

RESEARCH

Open Access



A learning style classification approach based on deep belief network for large-scale online education

Hao Zhang^{1,2*}, Tao Huang^{1,2†}, Sanya Liu^{1,2†}, Hao Yin³, Jia Li^{1,2}, Huali Yang^{1,2} and Yu Xia^{1,2}

Abstract

With the rapidly growing demand for large-scale online education and the advent of big data, numerous research works have been performed to enhance learning quality in e-learning environments. Among these studies, adaptive learning has become an increasingly important issue. The traditional classification approaches analyze only the surface characteristics of students but fail to classify students accurately in terms of deep learning features. Meanwhile, these approaches are unable to analyze these high-dimensional learning behaviors in massive amounts of data. Hence, we propose a learning style classification approach based on the deep belief network (DBN) for large-scale online education to identify students' learning styles and classify them. The first step is to build a learning style model and identify indicators of learning style based on the experiences of experts; then, relate the indicators to the different learning styles. We improve the DBN model and identify a student's learning style by analyzing each individual's learning style features using the improved DBN. Finally, we verify the DBN result by conducting practical experiments on an actual educational dataset. The various learning styles are determined by soliciting questionnaires from students based on the ILS theory by Felder and Soloman (1996) and the Readiness for Education At a Distance Indicator. Then, we utilized those data to train our DBNLS model. The experimental results indicate that the proposed DBNLS method has better accuracy than do the traditional approaches.

Keywords: Large-scale online education, Adaptive learning, Deep belief network, Learning style, High-dimensional

Introduction

With the deepening integration of big data [1] and education, the learning revolution, represented by MOOCs [2], Khan academy [3] and Flipped classroom [4], has had a strong impact on traditional forms of education and has highlighted the importance of large-scale online education in reshaping education reshaping—that is, globalized resources, modularized supply, personalized teaching and independent study can be implemented by

restructuring and process reengineering. Large-scale online education has made significant breakthroughs in teaching environments, content presentation, teaching modes, and learning evaluations. Large-scale online education also innovates and revolutionizes study approaches and is an effective way to promote educational equity and improve teaching quality. By offering new learning situations and technologies, the unique characteristics of MOOCs have given rise to a new learning model, personalized learning [5, 6]. In a MOOC environment, learners can select and order customized courses according to their own purposes and backgrounds. However, the materials in MOOCs are rather complicated; consequently learners can easily become trapped in “learning Trek” and “cognitive overload” situations [7] during the process of receiving information. Nevertheless, MOOCs provide the

* Correspondence: zhanghao@mail.ccnu.edu.cn

†Hao Zhang, Tao Huang and Sanya Liu contributed equally to this work and should be considered co-first authors

¹National Engineering Laboratory for Educational Big Data Central China Normal University, Wuhan, China

²National Engineering Research Center for E-Learning, Central China Normal University, Wuhan, China

Full list of author information is available at the end of the article

same learning materials and learning activities for all learners and ignore individual differences because they fail to analyze individual learning behaviors. As a result, MOOCs cannot achieve the ability to teach students according to their aptitudes [8]. By analyzing online learning behaviors, we can accurately identify a user's learning characteristics and recommend personalized resources to help them improve their quality of learning. At present, it is urgent to determine how to analyze and identify the learning features of individual users by mining big data in adaptive learning [9, 10].

In an online learning environment, individual differences between learners are clearly evident in terms of time, learning duration, selected learning content, online interactions, etc. Some learners prefer to complete learning tasks during the daytime, while others prefer learning at night. Some learners participate in discussions or group activities on discussion forums or through other social networking applications, while others prefer quiet study. Among all the individual characteristics, learning style is an important factor that affects learners' individual differences. In other words, different learners have different tendencies in terms of learning style [11].

To date, researchers have made important contributions regarding how to apply learning styles to online learning, especially in the field of learning style identification and prediction [12–15]. Most of these studies mainly collected and recorded real behavior data left by learners during the network learning process to build a set of learners' network learning behaviors, and used data mining algorithms, neural networks or simple calculation rules to automatically detect learning; thus some research results have been obtained. However, the central problem is that these studies were not based on large-scale online learning platforms; consequently, they do not meet the demands of modern online learning platforms for effective learning style detection based on a large number of complex network learner behaviors.

Therefore, it is urgent to solve problems such as how to identify learners' learning styles, how to guide individual learners to construct learning objectives and plans, and how to recommend specific learning resources that reflect an individual's needs and abilities. Identifying learning styles through learners' online behavior and classifying them correctly plays an important role in realizing adaptive learning. In a MOOC environment, the following three important problems must be solved to accurately identify and classify learning styles:

- 1) What type of learning style model is suitable in MOOC learning environments?
- 2) What is the relationship between learning behavior and learning styles in MOOC learning environments?

- 3) What classification method can be used to overcome the problem of inaccurate classification due to the high dimensionality of learning behavior data in MOOC environments?

Through intensive study of the relevant literature concerning learning styles, we found that the learning style model proposed by Felder-Silverman is suitable for MOOC online learning environments. Meanwhile, the online learning relationship between behavior characteristics and learning styles is complex. Through our observations, together with the results obtained from experts, we found that differences in student behavior can be described using the following four dimensions: information processing, perception, input, and understanding. Therefore, the characteristic indicators should be developed based on these four learning behavior dimensions to form the data mining dimension.

We propose a learning style classification based on a DBN for MOOCs, called DBNLS. In this method, the inherent characteristics of learners—their learning styles—are adopted as the classification criteria. First, we summarized and analyzed individual learner differences and preferences to build a learning style model suitable for MOOC learning environments. Then, we determined which learning-habit indicators should be used based on expert experience and linked the indicators to learning styles within individual sessions. The deep learning model DBN was used to learn the high-dimensional learning style features and model learning styles to classify students accurately. Finally, we collected network learning behavior data by analyzing the weblogs left during online learning sessions on StarC [16], a MOOC platform used at Central China Normal University. Meanwhile, we also conducted offline empirical research and collected learning style questionnaire data, which we used as training samples to train the DBN model. The trained DBNLS model was applied to classify students' learning styles. The results show that the method proposed in this paper is superior to the traditional methods.

The main contributions of this paper are as follows:

- 1) Learners' intrinsic characteristics and learning styles are introduced as an important standard for learner classification. Simultaneously, the explicit attributes of the network behavior are effectively mapped to the intrinsic characteristic learning style indicators. These network behavior indicators serve as important DBN inputs for classifying students' learning styles.
- 2) We introduce social interaction factors into the learning style model, with the view of including the network learning environment characteristics and the learner interactions in the learning platform.

The learning style model established in this paper focuses not only on the static learning resources available to students but also on how they interact with others. For example, some students tend to complete tasks through communication and discussion, while others tend to think and work independently.

- 3) We introduce a deep-learning algorithm into learning classification in the field of education. This approach effectively overcomes the problems of the sharp rise in computational complexity resulting from high-dimensional data in traditional classification methods and data overfitting conditions.
- 4) The research was conducted using a practical online learning activity. We collected and distinguished both online and offline data. The offline data served as training data used to train the model, which subsequently was used to classify students' online learning behaviors. The results show that the proposed mechanism is considerably more accurate than is the traditional classification model.

The rest of this paper is organized as follows: [Related work and background](#) introduces the related works and provides some background concerning learning styles, the restricted Boltzmann machine (RBM) learning model, and BP. We then present the learning style model and the classification model in [Learning style model in MOOCs](#). We present our DBNLS algorithm and its training model in [Learning style detection based on deep belief neural networks](#). The details of the experimental evaluation are described in Section 5, and this work is concluded in [Conclusions](#).

Related work and background

Research on learning style theory

The concept of "learning style", first defined by Herbert Thelen, has since evolved into dozens of learning style theories and has been put into practice in the field of education. With the teaching methods described in "Teaching Students in Accordance with Their Aptitude" and "Learner Centered", an increasing number of scholars have shifted their focus to the learning style of learners, which is hoped to be fully considered in the MOOC design process. In recent years, the rapid development of online educational tools such as MOOCs has inspired scholars to consider how to reflect different learning styles in online education so that appropriate materials and methods could be suggested accordingly to help improve learning efficiency [11–15, 17].

Theories on learning style have become relatively mature after a long research history. Many scholars have proposed sophisticated learning style models. According

to Curry's learning style model, all learning styles can be classified into three levels, namely, "teaching preference" the outer level, "information processing mode" the middle level, and "cognitive style" the innermost level [18]. Learning style models of this type include Kolb [19], Honey and Mumford [20], Dunn [21] and Felder-Silverman's learning style models [11]. In addition, other models such as cognitive styles, VARK, and Keefe's learning style model, propose different definitions of learning style.

Dunn's learning style model is the representative theory of the "onion model" at the outer level. Dunn was mainly concerned about the stimuli that influence learning activities [19]. These stimuli are related to the learning environment, social environment, physiological factors, psychological factors and emotional factors. However, all these stimuli are highly unstable and vulnerable as can easily be observed. In contrast, Kolb was interested in the learning process. He suggested that each learning process goes through four self-related phases. Learners exhibit different preferences towards these phases [19]. Based on the learning models in [19, 22], and [20], Felder-Silverman provided a brand-new learning style model [11] that focuses on learners' individual cognitive characteristics and provides a comprehensive description of learning styles by combining information processing, information cognition, information input and information understanding.

