Acta Technica 62 (2017), No. 7A, 1297–1308

Higher education evaluation based on bayesian self organizing neural network in big data environment

Yanqiu Liang[1]

Abstract. In order to improve effectiveness of higher education evaluation in big data environment, a higher education evaluation method based on Bayesian self-organizing neural network is presented in this paper. Firstly, a higher education evaluation index system is established. Then Bayesian self-organizing neural network classifier is able to be established in combination with index system and example data. And higher education evaluation is performed on the basis of Bayesian self-organizing neural network classifier. Next, for knowledge, which is improved on the basis that conventional Bayesian classification algorithm is insufficient and where mutual information is applied, characteristics are selected by means of relative credibility so as to delete redundant attributes in a way that obtains improved Bayesian algorithm and improves algorithm performance. In the end, effectiveness of the algorithm presented is verified by simulation experiment.

Key words. Big Data, Bayes, Self-organizing neural network, Higher education.

1. Introduction

Higher education is a bilateral activity between teachers and students to achieve certain higher education goals through information transmission, process control and strategy implementation. With a long history, higher education will remain a major mode of higher education for a long period of time, and also a major way for students to establish knowledge structure and develop cognitive structure and personality formation. Higher education quality evaluation can provide targeted information for the development of more scientific higher education strategies, and help promote the reform of higher education and improve the quality of higher education. The three-level index system is used for higher education evaluation. On the assumption of linear relation between the indexes, the second-level indexes are

[1]Normal College, Nanyang Institute of Technology, Nanyang City Henan Province, 473000, China

calculated on the basis of the third-level indexes, and level of first-level indexes is determined by means of the second-level indexes. Such level judgment process is a matter of hierarchical classification (technology which simulates human's concept learning and application). And judgment for higher education quality level based on classifier does not need assumption of linear relation. Thus, it has advantages in terms of reliability of evaluation and can explore a new idea of higher education evaluation.

There are already many well-known classifiers, such as neural networks, support vector machines, decision trees, and statistical discriminatory analysis and Bayesian networks. They are widely applied to numerous fields. However, such classifiers often require massive example data for learning. Example data on higher education are less generally. And continuous data are not suitable for discretization (discretization will make massive information lost). Naive Bayesian network (NBN for short) classifier, a probability classifier most suitable for classification and prediction of small example set, does not require massive example data for training. And it is able to directly process continuous attributes. The key issue to process continuous attributes in NBN classifier is conditional density estimation. At present, two major methods are used to estimate conditional density. One is to estimate attribute conditional density by means of Gaussian function. There may be a large difference between Gaussian function and practical density function. As a result, classification accurateness of classifier is impacted. The other one is to estimate attribute conditional density by means of Gaussian kernel function, which is apt to lead to over-fitting of examples. Thus, generalization capacity of classifier is also reduced.

In this paper, index system of higher education quality evaluation is developed from three perspectives, information transmission, process control and strategy of higher education. Classifier model of mint-hierarchical naïve Bayesian network (MHNBN for short) is established in accordance with practical conditions and demand of higher education quality evaluation. In order to avoid over-fitting of the examples which may be caused by attribute condition density estimation by Gaussian kernel function, the shape parameters are introduced into the Gaussian kernel function, and the classification and identification accuracy of the classifier is improved through optimization of the shape parameters.

2. Problem description

Higher education evaluation index system is established. Then MHNBN classifier is able to be established in combination with index system and example data. And higher education evaluation is performed on the basis of MHNBN classifier.

2.1. Higher education evaluation index system

The index system is the prerequisite for higher education evaluation. A three-level index system of higher education evaluation is established in accordance with educational cybernetics, systematic science principles, higher education mechanism and so on. The index system may be expanded hierarchically in accordance with practical demand.

(1) First-level index. Higher education (C) is divided into 4 grades, grade A (excellent), grade B (good), grade C (average) and grade D (bad).

(2) Second-level index. Second-level indexes subordinate to higher education include classroom information transmission (X_1), higher education control (X_2) and higher education strategy (X_3). They are all divided into three grades, grade A (good), grade B (average) and grade C (bad).

