APPLICATION OF DEEP NN OPTIMIZED BY MULTI-PARAMETER FUSION IN IDEOLOGICAL AND POLITICAL CONSTRUCTION OF PROFESSIONAL COURSES IN COLLEGES AND UNIVERSITIES

Rui Ma*

School of Financial Technology, Suzhou Industrial Park Institute of Service Outsourcing, Suzhou, Jiangsu, 215123, China

mar489@126.com

Reception: 23/10/2022 **Acceptance**: 29/12/2022 **Publication**: 23/01/2023

Suggested citation:

M., Rui. (2023). **Application of deep NN optimized by multi-parameter fusion in ideological and political construction of professional courses in colleges and universities**. *3C Tecnología. Glosas de innovación aplicada a la pyme, 12(1)*, 54-68. https://doi.org/10.17993/3ctecno.2023.v12n1e43.54-68

ABSTRACT

Curriculum ideology and politics is an inherent requirement to achieve the goal of "cultivating morality and cultivating people" in colleges and universities, and it is a beneficial exploration to realize the three-round education. The ideological and political construction of professional courses in colleges and universities not only teaches students knowledge and skills, but also helps students form correct values. Aiming at how to build ideological and political courses in colleges and universities, a design method based on multi-parameter fusion to gradually optimize deep NN is proposed. Firstly, the initial NN model without hidden layer is determined by analyzing the samples and categories, and then the hidden layer is gradually added on the basis of the initial NN to construct a deep NN with multi-parameter fusion optimization. Based on the TensorFlow framework, taking handwritten digit recognition as an example, a deep NN model is gradually designed. During the whole experiment, the network structure, activation function, loss function, optimizer, learning rate and sample batch size are continuously adjusted, and finally a multi-parameter design is designed. The fusion optimized deep NN model with high accuracy provides an effective idea for building a NN. As the learning rate increases, the performance of the NN gradually improves. In the training set and test set, the accuracy rate is almost the highest when the learning rate is 0.3, and the accuracy rate is 93.30% and 92.58% respectively when the number of iterations is 30, which shows that The NN optimized by multi-parameter fusion can be well applied to the ideological and political construction of professional courses in colleges and universities, and has strong application prospects.

KEYWORDS

Deep NN; TensorFlow; Activation function; Learning rate; Loss function

https://doi.org/10.17993/3ctecno.2023.v12n1e43.54-68

PAPER INDEX

https://doi.org/10.17993/3ctecno.2023.v12n1e43.54-68

1. INTRODUCTION

"Course Ideological and Political" is a new education and teaching concept, which is different from the traditional ideological and political course education method, and runs through various professional courses in a hidden education way[1]. Curriculum ideology and professional knowledge education and teaching go in the opposite direction. While teachers disseminate professional knowledge, they also focus on leading students to establish a correct value orientation[2]. Curriculum ideology and politics take professional courses as the carrier, shoulder the important task of value leadership, and make it play the overall effect of "1+1>2". In its construction, teachers and students jointly establish a correct world outlook, outlook on life and values. Although the school has an independent ideological and political curriculum, it is not integrated with the major, so it may lead to the phenomenon that students have a weak sense of social responsibility and professional ethics in the future[3]. The close connection of the three has higher requirements for college students to have correct value orientation and firm ideals and beliefs.

In foreign education and teaching, "ideological and political education" and "ideological and political theory courses" are not clearly used to define the content of their courses, but implicit education is embedded in the teaching of various disciplines. Moral quality and values[4]. The social action model of moral education proposed by American educator Fred Newman. He believes that a virtuous member of society should have material competence, interpersonal competence, and civic competence[5]. Thomas Ricorner's moral education model of perfect personality believes that perfect personality includes three aspects: moral cognition, moral emotion and moral behavior. Section[6]. In 1916, American educator John Dewey pointed out in "Principles of Morality in Education" that "Students are accompanied by ideals, attitudes, moral habits and other learning that are different from the formal curriculum in addition to the learning of the formal curriculum, that is, 'incidental learning'. It is proposed that education must pay attention to the influence of various factors outside the formal curriculum on students” [7]. Fred Newman believes that students can acquire the environmental competencies necessary for meaningful moral discourse only through the study of civic action courses[8]. Sukhomlinsky emphasized that moral education must adhere to the unity of theory and practice, and must run through all aspects of school teaching and education. Teachers must teach and educate people, so that teaching and education are organically unified. In 1991, American educator Thomas Rickner proposed that "'character education' should carry out implicit education, create a campus moral culture atmosphere, fruitful education, democratic classroom life, story discussion method and role simulation training and other moral development[9]. Herbert Heyman made relevant research on political education in his book "Political Socialization: A Study of the Psychology of Political Behavior", studying how individuals receive political education and then disseminate political ideas., and eventually formed a political concept[10]. Dewey pointed out in "Principles of Moral Education" that "the conscious moral teaching in the classroom is not as good as it used to be, and it has committed the error of equating the teaching of ethics with the manipulation and instillation of moral precepts.", put forward that