Felder also designed the Solomon Learning Style Scale based on the learning style model, which provides a good method for measuring learning styles. As a result, Felder-Silverman's model is not only widely used in practice but also suitable for a web-based learning environment. Moreover, the Solomon Scale enjoys fairly good reliability and validity and has become popular in the educational field [23, 24]. Although scholars have provided different definitions, they all include the three main characteristics of learning styles. First, learning style varies among individuals, which means that different learners tend to have different learning style preferences. Second, learning style formation is affected by stimuli from both the outside environment and the inner self, such as cultural discrepancies, family factors, educational factors, and physiological factors. Third, learning style affects learning behavior. Learners with different learning styles show differences in learning strategies and learning habits.

Because the measurement approach biased toward essential characteristics (cognitive characteristics) by Felder-Silverman fits web learning, and Felder-Silverman's model has been proved through numerous experiments to have gained a high frequency in practice, as shown in Table 1, this paper chooses Felder-Silverman as the base learning style model [25]. However, characteristics such as social

Table 1 The frequency of typical learning model uses in real environments

Learning style model ($n = 70$)	number	percent
Felder-Silverman	35	50
Cognitive styles	12	17.14
Kolb	6	8.57
VARC	5	7.14
Honey and Mumford	4	5.71
Other	3	4.29
Not specified	5	7.14

interactions in new learning environments, such as MOOCs, were not considered in Felder-Silverman's model even though they should be considered in web learning. Therefore, although based on the Felder-Silverman model, this paper will expand the social interaction dimension to make it more suitable for a MOOC teaching environment.

Learning style identification

The traditional questionnaire approach to measuring learning style does not fit well in a MOOC teaching environment—primarily because factors such as the subjective consciousness of the interviewees, failure to understand questions, and learning preferences measured at a specific point would negatively affect the accuracy of the results. Therefore, an increasing number of scholars both domestically and internationally have instead been studying learning styles through automatic detection methods [26–30]. An automatic detection method is a way to detect learners' learning styles automatically by collecting real data recorded about learners by web learning platforms and adapting data mining, neural networks or simple calculation rules to the set of learning behaviors that arise in web learning contexts [31–35].

The University of Vienna analyzed learning platform data and web logs formed during learning platform sessions to recognize learners' learning behaviors [31]. In [32], the authors predicted students' learning styles by analyzing log data using BP neural networks, while [33] used a Bayesian network approach to recognize learning styles of students attending artificial intelligence online courses and discovered a significant disparity in the accuracy of predicting learning styles from different dimensions [34] combined decision tree and hidden Markov models to evaluate learning behaviors and solve difficulties in sequence data to more accurately analyze of sequential and comprehensive learning styles from the understanding dimension. A contrast test was conducted in [26] on the teaching efficiency of adaptive learning styles. The results show that the students who

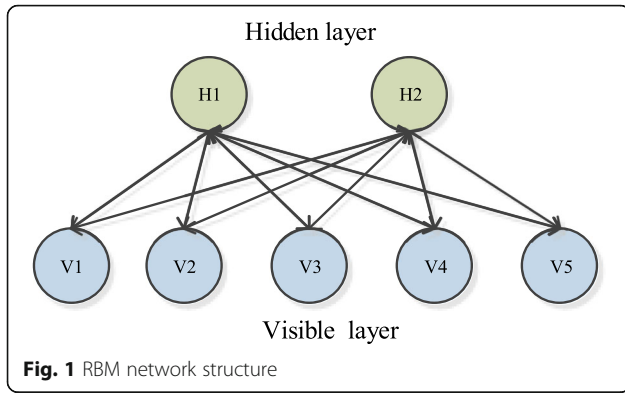
attended adaptive courses acquired a higher learning efficiency and performed better on examinations. The authors of [28] reported that standard achievement assessments could not only assess students' learning abilities but also detect individual learning characteristics and predict the results [29]. analyzed and processed the web pages visited by web-based learners to study learning styles, and [30] recorded web-based learners' learning demands and activities and explored their individualized features to study learning performance assessments.

Both the traditional learning algorithms and the back-propagation with adaptive learning rate (BPAL) network algorithm have two weak points. The first is that the raw data cannot include too many properties; the greater the number of properties are, the greater the difficulty is when composing a corresponding vector. When executing the classification algorithm, the computational complexity increases exponentially as the vector length increases. Second, the mapping relations between the properties of the raw data and learning styles cannot be too complex. Thus, these algorithms are unsuitable for complex mapping relations. The conventional methods fail to analyze and process the complex relations between the behavior data of web learning and learning styles. However, deep learning is a way to extract characteristics from vectors in a step-by-step manner, allowing more useful features to be studied by building a machine learning model with multiple hidden layers and enormous amounts of training data to promote the classification and prediction accuracy.

Deep belief networks

Deep belief networks (DBNs) were first proposed by Hinton et al. in 2006 [36]. A DBN analyzes the potential features of texts, images and voice by constructing a multilayer neural network model [37]. The training data proceeds through the network layer by layer, and each layer extracts more advanced features than the previous layers. Deep learning has substantial advantages over traditional neural network learning methods from two aspects. One is that each individual layer's training greatly promotes the training efficiency. The other is that they avoid the traditional neural networks' risk of becoming trapped in local minima under an unsupervised learning environment. The DBN model can be a combination of a multilayer RBM (unsupervised learning network) [38], BP (a supervised classifier) [39] or other prediction models.

As shown in Fig. 1, an RBM consists of two layers: a visible layer (visible units) and a hidden layer (hidden units). The connections between neurons have the following characteristics: no connections within the layer and a fully connected inner layer, where the inner layer includes both the visible layer neurons and the hidden



layer neurons. The term fully connected refers to the connection between every neuron in the visible and hidden layers. In an RBM each neuron has only two states: 0 or 1.

An RBM is an energy-based undirected generative model. Its energy function is formulated as follows for a given set of states (v, h) :

$$E_{\theta}(v, h) = -\sum_{i=1}^{n_v} b_i v_i - \sum_{j=1}^{n_h} a_j h_j - \sum_{i=1}^{n_v} \sum_{j=1}^{n_h} w_{ij} v_i h_j, \quad (1)$$

whose terms are described as follows:

$v = (v_1, v_2, \dots, v_{n_v})^T$: the state vector for the visible layer, v_i represents the state of the first i neuron in the visible layer;

$h = (h_1, h_2, \dots, h_{n_h})^T$: the state vector for the hidden layer, h_j represents the state of the first j neuron in the hidden layer;

$a = (a_1, a_2, \dots, a_{n_v})^T \in \mathbb{R}^{n_v}$: the bias vector of the visible layer, a_i indicates the bias of the i neurons in the visible layer;

$b = (b_1, b_2, \dots, b_{n_h})^T \in \mathbb{R}^{n_h}$: the bias vector of the hidden layer, b_j represents the bias of the j neurons in the visible layer;

$W = (w_{ij}) \in \mathbb{R}^{n_h \times n_v}$: the weight matrix between the hidden layer and the visible layer. w_{ij} represents the connection weights between the i neurons in the hidden layer and the j neurons in the visible layer.

The above shows the component form, but it can be rewritten in matrix form:

$$E_{\theta}(v, h) = -b^T h - a^T v - Wv h. \quad (2)$$

Using the energy function defined in Formula 1, the joint probability distribution of the state (V, H) can be given as follows:

$$P_{\theta}(v, h) = \frac{1}{Z_{\theta}} * e^{-E_{\theta}(v, h)}, \quad (3)$$

where

$$Z_{\theta} = \sum_{v, h} e^{-E_{\theta}(v, h)}. \quad (4)$$

which is a normalization factor, also called the partition function.

Through derivation, we obtain

$$P_{\theta}(h_k = 1 | v) = \text{sigmoid}\left(b_k + \sum_{i=1}^{n_v} w_{k,i} v_i\right) \quad (5)$$

$$P_{\theta}(v_k = 1 | h) = \text{sigmoid}\left(a_k + \sum_{i=1}^{n_h} w_{j,k} h_j\right) \quad (6)$$

The sigmoid function is a commonly used activation function in neural networks and is defined as follows:

$$\text{sigmoid}(x) = \left(\frac{1}{1 + e^{-x}}\right). \quad (7)$$

RBM training

Given a training sample, RBM training is intended to adjust the parameters $\theta \{W, a, b\}$ to fit the given training sample. For this parameter, the probability distribution of the corresponding RBM representation should fit the training data insofar as possible. The mathematical description is as follows:

Suppose the training sample set is

$$S = \{v^1, v^2, \dots, v^n\}, \quad (8)$$

where n_s is the number of training samples, $v^i = (v_1^i, v_2^i, \dots, v_{n_v}^i)$, $i = 1, 2, \dots, n_s$, and the samples are independent and identically distributed. The goal when training the RBM is to maximize the likelihood function.

$$\mathcal{L}_{\theta, S} = \prod_{i=1}^{n_s} P(v^i). \quad (9)$$

The product $\prod_{i=1}^{n_s} P(v^i)$ which addresses special problems can be derived from the strict monotonic property of the function $\ln x$; the maximization of $\mathcal{L}_{\theta, S}$ is equivalent to $\ln \mathcal{L}_{\theta, S}$. Therefore, the goal of training RBM is to maximize the likelihood function.

$$\ln \mathcal{L}_{\theta, S} = \ln \prod_{i=1}^{n_s} P(v^i) = \sum_{i=1}^{n_s} \ln P(v^i). \quad (10)$$

