Construction of Student model based on BP neural network

Yunxiang Liu

School of Computer Science and
Information Engineering
Shanghai Institute of Technology
Shanghai, China
yxliu@sit.edu.cn

Yuanyuan Zhang

School of Computer Science and
Information Engineering
Shanghai Institute of Technology
Shanghai, China
Zhyuyu0424@163.com

Guoqing Zhang

School of Computer Science and
Information Engineering
Shanghai Institute of Technology
Shanghai, China
749229869@qq.com

Abstract: With the development of personalized learning, the construction of student models is becoming more and more important. At present, there are still problems in the student model that the characteristics are single and the indicators of each dimension are not clear. In this paper, learners will be analyzed from the perspective of student characteristics. And BP (Back Propagation) neural network algorithm will be used to establish a personalized student model. This paper first constructs the feature system of the student model from six dimensions. Secondly, the initial data is obtained through questionnaire survey, and the data is initialized to obtain 30 feature vectors as input to BP neural network. The output of the network is a learner type, which is divided into 36 categories. The construction of the student model will have certain practical significance for realizing the effectiveness of personalized education in distance education.

Keywords: Student Model, Personalized Learning, BP Neural Network, Student Characteristics

A. INTRODUCTION

With the rise of educational big data and the development of online courses, personalized learning has become a research hotspot in the field of education. The student model is the basis for the realization of personalized learning in online courses and the core component of the intelligent education system. Student characteristics are key factors in the student model. The integrity and accuracy of features are directly related to the generation of effective individualized learning programs, which affects the learning efficiency and enthusiasm of learners. Therefore, research on student characteristics, modeling techniques, etc. in the student model is crucial.

The earliest foreign student model was proposed by Carbonell in 1970, which uses semantic networks to represent domain knowledge. At present, the common foreign student model have Overlay model, Differential model, Perturbation model, Stereotype model, Constraint-based model, Bayesian model, Fuzzy Student model, Machine Learning model, etc. These models are mostly in a state of knowledge of learners, learning styles or meta-cognitive characteristics as a reference, using different modeling methods to build models, so as to realize individualized teaching. For example: Dynamic Bayesian networks are used to construct student models to represent multiple skills in one model [1]. Ciolacu et al. (2018) used an early recognition system based on machine learning algorithms to predict the final grades of students before the final exam compared with foreign countries, the research on the construction of personalized student model in China is based on the reference of foreign literature, mainly for three aspects.

The first is to theoretically describe the student model or empirically study one or several features. For example: Sun & Zhang (2017) describe learners from four dimensions: individual attribute, cognitive ability, learning style and learning attitude [3]. Mou & Wu (2019) designed a six-dimensional feature analysis model, combined with scene-aware modeling method and frequent sequence mining algorithm to construct a personalized student model with scene characteristics [4]. The second is the construction of a student model based on modeling techniques. Wang (2017) uses coverage modeling technology and data-driven technology to construct a student model for learners' knowledge state, knowledge level, learning behavior and other personality characteristics, and update it in real time [5]. The third is the static and dynamic student model. For example, Li (2017) constructed a dynamic student model by using data mining technology to personalize the learning materials to learners [6].

This paper proposes a student classification model based on BP neural network by studying the related literature and examples of student models. The article will introduce the BP neural network student model from the three aspects of student feature index system construction, model construction and model validation.

II. STUDIES ON CHARACTERISTIC INDEX SYSTEM

At present, there is no uniform definition of student models, and researchers have different understandings of the construction of student models from different perspectives. In response to this problem, domestic and foreign authorities have issued relevant student model specifications: IMS LIP (Global Learning Consortium Learner Information Package)

specification^[7], IEEE PAPI (Public and Private Information) specification^[8], CELTS -11 specification^[9].

In order to accurately describe the true state of learners in online courses, this paper builds a learner's characteristic index system based on the CELTS-11 student model specification. This learner information structure is divided into six dimensions, including the dominant and recessive features of the learner, which can fully represent the learner's true learning state.

A. Demographic characteristics

In the study of demographic characteristics, Wang (2006) et al. attributed the experience to demographics^[10]. However, experience often has an important influence on learners' cognitive ability and learning style. Therefore, this article does not study empirical features separately. At the same time, age and gender are included in cognitive ability and academic qualifications, and no separate study is needed. Demographic characteristics of this paper are shown in the figure below.