https://doi.org/10.17993/3ctecno.2023.v12n1e43.54-68

moral education should focus on practicality[11]. In "Education and Democracy in the 21st Century", Neil Noddings explained the true connotation of democracy in education, and mentioned about patriotism, global citizenship, sublime School education in terms of moral feelings[12]. These all point to the practical significance of curriculum ideological and political construction.

The NN originated from McCulloch-Pitts (MCP) model, which is the earliest prototype of the artificial neural model[13]. In 1985, the perceptron algorithm was proposed, which enabled the MCP model to perform binary processing on multi-dimensional data, and then the back-propagation algorithm was proposed for the rapid development of modern NNs. Opened the door. In the 1980s, inspired convolutional NNs[14] and recurrent NNs[15] were successively proposed, and in the 1990s, LeNet[16] was applied to digit recognition and achieved good results. 2006 In 2009, due to the initial successful application of deep NN theory in machine learning, Hinton et al. proposed the concept of deep learning[17], which attracted people's attention. After years of development, it has gradually developed from a single-layer network to a multi-layer network. The multi-layer NN may contain hundreds of layers and hundreds of megabytes of training parameters. AlexNet, a deep learning architecture proposed in 2012, won the 2012 ILSVRC (image Net large-scale visual recognition challenge) crown, the error rate of Top-5 is reduced to 15.3%[18], and its effect is significantly ahead of traditional methods. In the following years, the recognition error rate has been continuously refreshed by new and deeper convolutional NNs. In 2014, VGGNet Obtained an average correct rate of 89.3%[19], and ResNet proposed by He et al. in 2016 reduced the classification error rate to 3.57%[20], while the SENet error recognition rate proposed by Hu Jie et al. in 2017 is only 2.25%. The introduction of various deep NN models has promoted the development of deep learning.

The rapid development of deep learning is inseparable from the design of better deep learning models, and people have gradually realized that the structure of deep learning models is the top priority of deep learning research[21]. The essence of deep learning is to build an artificial NN model with multiple hidden layers. The structure of artificial NN, whether shallow or deep, is mainly designed based on experiments and experience, but there is no set of specific theories to follow[22]. Based on the TensorFlow framework, this paper adopts a NN design method that is simple and then complex, and multi-parameter fusion is gradually optimized to optimize the ideological and political construction of professional courses in colleges and universities.

2. PRINCIPLES OF DEEP LEARNING MODELS

2.1. INTRODUCTION TO TENSORFLOW

Google has excellent performance in many fields related to computers, and the field of artificial intelligence is no exception[23]. TensorFlow is an excellent open source deep learning framework based on DistBelief developed by Google in 2015. The design of NN structure code is concise, and it is favored by more and more

https://doi.org/10.17993/3ctecno.2023.v12n1e43.54-68

developers. Not all tensFlow is written in Python, but many underlying codes are written in C++ or CUDA. It provides programming interfaces of Python and C++, and basic operations on threads and queues can be implemented from the bottom layer, and it can also be more convenient to call hardware resources. With the flexible architecture of TensorFlow, users can deploy to multiple platforms (CPU, GPU, TPU) for distributed computing and provide support for big data analysis. TensorFlow is also cross-platform and works on various devices (desktop devices, server clusters, mobile devices, edge devices).

2.2. DEEP NN MODEL DESIGN

2.2.1. DATA PREPROCESSING

In order to more easily extract the relevant information of the data during training, the data needs to be pre-processed. Data preprocessing includes normalization techniques, nonlinear transformations, feature extraction, discrete input, target coding, processing of missing data, and division of datasets, etc. [24]. The partition of the dataset is divided according to the evaluation model method validation and cross-validation. When the model method is validation, after selecting the data set, the data is generally divided into three subsets: training set, validation set and test set[25]. The size of the training set accounts for about 70% of the entire data set. T.

Choose the MNIST dataset, which is a well-known machine vision dataset for handwritten digits. There are two ways to obtain the MNIST dataset, one is to download it from Prof. Yann LeCun's official website, and the other is to use the official case of TensorFlow, and the MNIST dataset is included in TensorFlow[26]. The MNIST data set has 60,000 samples, of which 55,000 samples are training sets, and the other 5,000 samples are part of the validation set. The division of the entire dataset is shown in Table 1.