For simplicity, in the following, $\mathcal{L}_{\theta, S}$ is simplified to \mathcal{L}_S . Then, the maximum likelihood function evaluates the gradient

$$\frac{\partial \ln P(v)}{\partial \theta} = -\sum_h P(h|v) \frac{\partial E(v, h)}{\partial \theta} + \sum_{v, h} P(v, h) \frac{\partial E(v, h)}{\partial \theta}. \quad (11)$$

The maximum likelihood function of w_i, j, a_i, b_i for partial derivation is

$$\frac{\partial \ln P(v)}{\partial w_{i, j}} \approx P(h_i = 1|v) v_j - \sum_v P(v) P(h_i = 1|v) v_j \quad (12)$$

$$\frac{\partial \ln P(v)}{\partial a_i} \approx v_i - \sum_v P(v) v_i \quad (13)$$

$$\frac{\partial \ln P(v)}{\partial b_i} \approx P(h_i = 1|v) - \sum_v P(v) P(h_i = 1|v). \quad (14)$$

Contrast divergence (CD) is a standard method for training an RBM [40]. The steps in the k -step CD algorithm (abbreviated as CD- k) are simple. Specifically, approximately $\forall v \in S$ take an initial value of $v^{(0)} = v$; however, the implementation of k -step acquisition Gibbs sampling, which constitutes the first t steps ($t = 1, 2, \dots, k$), is executed as follows:

Use $P(h | v^{(t-1)})$ to sample $h^{(t-1)}$;

Use $P(v | h^{(t-1)})$ to sample v^t ;

Then, through k -step Gibbs sampling, we obtain the $v^{(k)}$ formula to approximate the corresponding desired item $\sum_v P(v)$ (or mean term), specifically as shown in (12), (13), and (14).

$$\frac{\partial \ln P(v)}{\partial w_{i, j}} \approx P(h_i = 1|v^{(0)}) v_j^{(0)} - P(h_i = 1|v^{(k)}) v_j^{(k)} \quad (15)$$

$$\frac{\partial \ln P(v)}{\partial a_i} \approx v_i^{(0)} - v_i^{(k)} \quad (16)$$

$$\frac{\partial \ln P(v)}{\partial b_i} \approx P(h_i = 1|v^{(0)}) - P(h_i = 1|v^{(k)}). \quad (17)$$

In fact, the above approximation can be regarded as using

$$CD_k(\theta, v) = -\sum_h P(h|v^{(0)}) \frac{\partial E(v^{(0)}, h)}{\partial \theta} + \sum_h P(h|v^{(k)}) \frac{\partial E(v^{(k)}, h)}{\partial \theta} \quad (18)$$

to approximate (11)

$$\frac{\partial \ln P(v)}{\partial \theta} = -\sum_h P(h|v^{(0)}) \frac{\partial E(v^{(0)}, h)}{\partial \theta} + \sum_h P(h|v^{(k)}) \frac{\partial E(v^{(k)}, h)}{\partial \theta}. \quad (19)$$

In this way, by using the stochastic gradient ascent method to maximize the log likelihood and estimating

the function value on the training data, the update criteria for each parameter can be described as follows:

$$\begin{aligned} \Delta W_{ij} &= \langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{recon} \\ &= \Delta W_{ij} \\ &\quad + P(h_i = 1|v^{(0)}) v_j^{(0)} - P(h_i = 1|v^{(k)}) v_j^{(k)} \end{aligned} \quad (20)$$

$$\Delta a_i = \langle v_i \rangle_{data} - \langle v_i \rangle_{recon} = \Delta a_i + v_i^{(0)} - v_i^{(k)} \quad (21)$$

$$\begin{aligned} \Delta b_j &= \langle h_j \rangle_{data} - \langle h_j \rangle_{recon} \\ &= \Delta b_j + P(h_i = 1|v^{(0)}) - P(h_i = 1|v^{(k)}). \end{aligned} \quad (22)$$

Learning style model in MOOCs

Learning style model

The related studies show that Felder-Silverman's learning style is suitable for use in web-based education environments, and the Solomon questionnaire that matches it has relatively high sophistication and popularity. From the perspectives of information processing, perception, input, and understanding, the learning style model classifies the learning styles into 16 categories, which are LS = {Procession, Perception, Input, Understanding}. Perception = {Sensitive/Intuitive}, Input = {Visual/Verbal}, Understanding = {Sequential/Global}. Nevertheless, the model considers only the interactions between learners and learning materials rather than the interactions between students as well as between students and teachers, and the latter is the defining characteristic in MOOC learning environments.

Opinions from domain experts were adequately considered to build a learning style model suitable for MOOC environments. Based on the Felder-Silverman model and considering the new features available in online learning environments, the learning style model used in this study adds social factors and independent/dependent/competitive types with reference to the measurement of interpersonal factors in the Readiness for Education at a Distance Indicator (READI) [41].

We built a learning style model suitable for MOOCs called MOOCLS, that comprises five dimensions and is able to provide a comprehensive description of learners in net-based learning environments. MOOCLS = {Procession, Perception, Input, Understanding, Society}, Procession = {Active/Reflective}, Perception = {Sensitive/Intuitive}, Input = {Visual/Verbal}, Understanding = {Sequential/Global}, Society = {social/solitary}. The dimensions in the model and their performance characteristics are listed in Table 2.

Collecting and mapping of e-learning behaviors

Most of the e-learning behavior data studied in the paper come from StarC (an e-learning platform) that serves as a

Table 2 The learning style model for moocs

Dimensions	Type	Description	Resources
Information Process	active	The active learners tend to improve learning efficiency with active use of study materials. They like to communicate with others and complete tasks or solve problems through teamwork.	Felder-Silverman
	reflective	Reflective learners tend to think and rethink independently. They like to complete tasks and solve problems on their own.	Felder-Silverman
Information Perception	sensitive	Sensitive learners tend to look into practical cases or some specific materials. They prefer to solve problems in standard ways.	Felder-Silverman
	intuitive	Intuitive learners prefer abstract materials. They are good at discovering the inner relationships of things. They are creative and like to solve problems in creative ways.	Felder-Silverman
Information Input	verbal	Verbal learners are good at acquiring information from text and communication.	Felder-Silverman
	visual	Visual learners prefer to get information from picture and videos.	Felder-Silverman
Information Understanding	sequential	Sequential learners prefer to study in a logical and sequential order. They like to focus on small-sized materials.	Felder-Silverman
	global	Global learners prefer autonomous learning. They like to master the general structure before beginning an in-depth study of specific points. They are non-sequential learners inclined to large-paced study.	Felder-Silverman
Social Interaction	social	Social learners like to communicate with others. They are inclined to complete tasks through teamwork.	READI
	solitary	Solitary learners like to think independently. They prefer self-analysis and rethinking. They are willing to make plans and set goals.	READI

fundamental education tool. StarC has made great contributions to resource sharing nationwide. Two hundred MOOC resources were made available online from March 2015 to August 2016, and 85,345 accounts were opened, 32,987 of which were active. The platform provides various functions for users to learn and communicate, such as knowledge maps, videos, tests, wikis, instant messaging, etc. All the data produced in the learning process are recorded in the system and constitute a set (F) of users' learning behavior features. We then map these learning behaviors according to the learning styles defined by the MOOCLS model and the experiences of e-learning experts to predict which learning style model the user belongs to based on the user's learning behavior features, as is shown in the Table 3.

Active/reflective learning style and e-learning behavior

A learning style can be recognized by users' learning behaviors on forums, wikis, and homework in the active/

reflective dimension. Generally, active learners tend to have more posts, post replies and post views in the forum, while reflective learners prefer to think in private and tend to read posts instead of posting. This approach can also be used to recognize active/reflective learners. A wiki is an important feature through which teachers can instruct students by providing key information. When teachers make an entry, students can revise and extend its definition, connotation and denotation, in the course of which each change to the wiki view and all the wiki edits are recorded. Active learners are inclined to try more wiki functions and continually revise the entry. In contrast, reflective learners analyze each question carefully and look up wiki entries more than they revise or update them. Regarding homework, data concerning the time an assignment is submitted before the homework deadline and the frequency of completed homework problems are collected.

Table 3 Matching of learning style dimension and learning behavior

Active/contemplative	sensory/intuitive	visual/verbal	sequence/integrated	social /alone
post	outline_view	video_view	outline_view	Post
post_replies	textbook_view	video_play_time	button_click	forum_visits
post_view	video_view	textbook_view	table of content	post_replies
wiki_edit	Hd	K-Map_view	introduction_view	post_views
wiki_view	Review hd	wiki_view		post_replied_by
time_before_hw_deadline	K-Map_view	wiki_edit		post_viewed_by
fr_hw_problem	introduction_view	introduction_view		self_hw_views
				chart_replies_by
				charts

Sensitive/intuitive learning style and E-learning behavior

The number of users' video views, textbook views, outline views, K-Map views, homework dealing, homework review, and introduction view. Sensitive learners are more likely to review their homework carefully, spend more time on doing homework and deal with practical cases such as video, textbook and K-map. Intuitive learners are more creative. They prefer abstract ideas and would like to spend more time on them. Therefore, we can recognize an intuitive learning style by the visit frequency for such abstract resources as K-maps, outline views, and course introductions.

