

Retraction

Retracted: College English Online Teaching Model Based on Deep Learning

Security and Communication Networks

Received 26 December 2023; Accepted 26 December 2023; Published 29 December 2023

Copyright © 2023 Security and Communication Networks. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] H. Ding, Y. Chen, and L. Wang, "College English Online Teaching Model Based on Deep Learning," *Security and Communication Networks*, vol. 2021, Article ID 8919320, 11 pages, 2021.

Research Article

College English Online Teaching Model Based on Deep Learning

Hui Ding ¹, Yajun Chen ², and Linling Wang ³

¹School of Foreign Languages, Changchun Normal University, Changchun 130032, Jilin, China

²School of Humanities and Social Sciences, Baotou Medical College, Baotou 014040, Inner Mongolia, China

³College of Computer Science and Technology, Changchun Normal University, Changchun 130032, Jilin, China

Correspondence should be addressed to Hui Ding; dinghui@ccsfu.edu.cn

Received 25 October 2021; Revised 24 November 2021; Accepted 29 November 2021; Published 21 December 2021

Academic Editor: Jian Su

Copyright © 2021 Hui Ding et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In today's era, online teaching plays an important part in the college English teaching. Deep learning, famous for its ability of imitating the learning process of human brains and obtaining the internal essential features or rules of voice, videos, images, and other data, can be applied to assist and improve the college English online teaching which involves a wide use of those data. Based on the combination of the multilayer neural network model and the k -means clustering algorithm, this paper designs a kind of deep learning method that can be used to assist and improve the college English online teaching. Experiments were designed to test the reliability of this deep learning method. The results show that the optimization algorithm designed in this paper, which can adjust the learning rate, will improve the common probability gradient descent algorithm. Besides, it is proved that the deep learning's efficiency of the CNN model is significantly higher than that of the MLP model. With the help of this deep learning method, it becomes feasible to apply the technologies related to the artificial intelligence to help teachers deeply analyze and diagnose students' English learning behavior, replace the teachers in part to answer students' questions in time, and automatically grade assignments in the process of the college English online teaching. Surveys and exams were then conducted to evaluate the effect of the application of the college English online teaching model based on deep learning on the students' learning cognition and their academic performance. The results show that the college English online teaching model based on deep learning can stimulate students' learning motivation and improve their academic performance.

1. Introduction

Today, the development of the information technology has greatly shaped the ways of teaching in colleges and online teaching has become an indispensable part of the teaching of most courses. College English teaching in China is no exception. The relevant research studies so far have in large part focused on how to choose and use proper online teaching platforms and different online teaching models to carry out the college English online teaching as well as the factors that affect the college English online teaching.

It should be pointed out that in the process of carrying out the college English online teaching, there is often a lack of effective learning supervision mechanism, quick learning behavior analysis, and the corresponding evaluation and feedback according to the individual characteristics of students, which bothers both teachers and students.

In order to improve the effect of the college English online teaching, many researchers turn to the artificial intelligence techniques. Sun et al. made use of the decision tree algorithm and neural networks to develop an implementation model of the English teaching assessment to help teachers to improve their teaching and students' English performance [1]. The results of their research prove that the relevant artificial intelligence techniques can play a key role in improving the online English teaching. Among the common artificial intelligence techniques, deep learning algorithms are very powerful. This paper concentrates on the comparison and analysis of the existing neural network algorithms and then based on the combination of the neural networks and clustering algorithms, it attempts to design a kind of deep learning method that is expected to be of help to the improvement of the college English online teaching.

The innovation of this paper lies in the design of a deep learning method that combines neural networks and clustering algorithms. With the help of this deep learning method, it becomes feasible to apply the technologies related to the artificial intelligence to deeply analyze and diagnose students' English learning behavior, replace the teachers in part to answer students' questions in time, and automatically grade assignments in the process of the college English online teaching.

2. Deep Learning Method Based on Artificial Neural Network and Clustering Algorithm

2.1. Status Quo of Deep Learning Research. In the last few years, there has been a large growth in research activity into deep learning. Hao et al. offered a general introduction to deep learning in neural networks, involving popular architectural models and training algorithms [2]. Due to the increase of the data volume and the improvement of the computing power, neural networks with more complex structures have received extensive attention and have been applied to various fields. A large number of research studies were focused on the practical applications, and abundant research results have been achieved. Deep learning algorithms have been used in the research of image analysis, text analysis, speech recognition, etc., offering solutions to many problems in the real life. Wang et al. reviewed the background knowledge of the transfer learning, various kinds of strategies used to perform the transfer learning, and how the transfer learning was applied in various subfields of medical image analysis [3]. The review proves that the latest advances in deep learning, especially advances in transfer learning, have made it possible to identify, classify, and quantify certain patterns from a large number of medical images. Lu et al. proposed a Char-CNNs (Character-level Convolutional Neural Network with Shortcuts) model as an attempt to offer an automatic solution to the identification of whether the text in social media contains cyberbullying [4]. Hourri et al. proposed a new approach to extract speaker characteristics by constructing CNN filters linked to the speaker to help people understand how deep learning models can be adapted to tackle the problem of speaker recognition [5].

Deep learning is also widely applied in the field of education. In the school education, how to monitor the students' learning state in the class is an important concern. Deep learning algorithms have been used by researchers to develop efficient ways to evaluate the students' learning state and behavior [6–8]. Interactions are the key factor to the success of the teaching and learning. For the abilities in natural language generation and processing, deep learning algorithms have been adopted by researchers to design systems to support and stimulate students' learning enthusiasm [9]. Besides, deep learning can be used to help students in their web-based study. Smith et al. conducted an investigation on how to make it possible to use Natural Language Processing (NLP) algorithms to automatically grade answers to open-ended inference questions in web-

based e-books, with the aim of making reading more appealing to children and increasing their comprehension [10].

2.2. Artificial Neural Network

2.2.1. Overview of the Artificial Neural Network. The research on the artificial neural network began in the 1940s, but at the beginning, its development was relatively slow [11]. In the 1980s, under the influence of Hopfield and his relevant research results, research on artificial neural networks received a revival. In the past decade, more and more progress has been made in the research work on the artificial neural networks.