(3) Third-level index. Third-level indexes subordinate to classroom information transmission include information transmission from teachers to students (X_{11}), semantic information transmission (X_{12}), pragmatic information transmission (X_{13}), information transmission from students to teachers, feed-forward information (X_{15}) and feed-backward information (X_{16}).

Third-level indexes subordinate to higher education control include knowledge structure control (concept (X_{21}), rule (X_{22}) and problem solving (X_{23})), cognitive structure control (cognitive operation (X_{24}), impetus supply (X_{25}), cognition strategy (X_{26}), method control (process control (X_{27}) and random control (X_{27}).

Third-level indexes subordinate to higher education strategy include teaching-based higher education (X_{31}), enlightenment-based higher education (X_{32}), deduction-based higher education (X_{33}), summarization-based higher education (X_{34}), and backtracking-based higher education (X_{35}).

2.2. Classifier mode of higher education evaluation

According to the higher education evaluation index system above, two levels of MHNBN classifier structure can be obtained, as shown in Fig. 1.

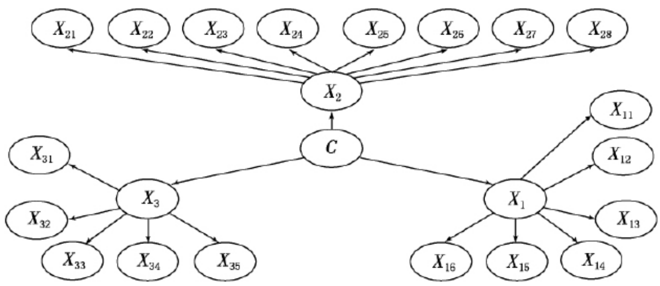


Fig. 1. MHNBN classifier structure used for higher education evaluation

Parameters are estimated on the basis of classifier structure and example data in a way that obtains MHNBN classifier used for higher education evaluation and inputs the latest information about higher education. Level of higher education can be obtained through classification operation.

3. Bayesian Higher Education Evaluation Based on Credibility of Mutual Information

3.1. Naive Bayesian classification

In NB classification, probability of type subordinate to sample is determined through posterior probability calculation. The basic idea is to estimate type of new sample by means of combined probability of attribute and type and on the basis of Bayesian equation and simplified assumption of probability theory.

Bayesian equation: Assume $A_1, A_2, \cdots, A_n$ are a group of mutually incompatible events, when event B can only occur with one of such events at a time, equation (1) is workable, namely

$$P(Ai|B) = \frac{P(B|Ai)P(Ai)}{P(B)} = \frac{P(B|Ai)P(Ai)}{\sum\limits_{i=1}^{n} P(B|Ai)P(Ai)}\,. \tag{1}$$

Where

$$P(Ci|X) = \frac{P(X|Ci)P(Ci)}{P(X)}\,. \tag{2}$$

$$P(X|Ci) = \prod_{j=1}^{n} P(xj|Ci)\,. \tag{3}$$

$P\left(x_j|C_i\right)$ is obtained from training set estimation. If A_j is classification attribute, $P\left(x_j|C_i\right)$ is equal to the proportion that attribute A_j is equal to x_j in training sample of type C_i.

3.2. Bayesian classification algorithm improved by mutual information

From a theoretical analysis, classification accuracy of Naive Bayes is higher than that of other classification algorithms such as decision tree and SVM. However, in terms of Bayesian classification model, different role of different attribute is not considered during sample classification. Attributes of redundant data will increase dimensionality of data, increase classification calculation, bring about noise impact and make classification accuracy lower [2]. On the basis of this case, there are algorithms for characteristics selection before classification. There are many commonly used methods for characteristics selection. However, in all of them, only the maximum of the sum of the relevancy between characteristic attribute A and each type $C_i (1 \leq i \leq m)$ is considered but this maximum is not able to be used to fully measure significance of each characteristic attribute A in classification. It shall also be considered whether the maximum measure of the relevancy between attribute A and type $C_i (1 \leq i \leq m)$ is significantly larger than the measure of the relevancy between that such attribute and other types $C_j (1 \leq j \leq m, i \neq j)$. That is, distribution of $MI(A_i, C_j)$ shall be considered. Otherwise, even though maximum relevancy is large, when there is high relevancy between such characteristic attribute and other types, such characteristic attribute may play a small role in classification, which means it is not suitable for classification.