Visual/verbal learning style and e-learning behavior

The users' video play time is recorded as well as the frequency of video views, K-map views, textbook views, wiki views, wiki edits and introduction views. Verbal learners are more inclined to read text messages such as course book listings, course introductions, wiki edits and views, while visual learners prefer materials such as graphs, tables and videos. As a result, the time they spend on video and the frequency with which they check K-maps tend to be larger.

Sequential/global learning style and e-learning behavior

Outline views, introduction views, button clicks and tables of contents views are collected to help determine whether the learning style is sequential or global. Global learners normally visit comprehensive introduction pages, such as outline views and introduction views, while sequential learners visit fewer such pages. Sequential learners prefer to click buttons, while global learners prefer to navigate by clicking catalogue links.

Social/solitary learning style and e-learning behavior

From a social interaction perspective, learners can be divided into social and solitary types. These two types are distinguishable by number of posts, forum visits, replies, post views, post replied to, posts viewed by, self-homework views, charts, and chart replies. Social learners like to communicate with others, while solitary learners prefer to study alone. Therefore, social learners tend to be more active in discussion areas. They communicate with others frequently by posting, replying to posts and using built-in chat tools. They also maintain good relationships with others through their posts, post replies and chats. Solitary learners are much less active than their social counterparts and more likely to focus on their own homework. These learners have fewer posts, post replied to, viewed by, and fewer chat messages.

Description of the learning style test

The learning style test recognizes a learner's learning style based on learning behavior features. It can be formulated as follows:

$$F \Rightarrow LS$$

where $LS \in \{00000, 00001, \dots, 11111\}$ and the user's learning style belongs to one of the learning style dimensions. For example, a 00001 means that a user's learning style could be recognized by fundamental features and properties from the following dimensions: information processing as Active, information perception as Sensitive, information input as Visual, information understanding as Sequential, and social interaction as Social.

Learning style detection based on deep belief neural networks

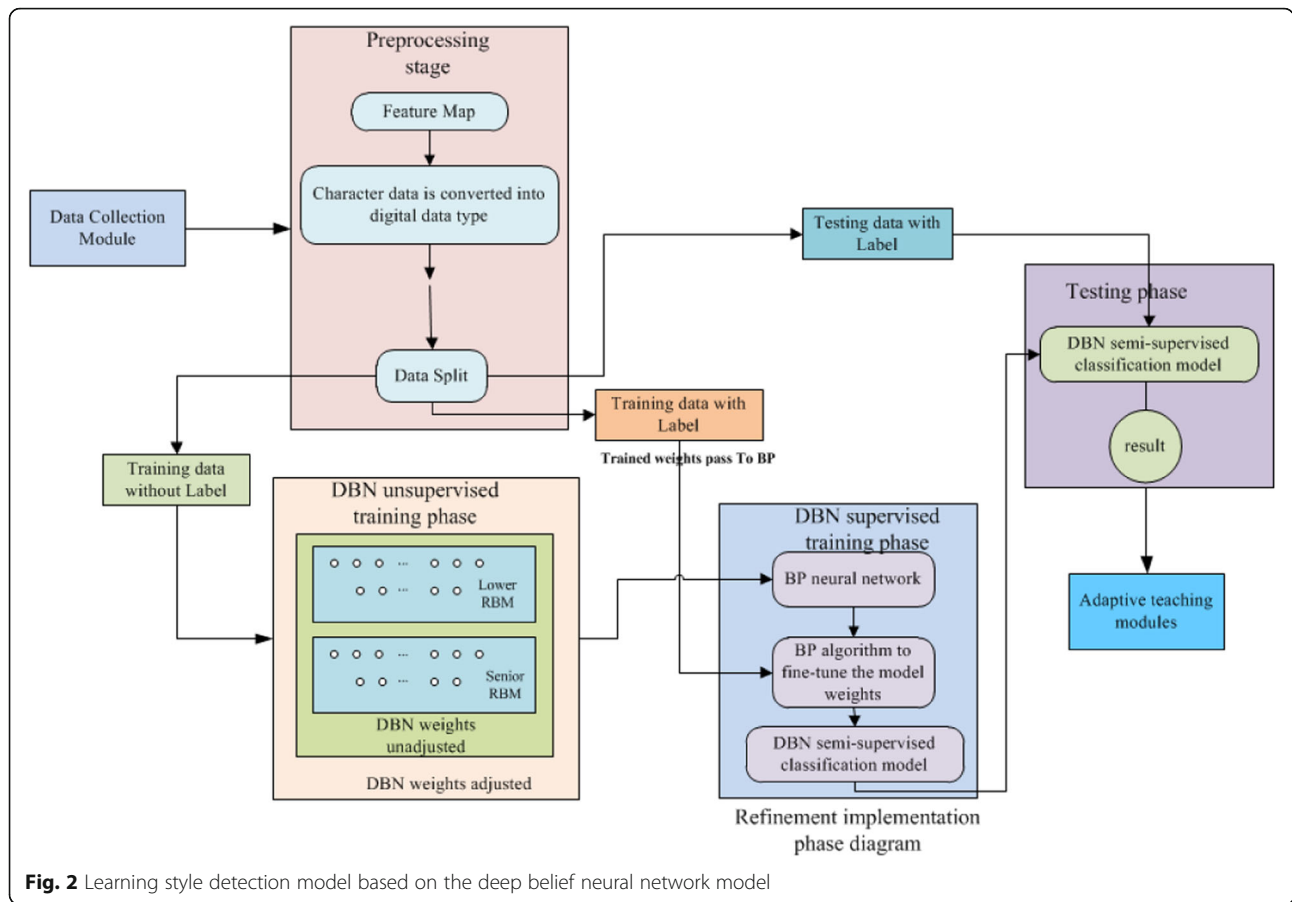
Predicting and identifying users' learning styles by analyzing a large number of e-learning behaviors is a complicated process in light of the complicated interactions in learning and the high-dimensionality of the data caused by remarkable differences. The traditional classification approaches cannot identify users' learning styles effectively and accurately. In this paper, the deep learning DBN model is used to identify the users' learning styles. Although deep learning in other areas has achieved great success, it has still not been widely applied in the field of educational technology.

DBNLS model

The DBNLS (DBN for MOOCLS) model is the core component of identifying learners' learning styles. DBNLS consists of a multilayer RBM and a BP network layer: the RBM implements feature extraction and the BP assists in fine-tuning the DBN and predicting learning styles. The core procedure can be formulated as follows. The first step is to obtain set of original learning behavior data by collecting and pretreating network behavior data. After training the DBN model with the extracted set, the features and weights of the e-learning activities are determined. Some of the learning styles are recognized through questionnaires and set as the standard dataset. Thus, the DBN model is fine-tuned and trained by adding the BP to fit the DBNLS to identify the learning style model in e-learning activities. The details are shown in Fig. 2.

The steps of DBNLS are as follows:

Step 1: Data preprocessing: Eliminate logged dates that deviate from the normal values, and normalize the dates from the characteristic value of learning



behavior according to the formula $x^* = \frac{x - \min}{\max - \min}$ (where min represents the minimum value with a single characteristic value, and max represents the maximum value with a single characteristic value) to obtain the initial dataset DB_{all} .

Step 2: Measure part of the active users' learning styles through the Index of Learning Style Questionnaire. The scale contains 44 questions. Each question is based on different learning style dimensions. With four dimensions in total, each dimension includes 11 questions, and each question corresponds to 2 alternative answers. Learning styles can be tested accurately using this scale. By labeling the online learning styles with the corresponding e-learning activity data from the questionnaire, we finally obtained a labeled activity set DB_{label} . We then divide the dataset into a training set DB_{train} and a testing set DB_{test} .

Step 3: We used the data from DB_{train} to train the DBN. BP received supervised training based on DB_{train} after the unsupervised DBN training.

Step 4: We used DB_{test} to test the performance and effect of the DBNLS model.

DBNLS model training

After training the DBNLS model the DBN can extract the behavioral characteristics of network learning behaviors and identify the learning styles of network learning behavior. The process of training the DBNLS learning style model is to adjust the parameters of each layer of the RBM to minimize the reconstruction error of the DBNLS model output. The model can abstract and extract the most original features of the sample data. The RBM parameter set of each layer of the DBNLS model

$$\text{in this paper is } \theta = \begin{cases} W = W + \eta \langle \frac{1}{n_s} \Delta W \rangle \\ a = a + \eta \langle \frac{1}{n_s} \Delta a \rangle \\ b = b + \eta \langle \frac{1}{n_s} \Delta b \rangle \end{cases}, \text{ where } W, a$$

and b are the weight matrix, the bias vector of the visible layer to the hidden layer, and the bias vector of the hidden layer to the visible layer, respectively. The DBNL model training is performed layer by layer, and the output of each RBM layer is used as the

input to the next layer. The gradient descent method is used during the training of each layer of the RBM to maximize the likelihood function of the RBM output. When the gradient descent reaches a reasonable value (or the training times reach a preset value), feature extraction is complete, and the DBMLS model has obtained a complete representation of the training data. At this point, the training process ends.

To perform learning style detection, a known analysis test sequence is input as a parameter into the DBMLS model, and the classifier output of the DBNLS model is used as the prediction result.