Artificial neural network [12] is the product of modern neuroscience research, which is designed to simulate the way the human nervous system stores and processes the information. It reflects some basic characteristics of the brain function. Artificial neural network belongs to a kind of algorithmic mathematical model that processes the information in a distributed and parallel way. It is a nonlinear, adaptive, self-organizing system which is composed of many simple processing units. It has the ability to learn and model nonlinear and complex relationships. Based on the complexity of the system, it achieves the purpose of information processing by adjusting the relationship between a large number of internal nodes. After learning the initial inputs and their relationships, it can infer invisible relationships between invisible data, thereby generalizing and predicting unknown data. Artificial neural networks can be applied in many fields, including the control field, the psychological field, the medical field, the educational field, the financial field [13], the traffic field, and the information field. Figure 1 shows some common applications of the artificial neural networks.

2.2.2. Artificial Neurons and Their Characteristics. The artificial neuron [14] is the basic processing unit of the artificial neural network. One type of the artificial neuron model is shown in Figure 2. It is composed of input signals, weights, bias, adder, and activation function. Each neuron is an information processing unit with multiple inputs but a single output. The input signal a_y ($y = 1, 2, \dots, n$) received from other neurons is transmitted through the weighted connection ω_{xy} which represents the weight corresponding to the y -th input of the x -th neuron and whose value also indicates the contribution of the signal input to the output. The total input value received by the neuron x will be compared with the threshold of the neuron, whose function is to adjust the input of the activation function $f(\bullet)$, and then be processed by the activation function to generate the output of the neuron.

The threshold of the neuron is denoted by θ_x . v_x , also known as the local sensing region of the neuron, is the value adjusted by the bias. Here, the bias is denoted by j_x . u_x is the output of the input signals after the linear combination and is the net input of the neuron x . The neuron x depicted in Figure 2 can be described by the following equations:

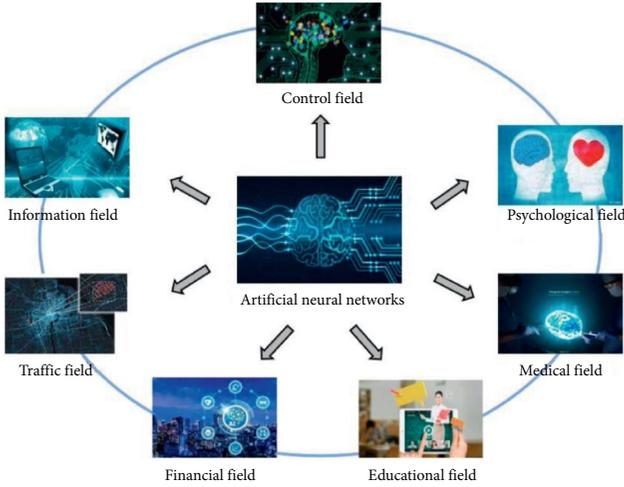


FIGURE 1: Applications of artificial neural networks.

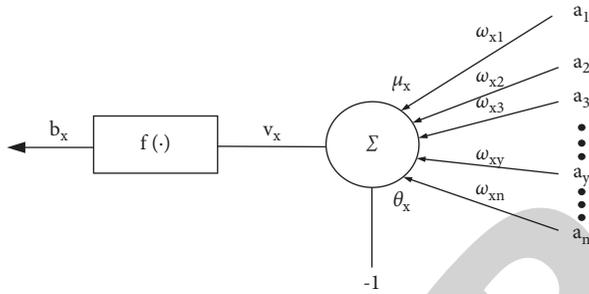


FIGURE 2: One type of the artificial neuron model.

$$\begin{aligned}
 v_x &= u_x + j_x, \\
 u_x &= \sum_y \omega_{xy} a_y, \\
 b_x &= f\left(\sum_y \omega_{xy} a_y + j_x\right).
 \end{aligned} \tag{1}$$

$f(\bullet)$ is the activation function, and b_x is the output of the neuron x . There are many kinds of activation function. Here, three common activation functions are introduced, namely the threshold function, the piecewise linear function, and the sigmoid function:

(1) Threshold function:

$$f(a) = \begin{cases} 1, & \text{if } a \geq 0, \\ 0, & \text{if } a < 0. \end{cases} \tag{2}$$

Threshold function, as shown in Figure 3, is usually called a step function. When the output of the neuron is 1 or 0, a step function can reflect the excitation or inhibition of the neuron.

(2) Piecewise linear function:

$$f(a) = \begin{cases} 1, & a \geq 0, \\ a, & -1 < a < 0, \\ -1, & a \leq -1. \end{cases} \tag{3}$$

As shown in Figure 4, this form of piecewise linear function can be regarded as an approximation of a nonlinear function.

(3) Sigmoid function:

The Sigmoid function is also called the S function, which is the most commonly used activation function in artificial neural networks. It is defined as follows:

$$f(a) = \frac{1}{1 + e^{-a}}. \tag{4}$$

Or,

$$f(a) = \frac{1 - e^{-a}}{1 + e^{-a}}. \tag{5}$$

The latter is shown in Figure 5.

2.2.3. Artificial Neural Network Models. As the most basic unit of neural networks, a single neuron has relatively simple structures and functions. However, the artificial neurons can form different artificial neural network models by changing the connection mode and the number and layers of neurons in different ways. In other words, the artificial neural network is composed of many artificial neurons and it has a powerful ability to process information. The artificial neural network often includes an input layer, a hidden layer, and an output layer. Many relationships between input and output are nonlinear, and no fixed relationship is imposed on the data.

There are various types of artificial neural network models, describing and simulating various levels of biological nervous systems from various angles. Schmidhuber summarized some common types of artificial neural network, including the feedforward neural networks, the recurrent neural networks, the convolutional neural networks, the bidirectional recurrent neural networks, and the deep belief networks [15]. Since the topological structure of the network is an important characteristic of the neural networks, much attention has been paid to the feedforward neural networks and the feedback neural networks, two common types of interconnected networks based on the connections of neurons:

(1) Feedforward neural network:

The feedforward neural network has a hierarchical structure composed of multiple layers, and neurons in the same layer are not connected to each other.