In view of this, with respect to concept of credibility in references of this paper, the method of characteristic selection based on the relative credibility of mutual information is proposed. The relative credibility is defined as follows:

$$R = \frac{MI_1 - MI_2}{MI_2}\,. \tag{4}$$

Where

$$MI = MI(A;Cj) = \sum_{i=1}^{n} p(ai, Cj)\log 2\frac{p(ai, Cj)}{p(ai)p(Cj)} \tag{5}$$

In the equation, n is number of values of attribute A, m is number of types $(1 \leq j \leq m)$, MI_1 is maximum of mutual information between attribute A and types, and MI_2 is the second maximum of mutual information between attribute A and types. Higher value of $(MI_1 - MI_2)$ means bigger role of attribute A in classification and higher relative credibility. MI_2 serves as denominator in order to define relative reliability as a dimensionless relative value.

The relative reliability R is introduced into the Bayesian algorithm as an attribute weight so as to obtain the improved NB algorithm:

$$P(Ci|X) = \arg\max P(Ci)\prod_{j=1}^{K} RjP(xj|Ci)\,. \tag{6}$$

3.3. Algorithm implementation steps

Step 1: Original travel data are preprocessed, such as discretization and null value processing.

Step 2: For the training samples, the relative credibility $R_i(1 \leq i \leq N)$ of each attribute A_i in classification is calculated by the formula (4). R_i is arranged from high to low, the first K attributes are selected as the optimal attributes, and $(N - K)$ attributes arranged behind are deleted so as to obtain a new attribute set $B = \{B_1, B_2, \cdots, B_K\}$ after characteristic selection. K, an artificial predetermined integer, serves as the number of selected characteristic attributes.

Step 3: $P(C_i) = S_i/S$ is calculated by statistical approach. S_i is number of samples of type C_i. $P(C_i)$ stands for probability where sample of type (C_i) occurs in sample set.

Step 4: For optimal attribute set B_K obtained in step 2, conditional probability $P(X|Ci) = \prod_{j=1}^{k} P(xj|Ci)$ of each attribute B_K is calculated. $P(x_j|C_i)$ is equal to the probability that attribute B_j is equal to x_j in training sample of type C_j.

Step 5: The relative credibility R based on mutual information is used as a weight to carry out Bayesian classification:

$$P(Ci|X) = \arg\max P(Ci) \prod_{j=1}^{K} RjP(xj|Ci)\,.$$

3.4. Structure of self-organizing neural network model

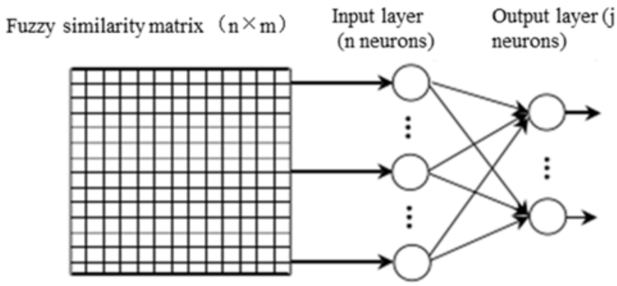


Fig. 2. Structure of self-organizing neural network model

Self-organizing neural network model is a multi-layer tree structure composed of input layer and competitive layer (ie output layer) [4]. Each input node of this model is associated with all neural trees and nodes by weight W, so as to reduce nonlinear dimensionality of input data. When the input maps to nodes of the same tree, the topological invariance is maintained. The number of neurons in the input layer is the number of rows or columns of the fuzzy similarity matrix (that is, the number of samples in the sample set), as shown in Fig. 2. By learning input iteratively, this structure can capture the mode characteristics contained in each input mode, self-organize them, and present the classification results at the competitive layer. When the network accepts a similar input with a memorized mode, the network recalls this mode and correctly classifies it. For a mode that does not exist in network memory, the self-organizing neural network memorizes this new mode on the premise that existing memory is not impacted.