Table 1. Division of the MNIST dataset

Data set	Number of samples	Sample tensor
Training set	55000	55000×84
Test set	10000	10000×784
Validation set	10000	10000×784

In the data set MNIST, each sample contains the gray value information and the corresponding label of the sample. Each image sample is composed of handwritten digits of 28×28 pixels. In order to simplify the model, through dimensionality reduction processing, the two-dimensional 28× 28 images are converted into a one-dimensional vector with 784 features, then the feature of the training set is a [55000,784] tensor, and the features of the test set are [10000,784] and [10000,784] respectively. tensor. The label corresponding to the training data set is a [55000, 10] tensor, where the 55000th means that there are 55,000 sample images in the training set, and 10

https://doi.org/10.17993/3ctecno.2023.v12n1e43.54-68

means that the label of each image sample in the training set is a hot encoding containing 10 types of numbers[27].

2.2.2. BUILD A PRELIMINARY MODEL THAT MEETS THE REQUIREMENTS

The samples in the MNIST data set are 28×28 two-dimensional, and the one-dimensional vector has 784 gray values, which determines that the number of neurons in the input layer of the NN model is 784. The MNIST dataset is a total of 10 categories of handwritten digits from 0 to 9, so output layer is 10. First design a simple NN without hidden layers.

The simple NN with multi-parameter fusion is optimized through experiments, and then the hidden layer is gradually increased on the basis of the NN without hidden layer.

2.2.3. CHOOSE ACTIVATION FUNCTION, LOSS FUNCTION AND OPTIMIZER

1. Activation function

The activation function enables the NN to have the ability of hierarchical nonlinear mapping learning, which can approximate any function and solve more complex problems. The relu function has been applied in the deep learning network. How to choose the activation function, there is no definite method, mainly based on some experience[28]. Several commonly used activation functions are as follows:

(1) sigmoid function. sigmoid is a commonly used nonlinear activation function, and its definition is shown in formula (1):

$$f(z) = 1/\left(1 + \mathrm{e}\,\hat{}\,(-z)\right) \tag{1}$$

The input z is mapped to the range between 0 and 1, but the sigmoid activation function in the DNN will origin the problem of incline explosion and gradient disappearance. The meeting is slow, and the sigmoid function has the disadvantages of exponentiation, which is time-consuming.

(2) tanh function. The problems of gradient disappearance and exponentiation still exist in the deep NN with tanh as the activation function. The analytical expression of the tanh function is shown in formula (2).

$$f(x) = \left(\mathrm{e}\,\hat{}\,x - \mathrm{e}\,\hat{}\,(-x)\right)/\left(\mathrm{e}\,\hat{}\,x + \mathrm{e}\,\hat{}\,(-x)\right) \tag{2}$$

(3) relu function. The analytical appearance of the relu function is revealed in formula (3).

$$\mathrm{relu}(x) = \max(0, x) \tag{3}$$

https://doi.org/10.17993/3ctecno.2023.v12n1e43.54-68

relu is a function that takes the maximum value between partitions. It is not derivable in the entire interval. It only determines the size of the input x and 0.

2. Loss function

The loss function is to estimate the difference between the predicted value ypre = f(x) of the designed NN model and the real value yhat. Usually Loss(yhat, ypre) is used to represent the loss function. Common loss functions are as follows:

(1) 0-1 loss function. The definition of the 0-1 loss function is shown in Equation (4).

$$\mathrm{Loss}(\mathrm{yhat}, \mathrm{ypre}) = \begin{cases} 1, & \mathrm{yhat} \neq y\,\mathrm{pre} \\ 0, & \mathrm{yhat} = \mathrm{ypre} \end{cases} \tag{4}$$

The 0-1 loss function does not consider the difference between the predicted value and the true value. If the prediction is correct, the value of the loss function is 0, otherwise the value of the loss function is 1.

(2) Squared loss function. The definition of the squared loss function is shown in formula (5) and (6).

$$Loss(yhat, ypre) = (yhat - ypre)^2 \tag{5}$$

$$H(p,q) = -\sum_{x} p(x)\log q(x) \tag{6}$$

2.2.4. TRAIN THE MODEL AND EVALUATE THE MODEL

The parameters of the NN model are weight (weight) and bias (threshold). The training model is to repeatedly adjust the weight and bias model parameter values through training samples and learning algorithms, so that the error between the actual output and the ideal output is less, and finally the NN is solved. parameters required by the problem. Among the learning algorithms for training models, the most representative one is the error backpropagation (BP) algorithm, which is widely used in multi-layer feedforward NNs. The methods of evaluating the model include validation and cross-validation, and different evaluation methods also determine the division of the dataset. Common model evaluation indicators for classification problems include confusion matrix, accuracy, precision, recall, specificity, etc., as shown in Table 2(Ying et al.).