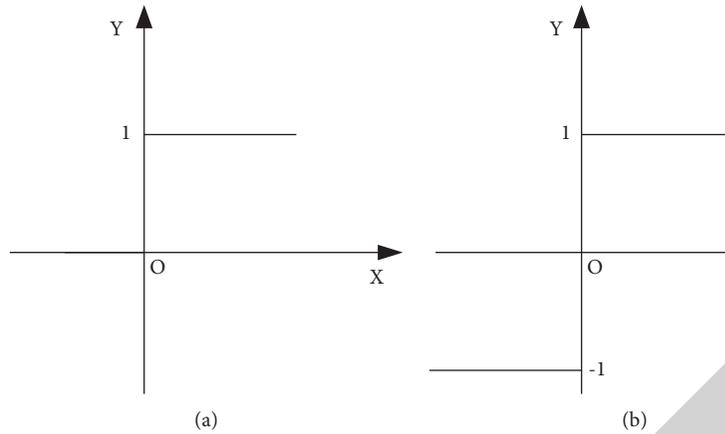


FIGURE 3: Threshold function.

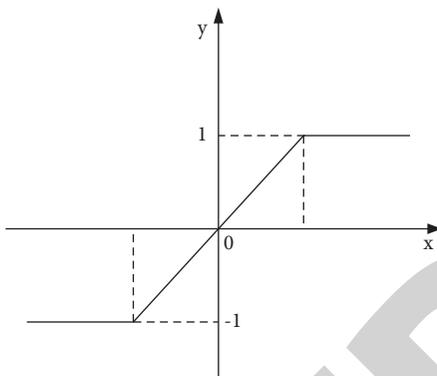


FIGURE 4: Piecewise linear function.

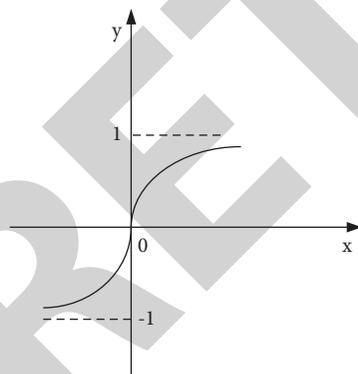


FIGURE 5: Sigmoid function.

The signal of the feedforward neural network has been transmitted forward and will not return. For any continuous function in a complete space, a neural network with a single hidden layer can always be found, so that the difference between it and the objective function at any point is arbitrary small. As shown in Figure 6, neurons are connected from one layer to the next. W and V , respectively, represent the weight of the hidden node of the input node and the output node. Y represents the output value after

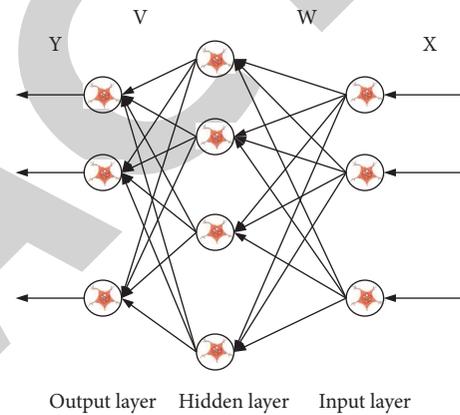


FIGURE 6: Feedforward neural network.

convergence, and X represents the input value of the node. There is no connection between neurons in the same layer. The signal loop from the input layer to the output layer is connected in one direction. In addition to the input layer, the neurons in the hidden layer and the output layer also realize certain operations, and they are called computing nodes. The BP neural network is a typical feedforward neural network.

(2) Feedback neural network:

Feedback neural network refers to the neural network with feedback loop in the topology. Typical feedback neural networks include Hopfield network, Elman network, CG network model, box-in-brain model, bidirectional associative memory, and so on [16]. In the feedback neural network, multiple neurons are connected to each other to form a connected neural network, and the output of some neurons is fed back to the same layer or the upper layer neurons, as shown in Figure 7.

2.2.4. Derivation of BP Algorithm. As the most widely used algorithm model in artificial neural network, the BP neural network imitates the process of human brain neurons'

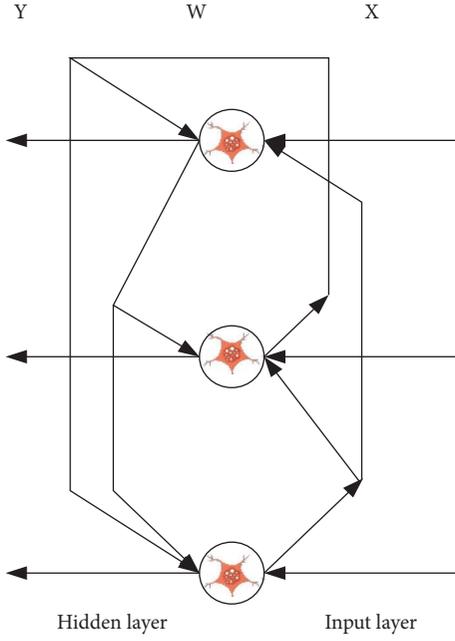


FIGURE 7: Feedback neural network.

reaction to external excitation signals. BP neural network establishes a multilayer perceptron model, and by using the learning mechanism of signal forward propagation and error back regulation, it successfully constructs an intelligent network model for processing nonlinear information through multiple iterative learning. The weight of the network is adjusted according to the deviation function of the sample training error weight. The network error is as follows:

$$Q = \frac{1}{2} \sum_y (b_y^t - b_y)^2. \quad (6)$$

Among them, b_y is the actual output of the y -th node and b_y^t is the expected output of the y -th output node. The connection weights between all the nodes of the neural network are expressed in vector form as $[w_{11}, w_{22}, \dots, w_{xy}, \dots]$, and the weight vector $[w_{11}, w_{22}, \dots, w_{xy}, \dots]$ is a function. The formula of the error function Q for each weight gradient is as follows:

$$\begin{aligned} \nabla Q &= \nabla f(w_{11}, w_{12}, \dots, w_{xy}, \dots) \\ &= \left[\frac{\partial f}{\partial w_{11}}, \frac{\partial f}{\partial w_{12}}, \dots, \frac{\partial f}{\partial w_{xy}}, \dots \right]^T. \end{aligned} \quad (7)$$