Model learning samples consist of samples with N classification indexes. Assume these dots in N-dimensional space are obviously of the same type or some samples with similar characteristics are close in N-dimensional space, these close samples will constitute one type, which forms one cluster in N-dimensional space. When input samples respectively belong to more than one type, N-dimensional space will be characterized by multi-clustered distribution. Each cluster stands for one type. Center of cluster is exactly center of type clustering. The distance between samples of the same type and center of such type clustering shall be small. Such distance can be measured with Euclidean distance:

$$D_j = \sqrt{\sum_{i=1}^{N} (x_i - W_{ij})^2}\,. \tag{7}$$

Where: x_i is classification index, W_{ij} is center of clustering of the jth dynamic type, and D_j is Euclidean distance.

Self-organizing neural learning algorithm does not require teacher signals and judges type of sample with Euclidean distance from sample to center of clustering. Steps of the algorithm are as follows:

Step 1: Threshold β is given and used to control roughness of classification. Higher β means rougher classification and a smaller number of types. Smaller β means refined classification and a larger number of types. Thus, trial calculation shall be performed for β to determine as the case may be.

Step 2: Assume number of original neurons at input layer is 1(namely j=1), one learning sample is randomly selected so as to assign connection weight as an initial value.

Step 3: One new learning sample is input. Euclidean distance D_j between it and center of clustering W_{ij} of each dynamic type is calculated.

Step 4: The output neuron with the smallest Euclidean distance D wins the competition:

$$D_j^* = \min\{D_j\} . \tag{8}$$

Step 5: In case $D_j^* < \beta$, the current input sample is deemed to belong to the dynamic type represented by the output neuron, and the connection weight W_{ij} is adjusted as follows:

$$W'_{ij} = (x_i - W_{ij})/h_j . \tag{9}$$

Where: W'_{ij} is adjusted value of W_{ij}, h_j is number of current samples of the jth dynamic type. Then step 3) is reached.

Step 6: In case $D_j^* \geq \beta$, such output neuron wins the competition. However, current input sample cannot be deemed to belong to the dynamic type represented by such output neuron and shall belong to a new type. Thus, output neurons shall be increased by one $j = j + 1$, which stands for a new dynamic type. Such input sample serves as initial value of $W_{i(j+1)}$. Then step 3 is reached.

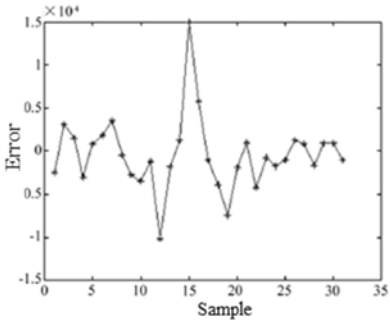
Step 7: A cycle goes on in this way until all samples have been learned. Number of output neurons of network model is that of types of all samples in the end. Connection weight is value of center of clustering of dynamic types.

The said learning algorithm shows that self-organizing neural network is characterized by plasticity and self-organization. Besides, network learning and training process is exactly process of dynamic classification for measured data. Network model, established after training is completed, is classification model. When new measured data are obtained, the network model can be input. And the dynamic type, which is represented by neuron at the output layer and wins the final competition, is the type to which the sample belongs. This is the dynamic identification process of the new data in the model.

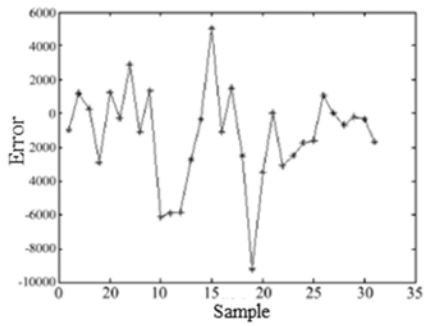
4. Experimental analysis

4.1. Standard index test

Hardware parameters: processor i7-6800HQ, internal storage 6G ddr3-1600, system win7 Ultimate. Algorithms are compared so as to select BP neural network algorithm. Samples in test set are respectively input to self-organizing neural network model and BP neural network model so as to perform simulation test and output error sum, root-mean-square error and error percentage of two model tests, as shown in Fig.3.