Fig.1 Demographic characteristics

B. Learning motivation

Learning motivation is a motivational tendency to initiate and maintain students' learning behaviors and to point them to certain learning goals. Learning motivation plays an important role in the formulation and implementation of learners' learning strategies. In order to be close to the real motivation of learners, this paper uses a research result of Tsinghua University Education Research Institute. The study of learner motivation is based on the "XuetangX" learners, and through the machine learning method, the motivation texts of MOOC (massive open online courses) learners are self-filled^[11]. Learning motivation feature information is shown in Fig.2.

C. Learning style

Learning style is a way for learners to express their personal characteristics when studying and solving learning tasks. In the process of online learning, providing learning strategies or learning materials that are appropriate to the learner's learning style will enhance the learner's interest in learning. At the same time, it is beneficial to improve the learning efficiency of learners. This paper uses the classic learning style measurement method - Solomon Learning Style Scale^[12] to investigate the learner's learning style. The learning style feature map is shown in Fig.3.

Fig. 2. Learning motivation characteristics information

Fig. 3. Learning style feature map

D. Cognitive ability

Cognitive ability is the ability of the human brain to process, store and extract information. And it is the most important psychological condition for people to complete activities. The strength of cognitive ability determines the difficulty level of

the course suitable for the learner. It is also a prediction of the learner's ability to learn. The cognitive abilities related to work and learning are language ability, calculation ability, perception ability, space ability, and reasoning ability. For example, a set of cognitive tests at the Princeton Educational Testing Service

Center in New Jersey. It involves test projects such as graphic matching, vocabulary comprehension, graphic categorization, and mathematical calculations, and examines cognitive abilities related to work-learning^[13]. The cognitive ability characteristics information is shown in the figure below.

Fig. 4. starting point capability characteristics information

E. Online behavior

The online learning behavior data hides the innumerable hidden features of the learner. Mining online data can

continuously improve the student's model in the system, so that it is infinitely close to the learner's real learning situation. A complete and accurate student model helps to personalize the

curriculum design while helping to improve learner learning efficiency. This paper analyzes online behavior from the perspective of login, information interaction, and homework^[14]. The online behavior information is shown in Fig.5.

F. Information literacy

Information literacy is a kind of ability that people need in the information age. This ability can determine when information is needed and how to get information, how to evaluate and effectively use the information you need. In the online course study, the learner's information literacy will continue to affect the learner's learning mood and learning effect. Considering the reliability of the research, the information literacy of this paper refers to the information quality standards of universities in Beijing^[15]. The information literacy standards are shown in Fig.6.

Fig. 5. online behavior characteristics information

Fig. 6. information literacy standard

III. STUDENT MODEL CONSTRUCTION

In the process of learning, learners' cognitive abilities and knowledge will constantly change. However, the changes of learner are uncertain and ambiguous, and learners cannot clearly perceive the degree of change. This paper chooses BP neural network algorithm to model students, considering the following two points.

(1) The BP neural network does not need to determine the mathematical equation of the mapping relationship between input and output in advance, and only learns certain rules through its own training, and obtains the result closest to the expected output value when given the input value.

(2) The neural network is particularly suitable for processing related data when it is necessary to bring in certain conditions and the information itself is uncertain and ambiguous.

A. BP neural network working principle

BP neural network is a multilayer feed forward neural network trained according to the error back propagation algorithm. The BP neural network generally includes a three-layer network of an input layer, an implicit layer, and an output layer. In the BP algorithm, the learning process consists of two processes: forward propagation of the signal and back propagation of the error. The calculation error is output positively, while the adjustment weight and threshold are

reversed. In the case of forward propagation, the sample is passed in by the input layer, processed by the hidden layer, and passed to the output layer. If the actual output does not match the expected output, the network will automatically “learn” and turn into the error back phase. Error back propagation is to propagate the output error back to the input layer through the hidden layer, and distribute the error to all the units of each layer, so as to obtain the error signal of each layer unit. Take the method of reducing the weight of the connection for the node that caused the error. At the same time, increase the connection weight of the node making the correct recommendations. This error signal serves as the basis for correcting each unit. The algorithm iterates over and updates the weights and errors until the error in the network output is reduced to an acceptable level or to a predetermined number of learnings. The working principle diagram of BP neural network is shown below.