For the error function Q , the process of gradient descent is the process of finding the lowest point along the surface. Generally speaking, the gradient value is calculated by estimating the deviation function of the error function with respect to each weight. Taking a 3-layer BP network as an example, according to the error calculation formula in (6), as well as the chain law of partial derivatives, the calculation formula for the partial derivative of the connection weight error is as follows:

$$\frac{\partial Q}{\partial w_{xy}} = \frac{\partial Q}{\partial b_y} \frac{\partial b_y}{\partial w_{xy}}. \quad (8)$$

Therefore, the partial derivative of the error with respect to the sample is as follows:

$$\frac{\partial Q}{\partial b_y} = -(b_y^t - b_y). \quad (9)$$

And, the output of the hidden layer is as follows:

$$b_y = \text{sig} \left(\sum_x w_{xy} j_x \right). \quad (10)$$

Then, the partial derivative of the output of the hidden layer to the connection weight is as follows:

$$\frac{\partial b_y}{\partial w_{xy}} = \text{sig}' \left(\sum_x w_{xy} j_x \right) j_x. \quad (11)$$

Substitute (9) and (11) into formula (8) to get

$$\frac{\partial b_y}{\partial w_{xy}} = -(b_y^t - b_y) \text{sig}' \left(\sum_x w_{xy} j_x \right) j_x \delta_y^c j_x, \quad (12)$$

where δ_y^c is the local gradient of the output layer.

$$\delta_y^c = -(b_y^t - b_y) \text{sig}' \left(\sum_x w_{xy} j_x \right). \quad (13)$$

Therefore, the current weight change from the input layer to the hidden layer can be obtained.

$$\begin{aligned} \Delta W_{xy} &= -\eta \frac{\partial Q}{\partial w_{xy}} \\ &= \eta \sum_p -(b_y^t - b_y) \text{sig}' \left(\sum_x w_{xy} j_x \right) j_x = \eta \sum_p \delta_y^c j_x. \end{aligned} \quad (14)$$

Similarly, the current weight change from the hidden layer to the output layer can be obtained.

$$\begin{aligned} \Delta W_{xh} &= -\eta \frac{\partial Q}{\partial w_{xh}} \\ &= \eta \sum_p \delta_x^j i_h. \end{aligned} \quad (15)$$

According to the formula

$$w(t) = w(t-1) - i \frac{\partial Q}{\partial w}, \quad (16)$$

the amount of change in each weight can be obtained.

2.3. Proof of k -Means Matrix Factorization. Affected by the number of clusters and the initial distribution, a simple calculation algorithm is necessary to be applied to solve the complicated problem of article data processing. Clustering [17] divides data into various categories or clusters based on

similarity. Aggregation algorithms are now used in various fields and play an important role in psychology, social sciences, information retrieval, pattern recognition, machine learning, data, etc. Traditional clustering algorithms are usually divided into grid clustering, hierarchical clustering, partition clustering, density clustering, and model clustering. Each clustering type can be classified into various categories according to the idea of the algorithm. According to the characteristics of traditional clustering algorithms generally used in data mining, this paper will summarize these algorithms and analyze the ideas and performance characteristics of the algorithms. Clustering algorithms can be developed into various fields based on the background of big data.

The k -means clustering algorithm [18] is the most basic and most commonly used clustering algorithm in machine learning. The basic meaning is that the given data set $X \in R^m$ is divided into subsets C_1, C_2, \dots, C_k , where k is the number of clustering categories. The basic mathematical formula of the k -means algorithm is as follows:

$$\sum_{x=1}^k \sum_{y=1}^n c_{xy} \|a_y - \mu_x\|^2. \quad (17)$$

$$\|A\|^2 = \sum_{l,x} a_{lx}^2 = \sum_y \|a_y\|^2 = \text{tr}[A^T A],$$

$$\sum_{x,y} c_{xy} \|a_y - \mu_x\|^2 = \sum_{x,y} c_{xy} a_y^T a_y - 2 \sum_{x,y} c_{xy} a_y^T \mu_x + \sum_{x,y} c_{xy} \mu_y^T \mu_y = T_1 - 2T_2 + T_3. \quad (20)$$

Simplify the T_1, T_2 , and T_3 terms to get

$$\begin{aligned} T_1 &= \sum_{x,y} z_{xy} a_y^T a_y, \\ &= \sum_{x,y} c_{xy} \|a_y\|^2 = \sum_y \|a_y\|^2 = \text{tr}[A^T A] \\ T_2 &= \sum_{x,y} c_{xy} a_y^T \mu_x, \\ &= \sum_{y,l} a_{xy} \sum_x \mu_{lx} c_{xy} \sum_y (A^T MC) = \text{tr}[A^T MC] \\ T_3 &= \sum_{x,y} c_{xy} \mu_x^T \mu_x, \\ &= \sum_{x,y} c_{xy} \|\mu_y\|^2 = \sum_x \|\mu_y\|^2 n_x. \end{aligned} \quad (21)$$

Here, n_x is the number of samples belonging to the x -th category. Expand the right side of the equation to get

$$\begin{aligned} \|A - MC\| &= \text{tr}[(A - MC)^T (A - MC)] \\ &= \text{tr}[A^T A] - 2\text{tr}[A^T - MC] \text{tr}[C^T M^T MC] \\ &= T_4 - 2T_5 + T_6. \end{aligned} \quad (22)$$

Here, a_y is each sample point and μ_x is the cluster center. When a_y belongs to the G_x , $c_{xy} = 1$; otherwise, $c_{xy} = 0$.

$$c_{xy} = \begin{cases} 1, & a_y \in G_x, \\ 0, & \text{otherwise.} \end{cases} \quad (18)$$

The proof of the possibility of matrix decomposition of the k -means clustering algorithm is the same as the proof that the purpose function of k -means can be decomposed into the next matrix form.

$$\sum_{x=1}^k \sum_{y=1}^n c_{xy} \|a_y - \mu_x\|^2 = \|A - MC\|^2. \quad (19)$$

A is a sample matrix, where each column represents a sample data; M is a cluster center matrix, where M_y is the y -th column of M , which represents a cluster center. We define each column of sample matrix A as a_y ; we define the 2-norm of A as the sum of the squares of the length of all sample vectors, shown as follows:

It can be found that $T_4 = T_1$ and $T_5 = T_2$. Therefore, we only need to prove $T_3 = T_6$, and the proof is as follows:

$$T_6 = \text{tr}[C^T M^T MC] = \text{tr}[M^T MCC^T]. \quad (23)$$

Further derivation can be obtained:

$$\begin{aligned} \text{tr}[M^T MCC^T] &= \sum_x (M^T M Z Z^T)_{xx}, \\ &= \sum_x \sum_l (M^T M)_{xl} (C C^T)_{lx} = \sum_x \|\mu_x\|^2 n_x. \end{aligned} \quad (24)$$

It can be seen that the proof of $T_3 = T_6$ is established, so the equation to be proved in the end is established.