(a) Prediction Error of BP Neural Network



(b) Prediction Error of Self-organizing Neural Network

Fig. 3. Prediction error comparison

As seen from the data of Fig. 3, performance of self-organizing neural network algorithm is about 25% higher than that of BP neural network algorithm in term of the index error sum. This shows the performance advantage of self-organizing neural network in evaluation. The proposed self-organizing neural network algorithm is superior to BP neural network algorithm in terms of fitting with real expectation curve, and its prediction error is lower.

Table 1. Results of experimental comparison

Test set	Index	Literature[13]	Literature[14]	Literature[15]	ODSDC
PKU	Accuracy	86.3%	88.5%	93.4%	96.7%
	Recall rate	75.6%	81.4%	91.5%	95.4%
	F measured value	80.6%	84.9%	92.4%	96.0%
Sogou	Accuracy	84.3%	89.7%	92.5%	95.8%
	Recall rate	85.2%	83.4%	93.6%	97.8%
	F measured value	84.7%	86.4%	93.0%	96.8%
FD	Accuracy	89.6%	93.4%	95.9%	98.1%
	Recall rate	82.3%	85.4%	90.5%	92.3%
	F measured value	85.8%	889.2%	93.1%	95.1%

Algorithms are compared so as to select literature [13 to 15]. In literature [13], nonnegative tensors are decomposed for large-domain knowledge corpus so as to match with ontology concept domain knowledge. In literature [14], domain knowledge classification technique is designed on the basis of data set in ontology diagram. These two algorithms are designed for non-domain knowledge. In literature [15] based on classification method for apparent domain knowledge and potential Chinese domain knowledge, the algorithm is especially aimed at Chinese domain knowledge classification. Threshold $\theta = 0$ and Table 1 shows results of experimental comparison.

It can be learned from Table 1 Results of Experimental Comparison that the ODSDC algorithm is superior to the contrast algorithm in indexes, accuracy of Chinese domain knowledge classification, recall rate and F measured value, in comparison with literatures [13–15]. Algorithms designed in literatures [13, 14] are respectively aimed at domain knowledge and ontology diagram. Thus, effects of their application to Chinese domain knowledge classification are not ideal. In literature [15], algorithm is designed for Chinese corpus. Thus, its algorithm effects are superior to those in literatures [13 and 14] but inferior to ODSDC algorithm.

4.2. Results of higher education evaluation

In order to verify the superiority of the quality evaluation model of remote higher education based on Bayesian self-organizing neural network, higher education quality evaluation models of literatures [3, 5 and 10] are selected for a contrast experiment. Fig.4 and 5 show their experimental results. From Fig. 4 and 5, it can be seen that, compared with the quality evaluation of current typical remote higher education, remote higher education quality evaluation based on Bayesian self-organizing neural network is more accurate and more reliable. A superior quality evaluation model for remote higher education is established.

