization process, ELM network is significantly faster, it is about 11 times faster than the BP network, and about 8 times faster than the RBF network.

Therefore, when we use the ELM to make early warning analysis of food safety, we can predict more quickly, more stable and with less error.

We analyze 15 inspection data of the dairy products in 2012. Fig. 4 shows the comparison of the actual risks of these dairy products and the risks predicted by AHP-ELM model. Table 6 are the risks data processed by AHP-ELM model and expected results.

We also select 15 inspection data of the dairy products from Analysis and Testing Institute of one province in 2014 to make early warning analysis of food safety. After the method in Section 2.1 preprocessed, then we use the AHP-ELM model to predict the 2014 dairy products safety risks. Compared with the 2012 dairy products safety risks in Table 4, the results are shown in Fig. 5:

### Table 6
Comparison of actual value and predictive value.

<table>
<thead>
<tr>
<th>ID</th>
<th>PA</th>
<th>PE</th>
<th>HA</th>
<th>HE</th>
<th>CA</th>
<th>CE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0715</td>
<td>1.0699</td>
<td>0.1193</td>
<td>0.1193</td>
<td>0.0938</td>
<td>0.0938</td>
</tr>
<tr>
<td>2</td>
<td>2.5996</td>
<td>2.5997</td>
<td>0.1512</td>
<td>0.1515</td>
<td>0.3155</td>
<td>0.3141</td>
</tr>
<tr>
<td>3</td>
<td>3.0799</td>
<td>3.0824</td>
<td>0.1794</td>
<td>0.1787</td>
<td>0.1575</td>
<td>0.1608</td>
</tr>
<tr>
<td>4</td>
<td>1.4650</td>
<td>1.4733</td>
<td>0.0274</td>
<td>0.0274</td>
<td>0.3245</td>
<td>0.3226</td>
</tr>
<tr>
<td>5</td>
<td>1.5195</td>
<td>1.5186</td>
<td>0.1690</td>
<td>0.1704</td>
<td>0.2865</td>
<td>0.2869</td>
</tr>
<tr>
<td>6</td>
<td>1.0050</td>
<td>0.9963</td>
<td>0.1303</td>
<td>0.1281</td>
<td>0.1800</td>
<td>0.1824</td>
</tr>
<tr>
<td>7</td>
<td>0.5348</td>
<td>0.5309</td>
<td>0.1442</td>
<td>0.1437</td>
<td>0.1259</td>
<td>0.1268</td>
</tr>
<tr>
<td>8</td>
<td>1.1209</td>
<td>1.1368</td>
<td>0.1245</td>
<td>0.1237</td>
<td>0.3264</td>
<td>0.3262</td>
</tr>
<tr>
<td>9</td>
<td>2.9966</td>
<td>2.9807</td>
<td>0.1306</td>
<td>0.1295</td>
<td>0.1445</td>
<td>0.1431</td>
</tr>
<tr>
<td>10</td>
<td>1.5533</td>
<td>1.5662</td>
<td>0.1309</td>
<td>0.1328</td>
<td>0.0889</td>
<td>0.0890</td>
</tr>
<tr>
<td>11</td>
<td>1.5844</td>
<td>1.5791</td>
<td>0.0932</td>
<td>0.0914</td>
<td>0.2988</td>
<td>0.2994</td>
</tr>
<tr>
<td>12</td>
<td>3.0612</td>
<td>3.0700</td>
<td>0.0954</td>
<td>0.0965</td>
<td>0.2328</td>
<td>0.2338</td>
</tr>
<tr>
<td>13</td>
<td>0.5348</td>
<td>0.5309</td>
<td>0.1442</td>
<td>0.1437</td>
<td>0.1259</td>
<td>0.1268</td>
</tr>
<tr>
<td>14</td>
<td>1.9880</td>
<td>1.9972</td>
<td>0.1624</td>
<td>0.1632</td>
<td>0.1465</td>
<td>0.1458</td>
</tr>
<tr>
<td>15</td>
<td>0.5467</td>
<td>0.5468</td>
<td>0.1102</td>
<td>0.1110</td>
<td>0.2241</td>
<td>0.2238</td>
</tr>
</tbody>
</table>