Fig. 7. BP neural network working principle diagram

B. BP neural network algorithm derivation process

Suppose there are n training samples $\{(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)\}$, and the batch update method is used to define the error function for the n samples.

$$E = \frac{1}{n} \sum_{i=1}^n Ei \quad (1)$$

Ei is the training error of a single sample:

$$Ei = \frac{1}{2} \sum_k^m (d_{ki} - y_{ki})^2 \quad (2)$$

Wherein, d_i is the desired output corresponding to the input sample x_i . Available from (1) and (2):

$$E = \frac{1}{2n} \sum_{i=1}^n \sum_{k=1}^m (d_{ki} - y_{ki})^2 \quad (3)$$

Each time the BP neural network algorithm updates the weights and offsets based on the error between the output value and the expected value:

$$W_{ij}^{(l)} = W_{ij}^{(l)} - \alpha \frac{\partial E}{\partial W_{ij}^{(l)}} \quad (4)$$

$$b_i^{(l)} = b_i^{(l)} - \alpha \frac{\partial E}{\partial b_i^{(l)}} \quad (5)$$

In the formula, α is the learning rate. And it is in the range of $(0, 1)$. The algorithm updates the weights and offsets after each iteration until the desired output is obtained.

C. Model Construction

1) Parameter settings

The training data in this paper was collected by online questionnaire survey. The 1000 data collected are divided according to the ratio of 7:3, 70% of the data is used for sample training, and 30% of the data is used for data verification. The network topology uses the classic three-layer design of BP neural network: input layer, hidden layer and output layer. The neurons in the input layer are the corresponding 30 feature vectors, and the output layer is 36 feature vectors. The 30 feature vectors corresponding to the input layer include education, learning place, cognitive ability, information literacy, 9 kinds of learning motivation, 8 types of learning styles, and 9 types of online behavior. The 36 feature vectors of the output layer are 36 learner types, such as: Master's degree in Area A, academic type, and strong ability type. The selection of the number of hidden layer nodes is very important. If the number of nodes is too small, the training effect will be poor. If there are too many nodes, the training time will be long and it will easily fall into the local minimum. Based on many experiments, the relationship between the number of nodes in different hidden layers and the model loss is obtained, as shown in the figure.

Fig. 8. Relationship between model loss and number of hidden layer nodes K

It can be seen from the figure that as the number of hidden layer nodes increases, the model loss becomes smaller and smaller, and the model training effect becomes better and better. According to the idea of the elbow method, we choose $K=40$ as the number of hidden layer nodes, and set the hidden layer to 2 layers. In addition, the selection of activation functions and learning rate related parameters in the BP neural network model is very important for model training. There are many activation functions, such as sigmoid function, Tanh function, Relu function, Softmax function and so on. The sigmoid function can map a real number to the $(0, 1)$ interval, which is suitable for the

two classifications. However, when it is back propagating, it is prone to the problem of gradient disappearance and it is impossible to complete the training of the deep network. The Tanh function works well when the features are significantly different, but it does not solve the gradient dispersion problem. The Relu function solves the partial gradient dispersion problem and converges faster than Tanh. The Softmax function is suitable

for the output of a multi-class neural network. Since the cross-entropy function derivation result is relatively simple and easy to calculate, it can solve the problem that some loss function learning is slow. Therefore, the Relu function is used as the activation function of the hidden layer, and the output layer uses the Softmax activation function. Relevant important parameters are shown in Table I.

TABLE I. RELEVANT IMPORTANT PARAMETERS

Input layer node number	Number of hidden layer nodes in the first layer	Number of hidden layer nodes in the second layer	Number of output layer nodes	Activation function	Learning rate	Number of iterations
30	40	40	36	Relu , Softmax	0.01	2000

Fig.10 is a block diagram of the model training detection. The specific training and detection process is:

Fig. 9. model training detection block diagram

2) Softmax cross entropy loss function

In a neural network, Softmax generally as classification task of the output layer , a plurality of neural networks obtained were normalized, so that the value obtained in (0,1) between . Let the output of the neuron be:

$$z_i = \sum_j w_{ij}x_{ij} + b \quad (6)$$

Where w_{ij} is the j th weight of the i th neuron, b is the offset, and z_i represents the i th output of the network. The i th output value of Softmax is as follows:

$$S_i = \frac{e^{z_i}}{\sum_k e^{z_k}} \quad (7)$$

Let p and q be two distributions about the sample set, where p is the true distribution of the sample set, q is the estimated distribution of the sample set, and the estimated distribution q is used to represent the average code length of the true distribution p . Definition of cross entropy:

$$H(p, q) = \sum_i^n p_i \log \frac{1}{q_i} = \sum_i^n -p_i \log q_i \quad (8)$$

Combine the concept of cross entropy function and Softmax to get the loss function of Softmax:

$$L = \sum -y_i \ln S_i \quad (9)$$

Wherein y_i is i th category true value, only the value 0 or 1.