According to the introduction of the DBN model, DBN training can be simplified into several independent RBM training sessions. The feature vector data are used as the input to fully train the first RBM; then, we fix the weight matrices and bias of the first RBM. The hidden layer neurons are set as the input vector to the second RBM. After fully training the second RBM, we stack the second RBM on top of the first RBM. The steps above are repeated until the entire training process is completed for all the RBM models. Finally, the hidden layer neuron state of the last RBM is set as the input to the BP neural network to supervise it efficiently.

DBNLS training can be roughly divided into two phases. The first phase is to initialize the model parameters, in which the given vector training sample $S(|S| = n_s)$, training cycle T , learning rate η , and k algorithm in CD- k is set before specifying the unit numbers n_v, m_h of the visible and hidden layers and initializing the vectors a, b as well as the weighting matrix W . Then, the training process is as follows:

```

FOR iter = 1, 2, ..., T DO
{
1. transfer CD - k, generate  $\Delta W, \Delta a, \Delta b$ .
2. refresh the parameters:  $W = W + \eta \langle \frac{1}{n_s} \Delta W \rangle$ ,
 $a = a + \eta \langle \frac{1}{n_s} \Delta a \rangle, b = b + \eta \langle \frac{1}{n_s} \Delta b \rangle$ .
}
    
```

To improve the efficiency and effectiveness of DBNLS training, the important parameters have been processed. The DBN classification algorithm is the core part of DBNLS; it determines the learning style identification performance of the whole DBNLS. The DBN classification algorithm is shown in Fig. 3.

```

DBN classification algorithm
Step1: Initialize the Gibbs sampling time k and the learning rate  $\epsilon$  in the
CD - k algorithm, the number of iterations epochs,
the training samples dataset, batch_size,
the number of visible layer units  $n_{ms}$ ,
the number of hidden layers and the number of hidden layer units,
DBN model parameters  $W, b, c$ , training period T.
Step2: FOR iter = 1, 2, ..., T DO
{
one). Transfer CD - k, generate  $\Delta W, \Delta b, \Delta c$ .
two). Refresh parameter:  $W = W + \epsilon \langle \frac{1}{dataset} \Delta W \rangle, b$ 
 $= b + \epsilon \langle \frac{1}{dataset} \Delta b \rangle, c = c + \epsilon \langle \frac{1}{dataset} \Delta c \rangle$ .
}
    
```

Small batch data During the course of the FOR loop in the second training step, the training sample set S is applied to generate the data $\Delta \theta'$ ($\Delta W, \Delta a, \Delta b$) one time through superposition of the partial derivative in various samples. By dividing set S into tens or hundreds of small batches and calculating them individually, the training efficiency can be improved considerably, and the advantages of GPUs and MATLAB matrix multiplication can be exploited. To address the problem that the gradient of the parameter differs for the different capacities of small batch data, we can divide the total gradient by the size of the small batch data to obtain the average gradient during the parameter updating process:

$$\theta' = \theta' + \eta \cdot \frac{\Delta \theta'}{B}. \tag{23}$$

In (23), B denotes the small batch data capacity. Every small batch of data should include samples from each class, and the samples function better when they are of equal size. The size of B is set to multiples of the number 10.

Learning rate The study of η plays a key role in the process of updating neural network parameters because its size directly affects the algorithm training performance. A large η improves the rate of convergence; however, it also can cause the algorithm to become unstable. In contrast, a smaller η improves algorithm stability, but slows down convergence [42]. In this paper, a momentum term was added to update the parameters to promote both the convergence speed and the performance equilibrium.

username	user_id	action	event_type	time	ip	context_course_id	source
xuxiao	5	dashboard view	/dashboard	2016-06-23T11:10:14.238.81	10.144.238.81		server
xuxiao	5	courses view	/courses	2016-06-23T11:10:14.238.81	10.144.238.81		server
xuxiao	5	courses view	courses/CCNU/HCO1/20150410/about	2016-06-23T11:10:14.238.81	10.144.238.81	CCNU/HCO1/20150410	server
xuxiao	5	courses enrollment	/change_enrollment	2016-06-23T11:10:14.238.81	10.144.238.81		server
xuxiao	5	courses enrollment	edx.course.enrollment.activated	2016-06-23T11:10:14.238.81	10.144.238.81	CCNU/HCO1/20150410	server
xuxiao	5	dashboard view	/dashboard	2016-06-23T11:10:14.238.81	10.144.238.81		server
xuxiao	5	courses view	courses/CCNU/CCNU_7/20150605/about	2016-06-23T11:10:14.238.81	10.144.238.81	CCNU/CCNU_7/20150605	server
xuxiao	5	courses view	/about	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	server
xuxiao	5	courses enrollment	/change_enrollment	2016-06-23T11:10:14.238.81	10.144.238.81		server
xuxiao	5	courses enrollment	edx.course.enrollment.activated	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	server
xuxiao	5	dashboard view	/dashboard	2016-06-23T11:10:14.238.81	10.144.238.81		server
xuxiao	5	courses view	/courseware	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	server
xuxiao	5	courses view	/courseware/	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	server
xuxiao	5	video load	load_video	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	browser
xuxiao	5	video play	play_video	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	browser
xuxiao	5	video pause	pause_video	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	browser
xuxiao	5	video stop	stop_video	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	browser
xuxiao	5	courses view	/courseware/	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	server
xuxiao	5	courses view	page_close	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	browser
xuxiao	5	video load	load_video	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	browser
xuxiao	5	video play	play_video	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	browser
xuxiao	5	video pause	pause_video	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	browser
xuxiao	5	video stop	stop_video	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	browser
xuxiao	5	courses view	/info	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	server
xuxiao	5	courses view	page_close	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	browser
xuxiao	5	discussion view	/discussion/forum	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	server
xuxiao	5	discussion reply	/discussion/threads//reply	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	server
xuxiao	5	discussion view	/discussion/forum/	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	server
xuxiao	5	discussion create	/discussion/create	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	server
xuxiao	5	discussion view	/discussion/forum//inline	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	server
xuxiao	5	discussion comments	/discussion/comments//upvote	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	server
xuxiao	5	wiki view	/course_wiki	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	server
xuxiao	5	wiki view	/wiki/CCNU.MATH001.2016_3_17/	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	server
xuxiao	5	wiki edit	iki/CCNU.MATH001.2016_3_17/_edit	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	server
xuxiao	5	wiki edit	iki/CCNU.MATH001.2016_3_17/_edit	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	server
xuxiao	5	wiki edit	/wiki/CCNU.MATH001.2016_3_17/	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	server
xuxiao	5	courses view	/progress	2016-06-23T11:10:14.238.81	10.144.238.81	XNU/MATH001/2016_3_1	server
xuxiao	5		/logout	2016-06-23T11:10:14.238.81	10.144.238.81		server
xuxiao	5		/	2016-06-23T11:10:14.238.81	10.144.238.81		server

Fig. 3 Sample weblog data

$$\theta' = \lambda\theta' + \eta \cdot \Delta\theta'. \quad (24)$$

In (24), λ is the momentum learning rate. The addition of λ causes parameter updating that is not totally in accord with the current sample gradient direction; instead, it is combined with the previous parameter change direction.

Number of hidden layer units Due to the lack of theoretical support, it is difficult to determine the RBM numbers of hidden layer units in DBN. Generally, the units can be determined according to the number of neurons in the input and output layers. If the hidden layer units are quite small, the ability of the RBM model to obtain information will be too weak to sum and reflect the characteristics of the training set. When there are too many hidden layer units, training will require more time and cause overfitting problems.

We determined the number of hidden layer units through experiments in this paper. The first step is to set the number of hidden layer units to a smaller value and then train on the samples. The reasonability of the results are judged according to the number of hidden layer units. When the training effect is poor, we can increase the number of hidden layer units until the results reach reasonable values.

The prediction of learning style

During the DBNL model training, the expressions of different network behavior characteristics are established. The MOOCLS learning style model is established in the form of parameter set $\theta = \{W, a, b\}$. When another set of network learning behavior characteristics are input, the

trained DBNLS can quickly identify the learning style the input network behavior characteristics belong to.

Experimental

We conducted a set of experiments to set the parameters and examine the effectiveness of our proposed DBNLS approach for identifying learning styles in terms of accuracy and quality.

The experimental environment

StarC is a free MOOC (Massive Open Online Course) platform open to K-12 education set up by the Central China Normal University. It is a research exchange and results application platform of the National Digital Learning Project and Technology Center, and it works to deliver free learning courses designed by first-rate teachers in our country to provide high-quality teaching materials for the majority of young students and to help educators establish an effective online learning community. Educators use the platform to design and release online courses and online teaching evaluations. During the process of online learning on StarC, interactive data of network behavior, such as independent learning journals, test performances and learners' essential information are produced and stored.

This study is based on the curriculum reform experiment carried out by Suzhou Educational Informatization Reform. Our candidates are more than 500's year students from 10 classes in Suzhou Junior High School and Zhenhua Junior High School. All the students are eager to learn. Before the online learning courses formally started, some targeted training was delivered by the teachers to help the students understand how to use an

online learning platform to conduct independent learning.

Data preparation

In addition to the behavior data collected by the StarC platform, we also collected offline questionnaires. We could associate the offline subjects’ names and student numbers using an online learning behaviors dataset. The questionnaires were mainly used as the class labels for the training dataset. The learner behaviors were trained through the DBNLS model, and finally, the classification results were obtained and compared with the class labels and used to assess the accuracy of the DBNLS model.