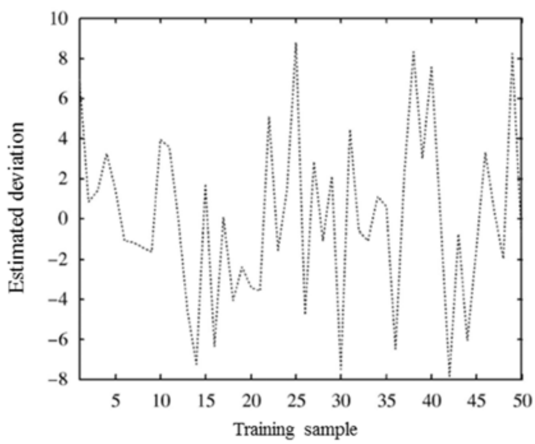


Fig. 4. Results of higher education quality evaluation

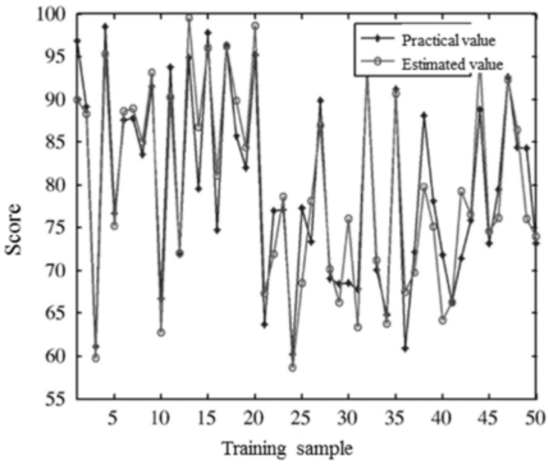


Fig. 5. Estimated deviation in higher education quality

5. Conclusion

In this paper, quality evaluation model for remote higher education, based on Bayesian self-organizing neural network, is proposed. In accordance with the defect that quality evaluation indexes of remote higher education are numerous, higher education evaluation index system is established. Then Bayesian self-organizing neural network classifier is able to be established in combination with index system and example data. And higher education evaluation is performed on the basis of Bayesian self-organizing neural network classifier. Next, for knowledge, which is improved on the basis that conventional Bayesian classification algorithm is insufficient and where mutual information is applied, characteristics are selected by means of relative credibility so as to delete redundant attributes in a way that obtains improved Bayesian algorithm and improves algorithm performance.

Acknowledgement

Henan province soft science research project in 2018 "Research on the integration of innovation and entrepreneurship education in Henan universities and the 2 construction of intelligent city".

References

[1] J. J. Faig, A. Moretti, L. B. Joseph, Y. Y. Zhang, M. J. Nova, K. Smith, and K. E. Uhrich: *Biodegradable Kojic Acid-Based Polymers: Controlled Delivery of Bioactives for Melanogenesis Inhibition*, Biomacromolecules *18* (2017), No. 2, 363–373.

[2] N. Arunkumar, V. Venkataraman, Thivyashree, Lavanya: *A moving window*

approximate entropy based neural network for detecting the onset of epileptic seizures. International Journal of Applied Engineering Research *8* (2013), No. 15, 1841–1847.

[3] Y. J. Zhao, L. Wang, H. J. Wang, and C. J. Liu: *Minimum Rate Sampling and Spectrum Blind Reconstruction in Random Equivalent Sampling.* Circuits Systems and Signal Processing *34* (2015), No. 8, 2667–2680.

[4] S. L. Fernandes, V. P. Gurupur, N. R. Sunder, N. Arunkumar, S. Kadry: *A novel nonintrusive decision support approach for heart rate measurement.* Pattern Recognition Letters. https: //doi.org/10.1016/j.patrec.2017.07.002 (2017).

[5] N. Arunkumar, K. Ramkumar, V. Venkatraman, E. Abdulhay, S. L. Fernandes, S. Kadry, S. Segal: *Classification of focal and nonfocal EEG using entropies.* Pattern Recognition Letters *94* (2017), 112–117.

[6] J. W. Chan, Y. Y. Zhang, and K. E. Uhrich: *Amphiphilic Macromolecule Self-Assembled Monolayers Suppress Smooth Muscle Cell Proliferation*, Bioconjugate Chemistry *26* (2015), No. 7, 1359–1369.

[7] M. P. Malarkodi, N. Arunkumar, V. Venkataraman: *Gabor wavelet based approach for face recognition.* International Journal of Applied Engineering Research *8* (2013), No. 15, 1831–1840.

[8] L. R. Stephygraph, N. Arunkumar: *Brain-actuated wireless mobile robot control through an adaptive human-machine interface.* Advances in Intelligent Systems and Computing *397* (2016), 537–549.

[9] W. S. Pan, S. Z. Chen, Z. Y. Feng: *Investigating the Collaborative Intention and Semantic Structure among Co-occurring Tags using Graph Theory.* International Enterprise Distributed Object Computing Conference, IEEE, Beijing (2012), 190–195.

[10] Y. J. Zhao, Y. H. Hu, J. J. Liu: *Random Triggering-Based Sub-Nyquist Sampling System for Sparse Multiband Signal.* IEEE Transactions on Instrumentation and Measurement *66* (2017), No. 7, 1789–1797.

[11] X. Du, L. Chen, D. Huang, Z. Peng, C. Zhao, Y. Zhang, Y. Zhu, Z. Wang, X. Li, G. Liu: *Elevated Apoptosis in the Liver of Dairy Cows with Ketosis.* Cellular Physiology & Biochemistry *43* (2017b), No. 2, 568–578.

Received May 7, 2017