IV. MODEL VALIDATION

A. Experimental environment configuration

The experimental environment configuration is shown in Table II.

TABLE II. EXPERIMENTAL ENVIRONMENT CONFIGURATION

Name	Related configuration
Central processor memory	Intel(R)-CPU-E5-2620-V4, 64G
GPU	2×Nvidia TitanX
GPU acceleration library	CUDA9.0, CUDNN7.0.3
Data processing	Python

B. Model training and evaluation results

Pre-processing the initial data to obtain input samples of the BP neural network. The learner's degree, undergraduate, master's and doctor's degrees are respectively recorded as 0, 1, and 2; the A and B areas are recorded as 1, 0, respectively. Learning motivation and learning style data are in the same way, labeled 0 or 1. For example, if a learner participates in an online course study due to cognitive interest (learning interests and hobbies), the item is recorded as 1, otherwise it is 0. The normal distribution method is used to compare and classify learners' cognitive ability, and finally they are divided into five criteria, labeled as 4, 3, 2, 1, and 0, respectively. For the processing of online behavior data, it is divided into three criteria: positive, general and passive, which are marked as 2, 1, and 0 respectively. The data collected by information literacy is also divided into three categories: strong, general, and weak by normal distribution methods, which are labeled as 2, 1, and 0, respectively. When the demographic characteristics are not considered, the learners are divided into six categories: strong academic ability, general academic ability, weak academic ability, strong application ability, general application ability,

and weak application ability. Join the learner's academic qualifications and learning factors, and finally divide it into 36 learner types, corresponding to 0, 1, 2...35. Such as: A area, undergraduate, strong ability type corresponding output 0, B area doctor, application type, weak ability corresponding output is 35.

The collected neural network model was used to classify and predict the 300 pieces of verification data, and compared with the test results. See Table III for some comparisons. The predictive output of the student model is the learner's assessment of the type of his or her own. The actual output is the learner type that inputs the eigenvalues into the BP neural network and is output by the network. X_1 - X_{30} in the table is the 30 eigenvalues after processing, which we can see in Table 3. Of the 10 verification data randomly selected, 8 data classification predictions are correct, and only 2 data classification prediction errors. It can be seen that the learner model based on BP neural network achieves a good prediction effect, with a correct rate of 80%, with high prediction accuracy, which can better classify learners and contribute to future applications and research.

TABLE III. MODEL PREDICTIVE OUTPUT VS. ACTUAL OUTPUT

Serial number	1	2	3	4	5	6	7	8	9	10
Input sample	X_1	1	0	0	1	1	2	1	0	2
	X_2	1	1	0	1	1	0	0	0	1
	X_3	1	0	1	1	0	1	0	0	1
	X_4	1	1	1	1	1	1	1	0	1
	X_5	1	0	1	0	0	1	0	0	1
	X_6	0	1	0	0	1	0	1	0	1
	X_7	1	1	1	1	1	1	1	1	1
	X_8	0	0	0	0	0	0	1	0	0
	X_9	1	1	1	1	1	1	1	1	1
	X_{10}	0	0	1	1	0	1	0	1	1
	X_{11}	0	1	0	0	1	0	1	0	0
	X_{12}	1	1	0	0	1	0	1	0	0
	X_{13}	0	0	1	1	0	1	0	0	1
	X_{14}	1	1	1	0	0	1	1	1	0
	X_{15}	0	0	0	1	1	0	0	0	1
	X_{16}	0	0	1	1	1	1	1	1	1
	X_{17}	1	1	0	0	0	0	0	1	0
	X_{18}	1	1	0	0	0	0	0	1	1
	X_{19}	0	0	1	1	1	1	1	0	0
	X_{20}	3	2	2	2	2	2	2	2	3
	X_{21}	1	1	0	1	1	1	0	1	2
	X_{22}	1	0	1	0	1	0	2	1	0
	X_{23}	2	1	1	1	2	0	2	0	1
	X_{24}	1	1	0	1	1	1	0	1	0
	X_{25}	1	0	0	0	0	1	1	0	1
	X_{26}	2	1	0	1	1	1	0	0	1
	X_{27}	1	1	1	1	2	1	2	1	1
	X_{28}	2	1	1	1	1	1	2	1	2
	X_{29}	1	1	1	1	1	1	2	1	1
	X_{30}	2	1	1	1	1	2	1	1	2
Forecast classification		6	4	19	7	9	31	21	22	12
Actual classification		6	4	19	7	10	31	21	23	12