The conclusion can be drawn from the proof process: k -means is a solution process of matrix factorization.

3. Deep Learning Test Experiment Based on Neural Network and Clustering Algorithm

3.1. Impact of Different Learning Rates on Network Convergence. In order to study the influence of different learning rates on network convergence, three different learning parameters of the MLP network are used for training. The output layer has only one neuron, and no

activation function is used. The loss function is the mean square error. Since the data set is very noisy, fewer neurons are used in order to avoid learning too much noise. The value of the cost function is the vertical axis, and Figure 8 is the convergence graphs of various learning rates obtained through two experiments.

By analyzing Figure 8, it can be found that when the learning rate is 0.02, the network will converge near Epoch27. If the learning rate increases and the learning rate is 0.25, the convergence speed of the network will increase and it will converge near Epoch18. The network learning rate 0.25 is significantly higher than the network learning rate 0.02. This shows that the larger the learning step size, the faster the network weight update and the faster the final network training time. After the learning rate rises to a certain level, the network training will vibrate and the network will become nonconvergent. As shown by the red curve, when the network learning rate is 3, the network training cannot converge. The choice of learning rate is very important for neural networks. The spatial distribution of the error function has multiple flat areas and minimum intervals. While shortening the time required for the network training process, the final recognition error meets the requirements, so this paper designs an optimization algorithm that can adjust the learning rate. In this way, the previous probability gradient descent algorithm will be improved.

3.2. Gradient Changes of Each Hidden Layer during MLP and Perceptron Training. The MLP network structure and the improved SGD algorithm [19] are used for parameter training of the MNIST data set, and the smallest batch of gradient data for each small sample is saved in the training process. A part of the gradient value of the hidden layer is representative. The number of hidden layers as a gradient change graph with the horizontal axis and the gradient value on the vertical axis is calculated. Figure 9 is the comparison of the gradient changes in the perceptron training and in the MLP training.

From the gradient curve of various hidden layers of the multilayer feedforward network in Figure 9, the numerical distribution of various minibatch becomes sparse and the gradient values of various samples of layer 1 are almost equal. But, the gradient values of different samples of layer 4 are larger. When the experimental error is small, it is necessary to consider changing the number of nodes in the previous hidden layer. In addition, the MLP network is obviously more efficient than the insight network, and the MLP network is used for deep learning.

3.3. Network Performance Analysis of CNN and MLP. Figure 10 is the recognition results of the CNN model and the MLP model in the validation set and the data set. The verification group and the test group each contained 10,000 images.

In this paper, the MLP model and the CNN model were used to run 500 epochs, respectively. The recognition rate of the observation table shows that the recognition rate of CNN

has increased rapidly in the first few eras. As the times increase, the growth rate of cognition will also increase. When the CNN is Epoch5, the test accuracy reaches 98.30%. When the MLP is Epoch15, the test accuracy is only 96.61%. The CNN model has a better learning effect than the MLP model. Since there are few trainable parameters in CNN, if the network is trained in turn, the training speed will be faster, the learning ability of the network will be improved, and the performance will also be improved.

3.4. Impact of Increasing the Weight Attenuation Term on the Generalization Ability of the CNN. In order to reduce the overmatching phenomenon of the network, this section mainly uses the method of increasing the weight attenuation term L1 to optimize the network parameters. In other words, the weighted attenuation term is added to the formula of the cross-entropy cost function, and the generalization of the network is improved by adjusting the function of various parameter values. As shown in Figure 11, this paper used a convolutional network structure (CNN1) to design 3 sets of different weighted attenuation control coefficients and conducted experiments by analyzing the number of neuron nodes in the fully connected layer without changing their weights.

Through the above three groups of experiments, we found that the experimental effect of the second group is slightly better than the experimental effect of the first group, but the difference between the two groups is not obvious and the experimental effect of the third group is very good. The average recognition rate is higher than the previous two, and the speed is fast. The model has relatively close experimental effects on the verification set and the test set, and the generalization ability of the network is relatively strong.

4. Investigation and Analysis of the College English Online Teaching Model Based on Deep Learning

4.1. College English Online Teaching Model Based on Deep Learning. The main feature of college English online teaching is to take students as the center, fully considering students' needs and providing them with convenient and intelligent teaching services. Therefore, any college English online teaching model must do its utmost to meet students' personalized learning needs and give feedback and evaluation to their learning behavior in time. The functions of automatic homework correction, online question answering, speech recognition evaluation, and data collection and analysis in the learning process can play an important role in the college English online teaching model. The realization and optimization of these functions are closely related to deep learning. The deep learning method combining clustering algorithm and neural network proposed in this paper can be used to continuously optimize these functions necessary to the college English online teaching model.