The questionnaire data

The questionnaire used in this study is divided into two parts. In the first part, we designed the offline questionnaire of this research based on the acquisition and mapping of online learning behaviors, according to the Index of Learning Styles quantization table [11], which describes a learning style model and refers to the maturity scale of each dimension in the DBNLS model. Van Zwanenberg’s research results indicate that the Index of Learning Styles scales have good reliability and validity [42]. The online learning styles model constructed for this research used all the items of the Index of Learning Styles scales to ensure both the consistency and integrity of the scales. In addition, it adds the social interdimension—11 questions for the social dimension of learning style. The second part of the questionnaire captures personal information. To enable the offline respondents to adapt the online data analysis objects, this part seeks to acquire personal information such as the testee’s real name, student number, and StarC platform user account, enabling one-to-one correspondence between the students’ learning style information and their online learning behavior data.

The scoring formula for this scale has three steps. The first step is to classify a and b answers to questions corresponding to each dimension in the questionnaire. Then, the second step is to determine the total number of answers to the questions in each dimension. Finally, the last step is to compare the larger sum with the smaller sum and combine the difference value with the letter of the larger sum. The final result is added to the last list. (The difference value means indicate the tendency level of the learning style; the letter denotes

different types of learning style). The design for measuring the items for social contact and solitude in the social dimension referred to learning-styles-online.com [43]. Each question adopts a three-level scale for measurement, where a “0” means unmatched and a “1” means matched. The forum of the questionnaire is shown in the Table 4.

To explain the class label role of the questionnaire in the DBNLS model, the measuring items in the questionnaire needed to be designed. There are 44 questions in the questionnaire corresponding to the Index of Learning Styles and 11 questions for each dimension for a total of 55 questions. The specific questionnaire question content refers to [43, 44] and the Learning Style Self-Test Questionnaire, as shown in the supplementary material below. With the assistance of the head teacher in the subject class and under the assumption that the questionnaire data matched the platform data, all the questionnaires were effective. To ensure that all the questionnaires were repeatedly confirmed and examined, 520 questionnaires were recycled, of which 508 are valid.

According to the Table 5, the measurement result “9a” in the information processing dimension means that the subject belongs to the “active type” and has a strong tendency level in learning style. The measuring result “5b” in the information perception dimension means that the subject belongs to the “conscious type” and has an average tendency level of learning style. The information input and information comprehension dimensions can also be analyzed in this way. For the social dimension, a 1 means the subject belongs to the “social level” learning style. In summary, the measurement result is =10,101 = 21. The online learning behavior data are used to obtain the corresponding learning style by comparing the data trained through the DBNLS model and the questionnaire results.

Network behavior data

The data produced by the learners during the whole learning process is stored in a MySQL database, as Mongo data and as a journal file. The user table is used to store students’ personal information, including birth date, grade, gender, learner types, while the journal file is used to record the online learning student interactive data (the students’ personal data was hidden). As shown in Fig. 3, this table reflect only part of the learning data

Table 4 Index of learning styles

Active type/reflective type			Feeling type/conscious type			Verbal type/visual type			Sequence type/synthetic type		
Questions 1	a	b	Questions 2	a	b	Questions 3	a	b	Questions 4	a	b
...				
total			total			total			total		

Table 5 Measurement results of the subjects

Username	Measurement results
xuxuan	9a,5b,11a,7b,1
...	...
yangyin	5a,7b,a,b,0

from one student. Typically, there will be hundreds of records after a student finishes an entire course.

The processing and choice of data have a large effect on the test learning style results; thus, these aspects are core steps in the learning style testing process. For this study, we preprocessed the log files and extracted the core property information, course information, forum information, wiki information and assignment submission information.

After communicating with four veteran teachers who participated in the experiment for this study, we confirmed the five types of learning style dimensions and the corresponding Internet learning behavior, which is shown in Table 2. Accordingly, we collected the number of visits, posts, readers, repliers, posts replied to and post read from the forum aspect; the compilation of wiki entries changed, the number of changed wiki entries and the number of read wiki entries from the wiki aspect; the advance times of assignment submissions, the number of exercises, the times the assignments were performed, checking and examining homework from the homework aspect; the number of message sent, messages replied to and message replies from the instant messaging aspect; the durations and times when video was played from the video aspect; and 24 other parameter values related to other aspects such as the number of visits to the course outline, course abstract, course catalog, course teaching materials,

knowledge map, the number of navigational clicks to move to the “previous/next page” and so on. These are considered to be core characteristic attributes.

We also obtained core personal information, such as the learners’ names and student IDs by linking the student ID to the userInfo table in the “school in CCNU” database, as shown in Fig. 4.

The preprocessing of data

In this paper, the learning style data are composed of two parts: training data and test data. The training data are divided into 30,000 unlabeled training data and 6150 labeled training data. In total, there are more than 30,000 records containing 32 types of learning styles. The test data comprise more than 2050 records.

The network learning behavior data, as discussed in the previous section, come from all the learners’ network behavior in the 10 courses of the “normal school” platform. These learners consist of two groups: 820 students who completed the learning style questionnaire and many other online learners who did not complete the learning style questionnaire. Based on the student IDs and usernames, we associated the learning style information with the students’ learning style questionnaires and the unlabeled network learning behavior data. Thus, we extracted 8200 different Internet learning behaviors. Then, we integrated the questionnaire data into the network learning behavior to form the 8200 labeled learning style data, as shown in Fig. 5, forming learning style tags that include _type information.

To improve the quality of the data. We eliminated some invalid learning behavior data. Then, we randomly divided the 8200 labeled learning style data into two groups at a 3:1 ratio: 6150 data items were