Table 2. Model evaluation indicators

Confusion Matrix		Goal		
		True Positives True Negatives	True Positives True Negatives	
Model	predicts positive samples	True Positive（TP）	False Positive（FP）	Positive predictive value or precision =TP/（TP+FP）

https://doi.org/10.17993/3ctecno.2023.v12n1e43.54-68

Predict Negative Samples	False Negative（FN）	True Negative（TN）	阴性预测值=（TN）（/ TN + FN）
	Recall =TP/（TP+FN）	Specificity=TN/（TN+FP）	Precision=（TP+TN）（/ TP+FP+FN+TN）

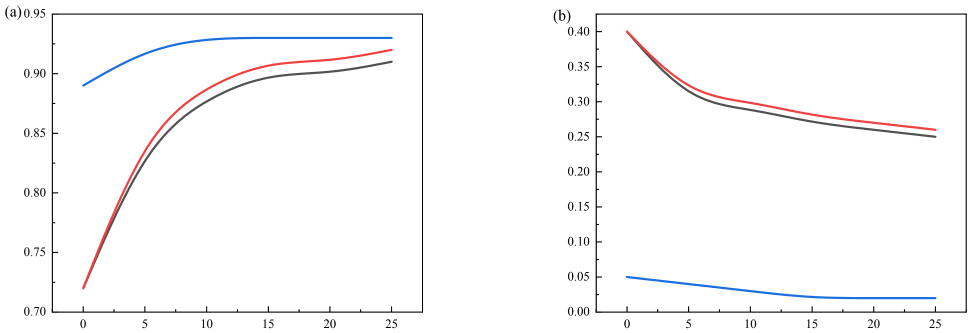
3. EXPERIMENTAL RESULTS AND ANALYSIS

Experiments were tested using TensorFlow on a Windows 10 system with Intel(R) Core(TM) i7- 6700HQ CPU@2.6 GHz 2.59 GHz, 8 GB RAM.

The sample image in the MNIST dataset is converted into a one-dimensional vector with 784 elements, and neurons number in the input layer is strongminded to be 784, and there are 10 categories of handwritten digits. A simple NN without a hidden layer has only an input layer [29].

(1) Comparison of loss functions. The learning rate is set to 0.1, the training model sample batch size is 100, and the number of iterations is 30. The optimizer uses the gradient descent method to compare the recognition accuracy of the cross entropy and the squared loss function on the simple NN.

Figure 1(a) is the curve relationship between the training model and the accuracy. The upper two curves are a group, which is the accuracy curve when the loss function is cross entropy [30-31]. The lower two curves are Accuracy curves for training and test sets when the loss function is a squared loss function. It could be gotten from Figure 1(a) that at iterations number, the training accuracy set is greater than that of the training set and test set when the loss function is the squared loss function. In Figure 1(b), the group with high loss value is the loss value curve of the training set and the test set when the loss function is a squared loss function. The group is the test set when the loss function is cross entropy. Figure 1(b) shows that the convergence speed is fast when the loss function is irritated information. From the comparison of the accuracy and loss value of cross entropy and squared loss function in Figure 1, it can be seen that irritated information is selected as the loss purpose of simple NN.

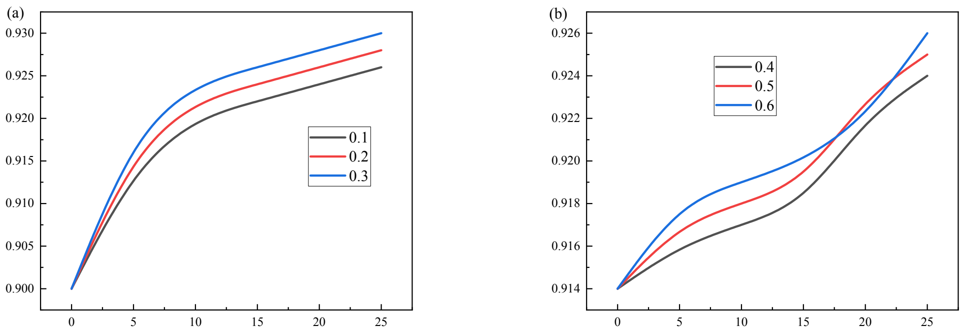


https://doi.org/10.17993/3ctecno.2023.v12n1e43.54-68

Figure 1. Comparison of cross entropy and square loss function, (a) Curve relationship between iteration times and accuracy; (b) Curve relationship between iteration times and loss value.