V. CONCLUSION

Based on the experiment result, this paper establishes a BP neural network student model by constructing the learner's feature system. The model predicts the learners and provides a basis for the learner's personalized learning program. The characteristic system of the student model covers the implicit and dominant features of the learner and can better represent the true state of the learner. The introduction of BP neural network can predict the learning type of the learner and provide a reliable basis for adjusting the learning state. The final result shows that the model has good training effect and good classification and recognition function, and the recognition rate reaches 80%. At the same time, the model still has errors, and the number of the type of learners is small, which leads to some inaccuracies in the description of learners. In the next step, the accuracy of the model will be further improved.

References

- [1] Ferreira H N M, Brant-Ribeiro T, Rafael D. Araújo, et al. An Automatic and Dynamic Student Modeling Approach for Adaptive and Intelligent Educational Systems using Ontologies and Bayesian Networks[C]// IEEE International Conference on Tools with Artificial Intelligence. IEEE, 2016.
 - [2] Ciolacu M, Tehrani AF, Binder L, et al. Education 4.0-Artificial Intelligence Assisted Higher Education: Early recognition System with Machine Learning to support Students' Success[C]//2018 IEEE 24th International Symposium for Design and Technology in Electronic Packaging (SIITME). IEEE, 2018: 23-30.
 - [3] Li Sun, Ting Zhang. Design and Analysis of Personalized Student Model in Network Education [J].Journal of Distance Education, 2017, 35(03):93-101.
 - [4] Zhijia Mou, Fati Wu. Research on electronic resources based on individualized learning recommendation student model [J]. Educational Technology Research, 2015, 36 (1): 69-76.
 - [5] Liping Wang. Research on student model and teaching model in adaptive learning system [D]. Northeast Normal University, 2017.
 - [6] Deyang Li. Research on resource push scheme in personalized learning system [D]. Northeast Normal University, 2017.
 - [7] IMS learner information package summary of changes version 1.0.1 final specification [EB/OL].[2005-01-04].https://www.imsglobal.org/profiles/lipv1p0p1/imsaip_sumcv1p0p1.html.
 - [8] IEEE Learning Technology Standards Committee (LTSC) [EB/OL]. [2018-12-01]. <http://ltsc.ieee.org/>.
 - [9] Ministry of Education Education Information Technology Standards Committee. CELTS-11: Student Model [EB/OL].<http://www.celtsc.org/content/xxzl/40288b88391ee7c701391ef e11810021.html#>, 2012-08-13/2019-7-19.
 - [10] Ying Wang, Yujie An, Huairong Huang. Fortune to learn who features build model study [J]. Open Education Research, 2006 (01): 47-53.
 - [11] Lei Zhang, Huafeng Zhang. Why MOOC-- based "school online" learning by data research online learning motivation [J]. Jiangsu Higher Education, 2019 (04): 50-57.
 - [12] Soloman BA, Felder R M. Index of learning styles questionnaire [J]. NC State University. Available online at: <http://www.engr.ncsu.edu/learningstyles/ilsweb.html> (last visited on 14.05. 2010), 2005, 70.
 - [13] Shen Jiliang, Chen Bo, Wang Dahua. Research on the development of basic cognitive ability in adulthood [J]. Acta Psychologica Sinica, 2000(01):54-58.
 - [14] Yibin Yin. Research on Multidimensional Characteristics of Adult Learners in Online Environment [D].East China Normal University, 2017.
- Na Ri, Wu Xiaowei, Lu Jihong. Current Status and Prospects of Research on Information Literacy Standards at Home and Abroad [J]. Library and Information Service, 2010, 54(03):32-35.