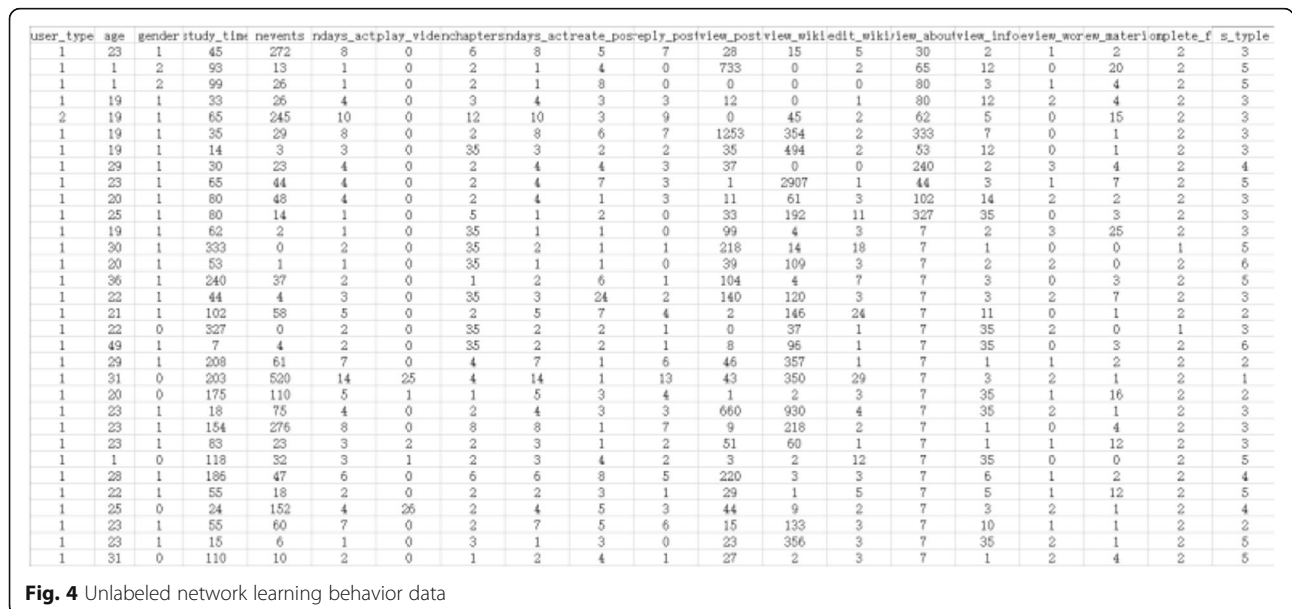


Fig. 4 Unlabeled network learning behavior data

user_type	age	gender	study_time	revents	ndays_act	play_vid	chapters	ndays_act	create_pos	reply_pos	view_post	view_wiki	edit_wiki	view_about	view_info	view_wor	view_mater	complete_f	s_type
1	23	1	45	272	8	0	6	8	5	7	28	15	5	30	2	1	2	2	3
1	1	2	93	13	1	0	2	1	4	0	733	0	2	65	12	0	20	2	5
1	1	2	99	26	1	0	2	1	8	0	0	0	0	80	3	1	4	2	5
1	19	1	33	26	4	0	3	4	3	3	12	0	1	80	12	2	4	2	3
2	19	1	65	245	10	0	12	10	3	9	0	45	2	62	5	0	15	2	3
1	19	1	35	29	8	0	2	8	6	7	1253	354	2	333	7	0	1	2	3
1	19	1	14	3	3	0	35	3	2	2	35	494	2	53	12	0	1	2	3
1	29	1	30	23	4	0	2	4	4	3	37	0	0	240	2	3	4	2	4
1	23	1	65	44	4	0	2	4	7	3	1	2907	1	44	3	1	7	2	5
1	20	1	80	48	4	0	2	4	1	3	11	61	3	102	14	2	2	2	3
1	25	1	80	14	1	0	5	1	2	0	33	192	11	327	35	0	3	2	3
1	19	1	62	2	1	0	35	1	1	0	99	4	3	7	2	3	25	2	3
1	30	1	333	0	2	0	35	2	1	1	218	14	18	7	1	0	0	1	5
1	20	1	53	1	1	0	35	1	1	0	39	109	3	7	2	2	0	2	6
1	36	1	240	37	2	0	1	2	6	1	104	4	7	7	3	0	3	2	5
1	22	1	44	4	3	0	35	3	24	2	140	120	3	7	3	2	7	2	3
1	21	1	102	58	5	0	2	5	7	4	2	146	24	7	11	0	1	2	2
1	22	0	327	0	2	0	35	2	2	1	0	37	1	7	35	2	0	1	3
1	49	1	7	4	2	0	35	2	2	1	8	96	1	7	35	0	3	2	6
1	29	1	208	61	7	0	4	7	1	6	46	357	1	7	1	1	2	2	2
1	31	0	203	520	14	25	4	14	1	13	43	350	29	7	3	2	1	2	1
1	20	0	175	110	5	1	1	5	3	4	1	2	3	7	35	1	16	2	2
1	23	1	18	75	4	0	2	4	3	3	660	930	4	7	35	2	1	2	3
1	23	1	154	276	8	0	8	8	1	7	9	218	2	7	1	0	4	2	3
1	23	1	83	23	3	2	2	3	1	2	51	60	1	7	1	1	12	2	3
1	1	0	118	32	3	1	2	3	4	2	3	2	12	7	35	0	0	2	5
1	28	1	186	47	6	0	6	6	8	5	220	3	3	7	6	1	2	2	4
1	22	1	55	18	2	0	2	2	3	1	29	1	5	7	5	1	12	2	5
1	25	0	24	152	4	26	2	4	5	3	44	9	2	7	3	2	1	2	4
1	23	1	55	60	7	0	2	7	5	6	15	133	3	7	10	1	1	2	2
1	23	1	15	6	1	0	3	1	3	0	23	356	3	7	35	2	1	2	5
1	31	0	110	10	2	0	1	2	4	1	27	2	3	7	1	2	4	2	5

Fig. 5 Tagged data

used as labeled training data for supervised model training; these data are denoted as train_label. The remaining 2050 data were used as a test datasets, denoted as test_label, to test the effects of the DBNLS model. Finally, the train_label test_label datasets were used to train test the DBNLS model.

Another part of the unlabeled data was collected from unknown learning style students who had also taken the 10 courses on the “normal school” platform. There are approximately 30,000 records in this part. An example of the form of the data is shown in Fig. 6. Compared with the labeled learning style

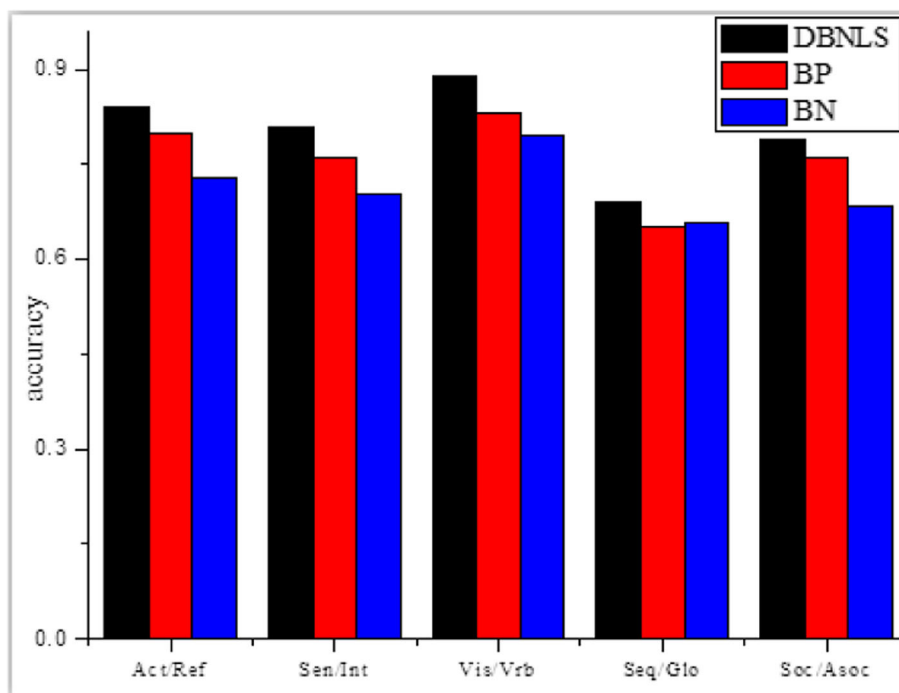


Fig. 6 Performance comparisons of different dimensions

data, unlabeled data did not include the last line (the types of learning styles) because their learning styles were unknown. We used this part of the learning style data plan for unsupervised DBNLS model training.

DBNLS classifier evaluation

To evaluate the effectiveness of the DBNLS model, we chose the ILS Standard Index of Learning Styles to collect the student learning style data of students. As explained before, these students were attending specific courses in the StarC Course Platform; thus, their online learning style behavior data was collected. We correlated the learning style data of each student with their online learning behaviors to form the learning style detection data, as shown in Table 6. We separated the data into two parts for training and testing the DBNLS model classifier.

Table 6 demonstrates the results of the DBNLS model. The column is the ground truth of the learning styles of each student, the row headers indicate the learning style predictions by DBNLS, and the values in the table show the instances numbers in which a student with the *i* learning style was falsely predicted as the *j* learning style, here $1 < i < = 32$ and $1 < j < = 32$:

$$D_p = \sum_{i=1}^n (TN + TP) / n \tag{25}$$

To further evaluate the DBNLS model, we decided to compute the classification accuracy. As is shown in the Table 7, the accuracy scores for the procession, perception, input, understanding, and sociality dimensions were 84%, 81%, 89%, 69%, 79%, respectively as is shown in the Table 8. TN and TP represent the true negative instances and true positive instances, respectively, and *n* is the total number of items in the dataset.

In addition, we also applied the BP neural network to perform the same classification experiments as the DBNLS model. The result show that DBNLS clearly performs better than the traditional BP neural network. García P used a Bayesian network to classify four types of learning styles; the average accuracy of DBNLS was also better than that achieved by BP.

We found that compared to the traditional BP neural network model, the standard deviation of the DBMS

Table 7 The classifier results

Model	Act/Ref	Sen/Int	Vis/Vrb	Seq/Glo	Soc/Asoc
DBNLS	84%	81%	89%	69%	79%
BP	80%	76%	83%	65%	76%
BN	72.73%	70.15%	79.54%	65.91%	

model’s accuracy is 0.1033, which is quite stable. As is shown in the Table 9, the standard deviation of the accuracy of the BP model is 0.8019, which is relatively large and indicates model uncertainty. Because the DBNLS model uses unsupervised pretraining of the RBM, the RBM is pretrained, and its weight matrix is adjusted to near-optimum values. Therefore, the stability of DBNLS is greater than that of the BP neural network model.

We further analyzed each data dimension and found that the detection rate for sequential/global learning styles reflected a significantly lower value in the understanding dimension. By analyzing the student logs, we found that most students read the whole course—that is, there was no difference among the syllabus and course profiles in terms of the number of visits. Moreover, most students did not skip units and read the entire unit. This phenomenon can be observed in the results obtained for the understanding dimension, where no global learners were discovered.

To summarize the study results, we can conclude that DBNLS performs better than do traditional methods for complex learning style classifications that consider more properties.

Conclusions

The differences between learners are determined according to each learner’s previous knowledge of the subject matter, their learning style, learning characteristics, preferences and goals. Being able to accurately classify online students through networking learning behavior analysis could enable more effective personalized support for students seeking information and learning in an online context. This paper proposes a learning style identification and classification method based on a DBN, called DBNLS. First, we built a student learning style model for MOOCS based on the existing Felder-Silverman model and expert experience. According to the proposed learning style model and its implications, we evaluated the correlations between network learning

Table 6 The results of the experiment

Prediction Actual	1	2	...	32
1	210	15	...	14
2	10	180	...	8
...
32	14	6	...	200

Table 8 Confusion matrix

Procession	Reflective	Predicted		Total
		Negative	Positive	
Active				
Actual	Negative	TN	FP	TN + FP
	Positive	FN	TP	FN + TP
Total		TN + FN	FP + TP	TN + TP + FN + TP

Table 9 The stability of the results

Number	DBNLS	BP
1	87.7%	80.2%
2	87.5%	79.1%
3	87.6%	81.3%
4	87.6%	79.6%
5	87.4%	79.9%
6	87.6%	80.8%
standard deviation	0.1033	0.8019

behaviors to learning style within individual sessions; then, we used the deep learning DBN model to learn those learning style features and model learning styles to classify students accurately.

We conducted several experiments on an actual educational dataset to verify the proposed method. First, some behavioral patterns in the various learning style dimensions were determined by conducting an experiment with learning content based on the ILS theory by Felder and Soloman (1996) and Readiness for Education At a Distance Indicator. We collected actual network learning behavior data from a MOOC learning platform and labeled the collected learning style data using ILS theory to analyze and classify them. Then, we utilized those data to train our DBNLS model. Compared with traditional classification methods (BP and BN), the proposed method achieves good accuracy and performance. However, the accuracy of the predictions regarding the understanding learning style dimensions is not perfect.