Fig. 4. Generalization curve of AHP-ELM.
one province in 2014, including pathogenic bacteria, heavy metals and chemical contaminants. For example, it can be seen from Fig. 5 that the pathogenic bacteria risk values of the fifth sample in 2012 and 2014 were 1.94 and 1.57, respectively, which indicate that the pathogenic bacteria risk of this product in 2014 was reduced. The heavy metal risk values of the fifth sample in 2012 and 2014 were 1.74 and 2.05, respectively, which indicate that the heavy metal risk of this product in 2014 was increased. The chemical contaminants risk values of the fifth sample in 2012 and 2014 were 2.76 and 2.26, respectively, which indicate that the chemical contaminants risk of this product in 2014 was reduced. Therefore, we should pay attention to the heavy metal pollution of this product.

For all dairy inspection samples of one province in 2014, the safety risks were caused by pathogenic bacteria, heavy metals, chemical pollutants, and their average values were 0.7969, 0.6842, 1.5698. Compared with the dairy products safety risks of this province in 2012, the risks increased in 2014, so we analyzed that there are hidden dangers of food safety issues. Among the three dimensions of pathogenic bacteria, heavy metals and chemical contaminants, the highest degree of risk are the chemical contaminants, so the possibility of chemical pollution of dairy products in the future is also the highest, it should cause the attention of the relevant departments.

5. Discussion

First, the food safety modeling and application approach based on the AHP-ELM was proposed. Compared with the traditional quantitative data analysis approach, the robustness and effectiveness of AHP-ELM model were validated through the dairy product inspection data source from the Analysis and Testing Institute of one province.

Second, the AHP method could help to extract data features and reduce data dimensions, which can make the quantitative analysis becomes simple. And by using the ELM network based on the fused data, the input-output model of the food safety risks could be established. The model would serve as an operation guide for the food safety risks discovery and early warning, it meets the needs of the national food safety policy for early warning. And it could also be used for predicting the potential risk value related to the input inspection data. This proposed modeling method could solve problems like, when using the AHP to evaluate the food safety risks, the data samples have too many input indicators that cannot directly analysis, and when using the ELM to predict the food safety risks, the potential risk of food can be predicted quickly without using AHP to fuse a lot of inspection data by category.

Third, the current data pre-processing method is mostly based on expert opinion, therefore, we will deal with more data for more objective, such as, using some data pre-processing methods like the principal component analysis(PCA).

Finally, the proposed modeling method is applied to analyze the inspection data and predict the risks of food safety effectively, and the parameters of model are adjusted by training samples. Therefore, we will improve our model so that parameters can be adjusted automatically, such as designing a self-organizing AHP-ANN model, which has more practical value and is more suitable for practical application.

6. Conclusions

This paper proposes an AHP-ELM approach that could help to build a multi-input-multi-output optimization and prediction model of food safety inspection data. This proposed algorithm analyzes the food safety risks based on the inspection data that have characteristics of high-dimension, complexity, discreteness and nonlinear properties. In order to reduce the difficulty of processing and forecasting data, we use the AHP method to filter out redundant information and extract the characteristic variables of data as the output of the ELM. The AHP-ELM approach solves the lag problem of the traditional quantitative data analysis. Moreover, in terms of accuracy and training time, through the inspection data modeling of dairy products in 2012, we verify that the robustness and effectiveness of the AHP-ELM model are better than the AHP-BP model and AHP-RBF model. Finally, the AHP-ELM approach is applied in predicting the safety risks of dairy products in 2014. The results show that the proposed method could obtain the risks of...
pathogenic bacteria, heavy metal and chemical contaminants. Meanwhile, the food inspection departments could focus on items with higher risk values by comparing the risk values of previous years. And the food inspection departments could utilize these results to effectively improve the early warning efficiency of food safety.

For further research, we will study the principal component analysis (PCA) method to optimize the data pre-processing process. We need to reduce the data dimension and maximum retain the useful information of the original variables. Moreover, the particle swarm optimization (PSO) will be used to analyze and predict the food safety risks to compare with current studies. Furthermore, we will use some self-organizing neural networks for the food safety modeling.

Acknowledgments

This research was partly funded by National Natural Science Foundation of China (61763046, 61603025), the Natural Science Foundation of Beijing, China (4162045) and the Technology Support Foundation of China (61673046, 61603025), the Natural Science Foundation of Beijing, China (4162045) and the Technology Support Foundation of Beijing, China (4162045).

References