This paper then chose two classes in a college as the test class and the control class and spent one term applying two

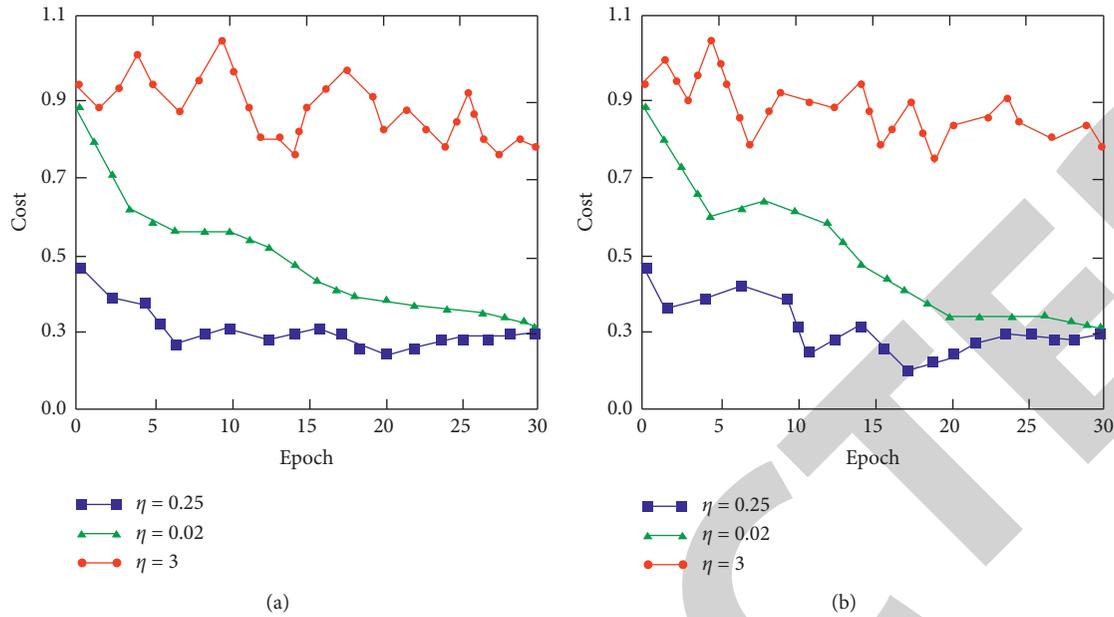


FIGURE 8: Curves of the influence of different learning rates on network convergence. (a) The first experimental test. (b) The second experimental test.

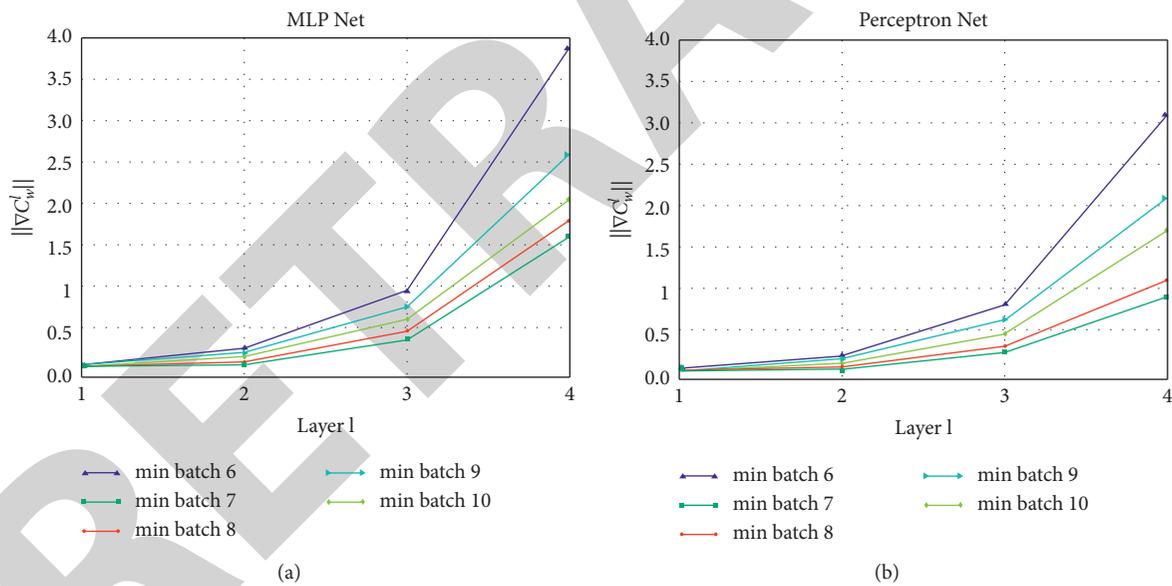


FIGURE 9: (a) Gradient change graph during MLP network training. (b) Gradient change graph during perceptron network training.

kinds of college English online teaching models in the two classes, respectively. In the test class the college English online teaching model based on deep learning was applied, while in the control class, the traditional college English online teaching model was applied. After that, this paper conducted a questionnaire survey and an exam to test the impact of the college English online teaching model based on deep learning on students' learning cognition and academic performance.

4.2. Survey of Students' Learning Cognition. After a term's study, this paper carried out a questionnaire survey on students' learning cognition. A total of 67 questionnaires were collected in the test class and 64 in the control class. The contents and results of the questionnaire are shown in Table 1. It can be seen from the table that for the students in the test class who adopted the college English online teaching model based on deep learning, the number of them who chose C when they answered the six questions in the

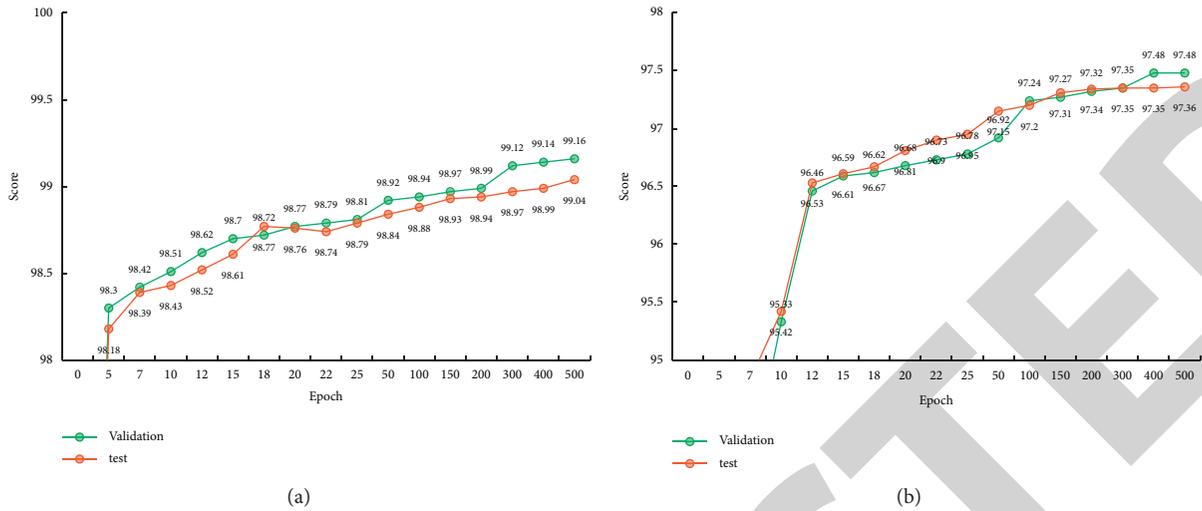


FIGURE 10: The recognition results of CNN and MLP on the training set and test set. (a) CNN model test results. (b) MLP model test results.