In the future, we plan to obtain more learning behaviors to improve the accuracy of predictions regarding the understanding learning style dimensions. Then, we plan to apply the DBNLS to other personalized learning systems, such as personalized recommender systems and personalized learning navigation.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s13677-020-00165-y>.

Additional file 1.

Abbreviations

DBN: Deep Belief Network; RBM: Restricted Boltzmann machines; LS: Learning style; DBNLS: DBN for MOOCLS; MOOCLS: MOOC learning style model

Acknowledgments

We are grateful to the volunteers for capturing the data.

Authors' contributions

Hao Zhang provided the idea for the paper and wrote the paper; Tao Huang and Sanya Liu helped to revise the paper; Hao Yin, Jia Li, and Yu Xia implemented the model for this paper. Huali Yang was responsible for the questionnaires. The authors read and approved the final manuscript.

Authors' information

Hao Zhang, Tao Huang, Sanya Liu, Huali Yang, Jia Li and Yu Xia are employees of the National Engineering Laboratory for Educational Big Data Central China Normal University and National Engineering Research Center for E-Learning, Central China Normal.

Hao Yin is the employee of College of Information Engineering, Shenyang University, Shenyang, China.

Funding

This research is supported by the National Key Research and Development Program of China (2017YFB1401300, 2017YFB1401304, and 2017YFB1401303), the National Natural Science Foundation of China (61702211, 61977033 and L1724007), the Hubei Provincial Science and Technology Program of China (2017AKA191), and the Fundamental Research Funds for the Central Universities (CCNU19QD004 and CCNU17GF0002).

Availability of data and materials

The datasets supporting the results of this article are included within the article.

Competing interests

The authors declare that they have no competing interests.

Author details

¹National Engineering Laboratory for Educational Big Data Central China Normal University, Wuhan, China. ²National Engineering Research Center for E-Learning, Central China Normal University, Wuhan, China. ³College of Information Engineering, Shenyang University, Shenyang, China.

Received: 16 July 2019 Accepted: 13 March 2020

Published online: 18 May 2020

References

- McAfee A, Brynjolfsson E, Davenport TH et al (2012) Big data. The management revolution. *Harv Bus Rev* 90(10):61–67
- Margaryan A, Bianco M, Littlejohn A (2015) Instructional quality of massive open online courses (MOOCs). *Comput Educ* 80:77–83
- Muñoz-Merino PJ, Valiente JAR, Kloos CD (2013) Inferring higher level learning information from low level data for the Khan Academy platform. In: Proceedings of the third international conference on learning analytics and knowledge. ACM, pp 112–116
- Kong SC (2014) Developing information literacy and critical thinking skills through domain knowledge learning in digital classrooms: an experience of practicing flipped classroom strategy. *Comput Educ* 78:160–173
- DeBoer J et al (2014) Changing “course” reconceptualizing educational variables for massive open online courses. *Educ Res* 43(2):74–84
- Lin CF et al (2013) Data mining for providing a personalized learning path in creativity: an application of decision trees. *Comput Educ* 68:199–210
- Chu H-C (2014) Potential negative effects of mobile learning on students' learning achievement and cognitive load—a format assessment perspective. *J Educ Technol Soc* 17:322–344
- Zheng Y (2015) Real difficulties in the “teaching students in accordance with their aptitude”. *Res Educ Dev* 6:10
- Siemens G, Long P (2011) Penetrating the fog: analytics in learning and education. *Educ Rev* 46(5):30
- Picciano AG (2012) The evolution of big data and learning analytics in American higher education. *J Asynchronous Learn Net* 16(3):9–20
- Felder RM, Silverman LK (1988) Learning and teaching styles in engineering education. *Eng Educ* 78(7):674–681
- Hmedna B, El Mezouary A, Baz O (2019) A predictive model for the identification of learning styles in MOOC environments. *Clust Comput*:1–26
- El Aissaoui O, El Madani YEA, Oughdir L et al (2019) A fuzzy classification approach for learning style prediction based on web mining technique in e-learning environments. *Educ Inf Technol* 24(3):1943–1959
- Khamparia A, Pandey B (2018) SVM and PCA based learning feature classification approaches for E-learning system. *Int J Web Based Learn Teach Technol* 13(2):32–45
- Allioui YE (2019) Advanced prediction of learner's profile based on Felder-Silverman learning styles using web usage mining approach and fuzzy c-means algorithm. *Int J Comp Aided Eng Technol* 11(4–5):495–512
- Starc <http://sdxt.starc.com.cn>

17. Graf S, Kinshuk K (2007) Providing adaptive courses in learning management systems with respect to learning styles. In: E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education. Association for the Advancement of Computing in Education (AACE), pp 2576–2583
18. Carruthers SA, Young A (1980) Preference of condition concerning time in learning environments of rural versus city eighth grade students. *Learn Styles Netw Newslett* 1:1
19. Shea TC (1983) An investigation of the relationship among preferences for the learning style element of design, selected instructional environments and reading achievement of ninth grade students to improve administrative determinations concerning effective educational facilities. Unpublished doctoral dissertation, St. John's University
20. Krinsky JS (1982) A comparative study of the effects of matching and mismatching fourth-grade students with learning style preferences for the environmental element of light and their subsequent reading speed and accuracy scores [D]. ProQuest Information and Learning
21. Murrain PG (1983) Administrative determinations concerning facilities utilization and instructional grouping: An analysis of the relationship (s) between selected thermal environments and preferences for temperature, an element of learning style, as they affect word recognition scores of secondary school students. Diss. ProQuest Information & Learning
22. Curry L (1983) An organization of learning styles theory and constructs
23. Kolb DA (2014) *Experiential learning: Experience as the source of learning and development*. FT press
24. Mumford A, Honey P (1982) *The manual of learning styles*. P. Honey, Maidenhead
25. Dunn R, Dunn K, Price G (1985) In: Lawrence KS (ed) *Manual: learning style inventory, vol 1*. Price Systems
26. Hummel KA, Hlavacs H (2003) Anytime, anywhere learning behavior using a web-based platform for a university lecture. In: *Proceedings of the SSGRR 2003 Winter Conference, L'Aquila*
27. Villaverde JE, Godoy D, Amandi A (2006) Learning styles' recognition in e-learning environments with feed-forward neural networks. *J Comput Assist Learn* 22(3):197–206
28. García P et al (2007) Evaluating Bayesian networks' precision for detecting students' learning styles. *Comput Educ* 49(3):794–808
29. Cha HJ, Kim YS, Park SH, Yoon TB, Jung YM, Lee JH (2006) Learning style diagnosis based on user interface behavior for the customization of learning interfaces in an intelligent tutoring system. In: *Lecture notes in computer science, proceedings of the international conference on intelligent tutoring systems, vol 4053*. Springer, Berlin, pp 513–524
30. Graf S, Kinshuk K (2007) Providing adaptive courses in learning management systems with respect to learning styles. In: Bastiaens T, Carliner S (eds) *Proceedings of world conference on E-learning in corporate, government, healthcare, and higher education*. AACE, Chesapeake, pp 2576–2583
31. Myers IB, McCaulley MH, Most R (1985) *Manual, a guide to the development and use of the Myers-Briggs type indicator*. Consulting Psychologists Press
32. Carver CA, Howard RA, Lane WD (1999) Enhancing student learning through hypermedia courseware and incorporation of student learning styles. *IEEE Trans Educ* 42(1):33–38
33. Kuljis J, Liu F (2005) A comparison of learning style theories on the suitability for elearning. *Web Technol Appl Serv* 2005:191–197
34. Akbulut Y, Cardak CS (2012) Adaptive educational hypermedia accommodating learning styles: a content analysis of publications from 2000 to 2011. *Comput Educ* 58(2):835–842
35. Hecht-Nielsen R (1992) Theory of the backpropagation neural network. In: *Neural networks for perception*. Academic Press, pp 593–605
36. Hinton GE (2012) A practical guide to training restricted Boltzmann machines. In: *Neural networks: Tricks of the trade*. Springer, Berlin, p. 599–619
37. Graf S, Liu TC, Kinshuk K (2010) Analysis of learners' navigational behavior and their learning styles in an online course. *J Comp Assisted Learn* 26(2): 116–131
38. Glaser R (1963) Instructional technology and the measurement of learning outcomes. *Am Psychol* 18:510–522
39. Lo JJ, Chan YC, Yeh SW (2012) Designing an adaptive web-based learning system based on students' cognitive styles identified online. *Comp Educ* 58(1):209–222
40. Lee WI, Shih BY, Chung YS (2008) The exploration of consumers' behavior in choosing hospital by the application of neural network. *Expert Syst Appl* 34(2):806–816
41. Ruangkanokmas P, Achalakul T, Akkarajitsakul K (2016) Deep belief networks with feature selection for sentiment classification [C]//7th international conference on intelligent systems, Modelling and simulation
42. Fischer A, Igel C (2012) An introduction to restricted Boltzmann machines. In: *Iberoamerican congress on pattern recognition*. Springer, Berlin, pp 14–36
43. Discover your learning styles—graphically [OL]. <https://learning-styles-online.com/>
44. Surjono HD (2015) The effects of multimedia and learning style on student achievement in online electronics course. *Turkish Online J Educ Technol* 14(1):116–122

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