(2) Different learning rates. The training model sample batch scope is 100, the iterations number is 30, the optimizer adopts the incline ancestry method, and the damage function is cross entropy. Figure 2 shows the contrast of correctness created on different learning rates. The upper row is the test result, and the lower row is the test result of the test set.

Figure 2(a) shows three curves of the accuracy as a function when the learning rates are 0.1, 0.2 and 0.3, respectively. The untried consequences presented with the increase of the learning rate, the performance of the NN gradually improves. In the training set and the test set, when the learning rate is 0.3, the accuracy is almost the highest, and the number of iterations is 30. The accuracy is 93.30% and 92.58, respectively. %. Figure 2(b) is the curve of the accuracy as a function of the number of iterations when the learning rate is 0.4, 0.5 and 0.6, respectively. The experimental results show that the accuracy fluctuates greatly when the learning rate is 0.5 and 0.6 in the training set, especially when the learning rate is 0.5, the accuracy is unstable. When the number of iterations is 30, the training set with a learning rate of 0.6 has the highest accuracy, and a learning rate of 0.5 has the highest accuracy on the test set. Figure 2(c) is the curve of the accuracy versus the number of iterations when the learning rates are 0.3, 0.5 and 0.6, respectively. The experimental results show that the accuracy of different learning rates is not much different whether it is the training set or the test set. The accuracy of the learning rate of 0.5 and 0.6 is slightly higher than that of the learning rate of 0.3, but the accuracy of the learning rate of 0.5 and 0.6 in the training set is slightly higher. The rate fluctuates greatly. It can be seen from the analysis in Figure 2 that the learning rate is between 0.3 and 0.4, so the learning rate is 0.3.



https://doi.org/10.17993/3ctecno.2023.v12n1e43.54-68

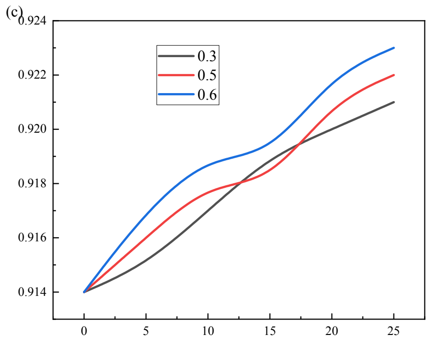
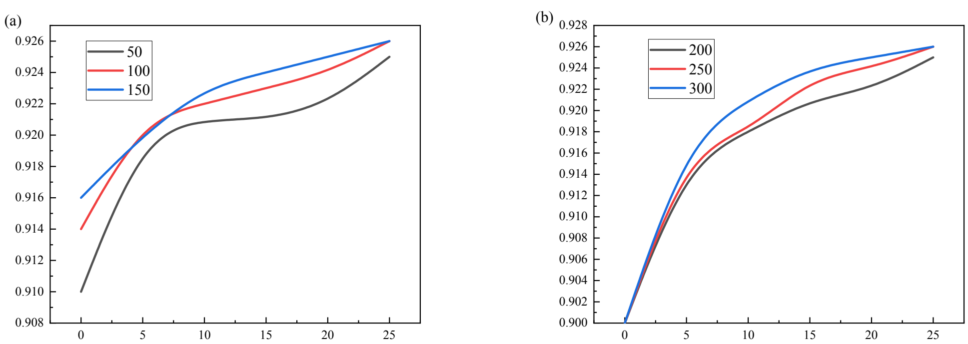


Figure 2. Effect of different learning rates on the accuracy of NN, (a) The learning rates were 0.1, 0.2 and 0.3; (b) The learning rates were 0.4, 0.5 and 0.6; (c) The learning rates were 0.3, 0.5 and 0.6.