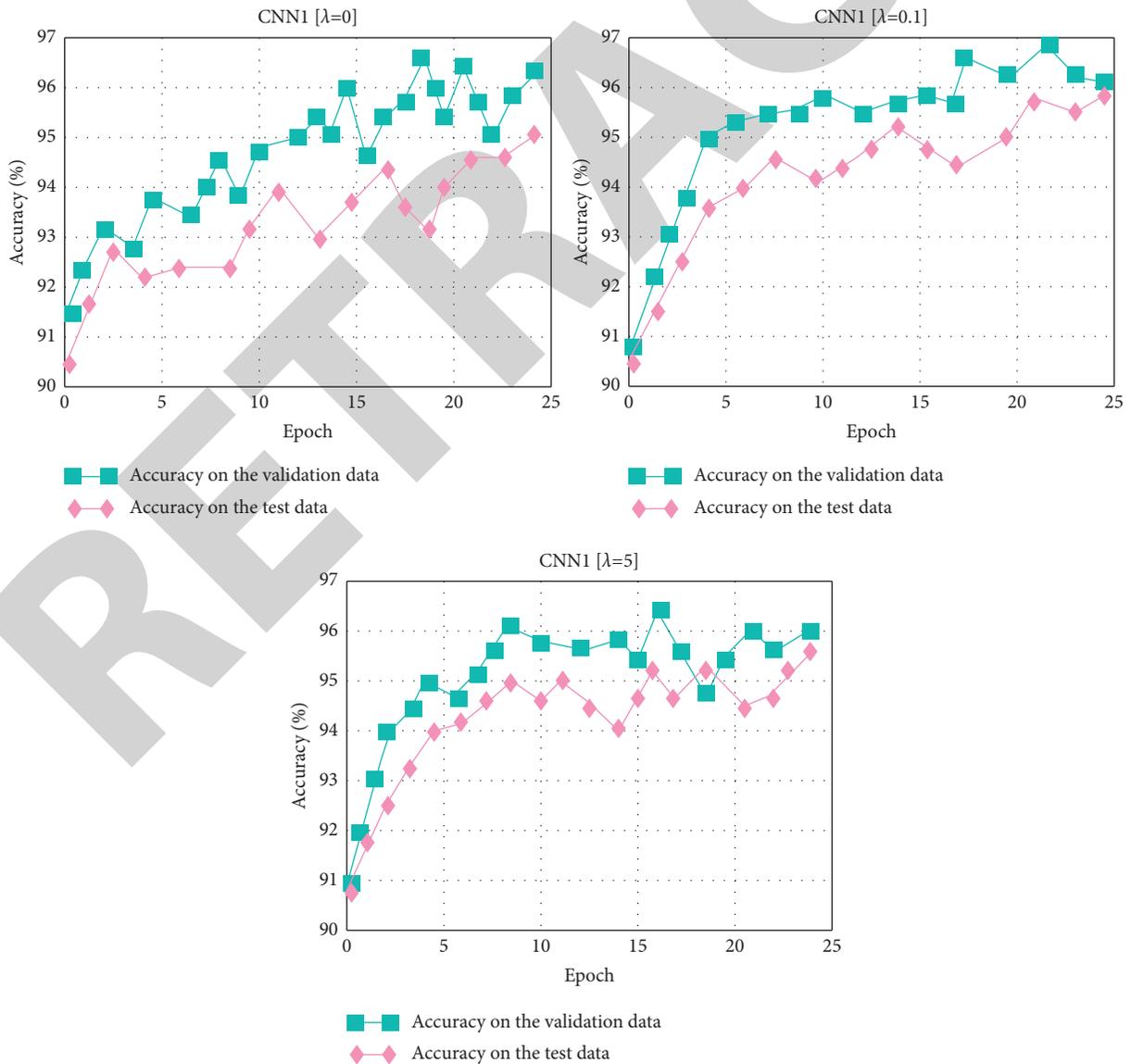


FIGURE 11: The recognition results of CNN1 models with different λ in three groups of experiments on the data set.

TABLE 1: Survey results of students' learning cognition.

Contents	Options	Testgroup	Controlgroup
After a term's study, have you fully realized why you study? (learning motivation)	A. Yes	26	21
	B. Basically	39	27
	C. No	2	16
After a term's study, can you now take the initiative to study? (learning attitude)	A. Yes	37	11
	B. Basically	30	24
	C. No	0	29
After a term's study, do you think you have mastered what were taught? (the acquisition of knowledge)	A.	32	13
	Completely		
	B. Basically	32	35
	C. No	3	16
After a term's study, can you now apply what you have learned in the real life? (practical ability)	A. Yes	33	17
	B. Basically	34	32
	C. No	0	15
After a term's study, can you now know how to learn? (learning method)	A. Yes	36	17
	B. Basically	30	19
	C. No	1	28

TABLE 2: Evaluation of students' academic performance.

Category	Group	Highest score	Lowest score	Average score
Listening	Test group	99	81	87.5
	Control group	85	65	72.33
Reading	Test group	96	80	88.5
	Control group	87	63	71.67
Translation	Test group	99	83	89.33
	Control group	81	64	75.8
Writing	Test group	93	79	85.5
	Control group	76	58	65.67

questionnaire is far less than that of the students in the control class. The majority of the students in the test class tended to have a better understanding of why they study, what to study, and how to study after a term's study. The results show that the college English online teaching model based on deep learning can have a positive impact on students' learning cognition and stimulate their learning motivation.