(3) Batch size. The learning rate is 0.3, the number of iterations is 30, the optimizer uses gradient descent, and the loss function is cross-entropy to compare the effects of different batch sizes (50, 100, 150, 200, 250, 300) on the NN model. Figure 3 shows the comparison of the accuracy of different batch sizes. The upper row is the test results of the training set with different batch sizes, and the lower row is the test results of the test set with different batch sizes. Figure 3(a) are three plots of accuracy versus number of iterations for batch sizes of 50, 100, and 150. The experimental results show that in the training set, the accuracy rate of batch size 50 is higher than that of batch size 100 and 150, while the accuracy rate of batch size 100 in the test set is higher than that of batch size 50 and 150. Accuracy. Figure 3(b) shows the curve relationship between the accuracy rate and the number of iterations when the batch size is 200, 250 and 300 respectively. In the training set, when the number of iterations is 30 and the batch size is 200, the accuracy rate is 93.03%, which is higher than the batch size of 250 and 200. The accuracy rates corresponding to 300 are 92.89% and 92.85%, and the accuracy is also the highest when the number of iterations in the test set is 30 and the batch size is 200. Figure 3(c) shows the curve relationship of the accuracy rate with the number of iterations when the batch size is 100, 200 and 300 respectively. Whether it is the training set or the test set, the accuracy rate is the highest when the batch size is 100. It can be seen from the analysis in Figure 3 that the model with a batch size of 100 has the best performance.



https://doi.org/10.17993/3ctecno.2023.v12n1e43.54-68

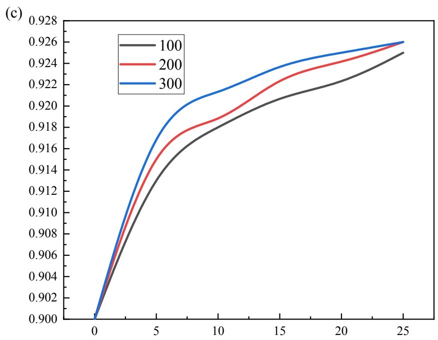


Figure 3. Effect of different batch sizes on the accuracy of NN, (a) Batch sizes 50, 100 and 150; (b) batch sizes 200, 250 and 300; (c) batch sizes 100, 200 and 300.

On the basis of the single hidden layer NN optimized by multi-parameter fusion, a hidden layer is added to continue to optimize the NN model. The sample batch size is 100, the number of iterations is 30, the loss function is cross entropy, the optimizer is the AdaGrad algorithm, there are two hidden layers, the number of neurons in the hidden layer is 500 and 300, and the activation function of the hidden layer is relu, and compare the effects of different learning rates (0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7) on the NN model in the test set. It can be seen from Table 5 that the accuracy of learning rates of 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, and 0.45 is higher than that of learning rates of 0.5, 0.55, 0.6, 0.65, and 0.7. The learning rates of 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, and 0.4 have little difference in accuracy, and their accuracy curves with the number of iterations almost overlap each other, with learning rates of 0.2 and 0.25 having the best performance. The experimental results show that the learning rate of the multi-layer hidden layer NN is set to 0.2. The sample batch size is 100, the number of iterations is 30, the loss function is cross entropy, the learning rate is 0.2, the optimizer is the AdaGrad algorithm, two hidden layers, and the number of neurons in the hidden layer is 500 and 300, respectively, in the test set Compare the effect of different activation functions (sigmoid, relu, selu, and tanh) of the hidden layer on the NN model.

Table 3. Different learning rates in the test set

Number	Accuracy of different learning rates (%)										
	0.10	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.60
1	0.964	0.965	0.965	0.967	0.963	0.960	0.954	0.926	0.234	0.402	0.295
2	0.975	0.974	0.974	0.974	0.973	0.967	0.963	0.950	0.911	0.910	0.906
3	0.977	0.978	0.978	0.976	0.972	0.971	0.968	0.959	0.945	0.944	0.941
5	0.978	0.981	0.981	0.979	0.976	0.976	0.972	0.966	0.961	0.956	0.952
10	0.983	0.983	0.983	0.983	0.980	0.975	0.976	0.970	0.970	0.964	0.957
15	0.983	0.985	0.985	0.985	0.983	0.983	0.977	0.972	0.967	0.968	0.964
20	0.984	0.986	0.986	0.987	0.985	0.984	0.982	0.973	0.975	0.971	0.964
25	0.985	0.987	0.987	0.989	0.988	0.986	0.984	0.977	0.976	0.969	0.966
30	0.986	0.988	0.988	0.990	0.989	0.988	0.986	0.980	0.977	0.971	0.967