4.3. Evaluation of Students' Academic Performance. At the end of the term, the two classes took the same exam, which involved four sections, listening, reading, translation, and writing, to evaluate their academic performance. The relevant test results are shown in Table 2. By analyzing the test results of the two classes, it can be concluded that on average, students in the test class performed much better in the exams of listening, reading, translation, and writing than those in the other class. Besides, students who got the highest score in the exams of listening, reading, translation, and writing were all from the test class. This shows that the college English online teaching model based on deep learning can effectively improve students' academic performance.

5. Conclusions

In order to improve the college English online teaching, this paper designed a kind of online deep learning method based

on the combination of the multilayer neural network model and the k -means clustering algorithm, which can be used to help constantly improve the artificial intelligence technologies necessary for the college English online teaching. Experiments were designed to test the reliability of this deep learning method. The results show that the optimization algorithm designed by this paper, which can adjust the learning rate, will improve the common probability gradient descent algorithm. Besides, it is proved that the deep learning's efficiency of the CNN model is significantly higher than that of the MLP model. Then the college English online teaching model based on deep learning was applied in practice. From the results of the survey and the exam taken by the students, it is found that students who received the college English online teaching model based on deep learning had more positive attitudes towards learning and they behaved better in the exam of listening, reading, translation, and writing. In other words, the college English online teaching model based on deep learning can stimulate students' learning motivation and improve their academic performance.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the present study.

Acknowledgments

This work was supported by a foundation from 2020 Education and Science Research of Inner Mongolia Autonomous Region during the 13th Five-Year Plan Period, “Research on the Construction of College English Online Open Courses under the Flipped Classroom Mode” (Project Number: NGJGH2020198); a foundation from 2021 Asking, Learning and Practicing Program of Baotou Medical College, “Research on the Construction and Application of College English Online Open Course Resources Based on Autonomous Learning” (Project Number: 2021BYWWJ-ZC-18); and a foundation from 2020 Scientific Research Fund Project of Baotou Medical College, “Research on College English Blended Teaching Model Based on Medical Humanism—Taking Baotou Medical College as an Example” (Project Number: BYJJ-QWB 202019).

References

- [1] Z. Sun, M. Anbarasan, and D. P. Kumar, “Design of online intelligent English teaching platform based on artificial intelligence techniques,” *Computational Intelligence*, vol. 37, no. 3, pp. 1166–1180, 2020.
- [2] X. Hao, G. Zhang, and S. Ma, “Deep learning,” *International Journal of Semantic Computing*, vol. 10, no. 3, pp. 417–439, 2016.
- [3] J. Wang, H. Zhu, S.-H. Wang, and Y. D. Zhang, “A review of deep learning on medical image analysis,” *Mobile Networks and Applications*, vol. 26, no. 1, pp. 351–380, 2020.
- [4] N. Lu, G. Wu, Z. Zhang, Y. Zheng, Y. Ren, and K. K. R. Choo, “Cyberbullying detection in social media text based on character-level convolutional neural network with shortcuts,” *Concurrency and Computation*, vol. 32, no. 23, pp. 1–11, 2020.
- [5] S. Hourri, N. S. Nikolov, and J. Kharroubi, “Convolutional neural network vectors for speaker recognition,” *International Journal of Speech Technology*, vol. 24, no. 2, pp. 389–400, 2021.
- [6] A. A. Mubarak, H. Cao, and S. A. M. Ahmed, “Predictive learning analytics using deep learning model in MOOCs’ courses videos,” *Education and Information Technologies*, vol. 26, no. 1, pp. 371–392, 2020.
- [7] L. Chen and B. Ding, “Evaluation technology of classroom students’ learning state based on deep learning,” *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 6999347, 8 pages, 2021.
- [8] Y. C. Semerci and D. Goularas, “Evaluation of students’ flow state in an E-learning environment through activity and performance using deep learning techniques,” *Journal of Educational Computing Research*, vol. 59, no. 5, pp. 960–987, 2021.
- [9] C. Li and W. Xing, “Natural language generation using deep learning to support MOOC learners,” *International Journal of Artificial Intelligence in Education*, vol. 31, no. 2, pp. 186–214, 2021.
- [10] G. G. Smith, R. Haworth, and S. Žitnik, “Computer science meets education: Natural Language Processing for automatic grading of open-ended questions in eBooks,” *Journal of Educational Computing Research*, vol. 58, no. 7, pp. 1227–1255, 2020.
- [11] Y. Wu and J. Feng, “Development and application of artificial neural network,” *Wireless Personal Communications*, vol. 102, no. 2, pp. 1645–1656, 2018.
- [12] I. A. Basheera and M. Hajmeer, “Artificial neural networks: fundamentals, computing, design, and application,” *Journal of Microbiological Methods*, vol. 43, no. 1, pp. 3–31, 2000.
- [13] X. Zhang, “Construction and simulation of financial audit model based on convolutional neural network,” *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 1182557, 11 pages, 2021.
- [14] H. Simon, *Neural Networks and Learning Machines*, Prentice-Hall, Hoboken, NY, USA, 3rd edition, 2008.
- [15] J. Schmidhuber, “Deep learning in neural networks: an overview,” *Neural Networks*, vol. 61, pp. 85–117, 2015.
- [16] M. S. Ansari, *Non-Linear Feedback Neural Networks: VLSI Implementations and Applications*, Springer, New York, NY, USA, 2014.
- [17] A. Likas, N. Vlassis, and J. J. Verbeek, “The global k-means clustering algorithm,” *Pattern Recognition*, vol. 36, no. 2, pp. 451–461, 2003.
- [18] S. Vassilvitskii, *K-Means: Algorithms, Analyses, Experiments*, Stanford University, Stanford, CA, USA, 2007.
- [19] S. Zhang, A. E. Choromanska, and Y. LeCun, “Deep learning with elastic averaging SGD,” in *Proceedings of the 28th International Conference on Neural Information Processing Systems*, pp. 685–693, Montreal, Canada, December, 2015.