https://doi.org/10.17993/3ctecno.2023.v12n1e43.54-68

4. CONCLUSION

Curriculum ideology and politics is an inherent requirement to achieve the goal of "cultivating morality and cultivating people" in colleges and universities, and it is a beneficial exploration to realize the three-round education. In this study, a deep NN model with multi-parameter fusion optimization was constructed and applied to the ideological construction of college professional courses. The following conclusions were drawn: (1) At any number of iterations, when the loss function is cross entropy, the training set and the test set the accuracy of the NN is greater than that of the training set and the test set when the loss function is a squared loss function; (2) With the increase of the learning rate, the performance of the NN gradually improves. In the training set and the test set, when the learning rate is 0.3 The accuracy rates are almost all the highest. The iteration times are 30, and the accuracy rates are 93.30% and 92.58%, respectively. When the iteration times are 30, the training set with a learning rate of 0.6 has the highest accuracy, and a learning rate of 0.5. The test set has the highest accuracy; (3) batch The curve relationship between the accuracy rate and the number of iterations when the number of times is 200, 250 and 300 respectively. In the training set, the number of iterations is 30 and the accuracy rate is 93.03% when the batch size is 200, which is higher than the accuracy rate of 92.89% corresponding to the batch size of 250 and 300. and 92.85%, and the accuracy rate is also the highest when the number of iterations in the test set is 30 and the batch size is 200, and the model with a batch size of 100 has the best performance.

5. CONFLICT OF INTEREST

The authors declared that there is no conflict of interest.

REFERENCES

(1) Min, W. U. (2020). **Research on the Integration of Curriculum Thought and Politics into the Teaching Practice of Business Etiquette in Higher Vocational Education[J]**. *Journal of International Education and Development,* 4(9), 20-24. https://doi.org/10.47297/wspiedWSP2516-250004.20200409

(2) Vickers, E. (2009). **Selling 'Socialism with Chinese Characteristics' 'Thought and Politics' and the legitimisation of China's developmental strategy**. *International Journal of Educational Development*, 29(5), 523-531. https://doi.org/10.1016/j.ijedudev.2009.04.012

(3) Madeleine, Arnot, Jo-Anne, & Dillabough. (1999). **Feminist Politics and Democratic Values in Education. Curriculum Inquiry**. *Special Series on Girls and Women in Education*, 29(2), 159-189, https://doi.org/10.1111/0362-6784.00120

(4) Cakal, H., Hewstone, M., Schw?R, G., & Heath, A. (2011). **An investigation of the social identity model of collective action and the 'sedative' effect of intergroup contact among Black and White students in South Africa**. *British Journal of Social Psychology*, 50(4), 606-627. https://doi.org/10.1111/j.2044-8309.2011.02075.x

(5) Shi, J., Hao, Z., Saeri, A. K., & Cui, L. (2015). **The dual-pathway model of collective action: Impacts of types of collective action and social identity**. *Group Processes & Intergroup Relations*, 18(1), 45-65. https://doi.org/10.1177/1368430214524288

(6) Park, H. S., Gonsier-Gerdin, J., Hoffman, S., Whaley, S., & Yount, M. (1998). **Applying the Participatory Action Research Model to the Study of Social Inclusion at Worksites**. *Journal of the Association for Persons with Severe Handicaps*, 23(3), 189-202. https://doi.org/10.2511/rpsd.23.3.189

(7) Oel, P. V., Mulatu, D. W., Odongo, V. O., Willy, D. K., & Van, D. (2019). **Using Data on Social Influence and Collective Action for Parameterizing a Geographically-Explicit Agent-Based Model for the Diffusion of Soil Conservation Efforts**. *Environmental Modeling & Assessment, (1)*. https://doi.org/10.1007/S10666-018-9638-Y

(8) Zarpour, M. T. (2013). **Discourse and Dissent in the Diaspora: Civic and Political Lives of Iranian Americans**. http://hdl.handle.net/1903/14216

(9) Lickona, T. (1993). **The Return of Character Education**. *Educational Leadership Journal of the Department of Supervision & Curriculum Development N.e.a*, 51(3), 6-11. https://doi.org/10.1177/0013161X93029004010

(10) Hyman, H. (1971). **Political Socialization**. *International Journal of Psychology*, 6. https://doi.org/10.1080/00207597108246696

(11) Robin, A., & Hodgkin. (2006). **Where Law and Order Start: The Genesis of Boundaries and Norms**. *Journal of Moral Education*. https://doi.org/10.1080/0305724820110205

(12) Olssen, M., Codd, J., & O'Neill, A. M. (2004). **Education Policy: Globalization, Citizenship and Democracy**. *SAGE Publications*, https://doi.org/10.4135/9781446221501

(13) Pospíchal, J., & Kvasnička, V. (2015). **70th Anniversary of Publication: Warren McCulloch & Walter Pitts - A Logical Calculus of the Ideas Immanent in Nervous Activity**. *Springer International Publishing*. https://doi.org/10.1007/978-3-319-10783-7_1

(14) Fukushima, K., Miyake, S., & Ito, T. (1988). **Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Visual Pattern Recognition**. *IEEE Transactions on Systems Man and Cybernetics*, SMC-13(5), 826-834. https://doi.org/10.1007/978-3-642-46466-9_18

(15) Hu, C., Yan, Z., Jiang, J., Zhang, S., & Gu, T. (2022). **Traditional Chinese Medicine Information Analysis Based on Multi-task Joint Learning Model**. https://doi.org/10.1007/978-981-16-6963-7_25

(16) Lecun, Y., & Bottou, L. (1998). **Gradient-based learning applied to document recognition**. *Proceedings of the IEEE*, 86(11), 2278-2324. https://doi.org/10.1109/5.726791

(17) Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). **A Fast Learning Algorithm for Deep Belief Nets**. *Neural Computation*, 18(7), 1527-1554. https://doi.org/10.1162/neco.2006.18.7.1527

(18) Lawrence, S., Burns, I., Back, A., Tsoi, A. C., & Giles, C. L. (2012). **Neural Network Classification and Prior Class Probabilities**. https://doi.org/10.1007/978-3-642-35289-8_19

(19) Zhang, P., Wei, L., Wang, H., Lei, Y., & Lu, H. (2018). **Deep Gated Attention Networks for Large-scale Street-level Scene Segmentation**. *Pattern Recognition*, 88. https://doi.org/10.1016/j.patcog.2018.12.021

(20) He, K., Zhang, X., Ren, S., & Sun, J. (2016). **Deep Residual Learning for Image Recognition**. *IEEE*. https://doi.org/10.1109/CVPR.2016.90

(21) Yi, S., Wang, X., & Tang, X. (2014). **Deep Learning Face Representation by Joint Identification-Verification.** *Advances in neural information processing systems*, 27. https://doi.org/10.48550/arXiv.1406.4773

(22) Zhang, S., He, Y., Wei, J., Mei, S., & Chen, K. (2019). **Person Re-identification with Joint Verification and Identification of Identity-Attribute Labels**. *IEEE Access*, 7, 126116-126126. https://doi.org/10.1109/ACCESS.2019.2939071

(23) González-Muñoz, M., Bastida, S., & Sánchez-Muniz, F. (2003). **Short-term in vivo digestibility assessment of a highly oxidized and polymerized sunflower oil**. *Journal of the Science of Food and Agriculture*, 83(5), 413-418. https://doi.org/10.1002/jsfa.1383

(24) Canziani, A., Paszke, A., & Culurciello, E. (2016). **An Analysis of Deep Neural Network Models for Practical Applications**. https://doi.org/10.48550/arXiv.1605.07678

(25) A, C. W., A, X. W., Jz, A., Liang, Z. A., Xiao, B. A., Xin, N. B., Ehd, A. (2021). **Uncertainty Estimation for Stereo Matching Based on Evidential Deep Learning**. https://doi.org/10.1016/j.patcog.2021.108498

(26) Cai, W., Zhai, B., Liu, Y., Liu, R., & Ning, X. (2021). **Quadratic polynomial guided fuzzy C-means and dual attention mechanism for medical image segmentation**. *Displays*, 70, 102106. https://doi.org/10.1016/j.displa.2021.102106

(27) Miao, J., Wang, Z., Ning, X., Xiao, N., Cai, W., & Liu, R. (2022). **Practical and secure multifactor authentication protocol for autonomous vehicles in 5G**. *Software: Practice and Experience*. https://doi.org/10.1002/SPE.3087

(28) Ning, X., Gong, K., Li, W., & Zhang, L. (2021). **JWSAA: joint weak saliency and attention aware for person re-identification**. *Neurocomputing*, 453, 801-811. https://doi.org/10.1016/j.neucom.2020.05.106

(29) Yu, Z., Li, S., Sun, L. N. U., Liu, L., & Haining, W. **Multi-distribution noise quantisation: an extreme compression scheme for transformer according to parameter distribution**. https://doi.org/10.1080/09540091.2021.2024510

(30) Medina, R., Breña, J. L., y Esenarro, D. (2021). **Efficient and sustainable improvement of a system of production and commercialization of Essential Molle Oil (Schinus Molle)**. *3C Empresa. Investigación y pensamiento crítico*, 10(4), 43-75. https://doi.org/10.17993/3cemp.2021.100448.43-75

(31) Xiong Xiaokun. (2022). **Research on tourism income index based on ordinary differential mathematical equation**. *Applied Mathematics and Nonlinear Sciences, 7(1), 653-660*. https://doi.org/10.2478/AMNS.2021.2.00113.