Hindawi

Research Article

Research on the Guidance of Youth Labor Education Based on the "Combination of Education and Production Labor" Program Based on the Deep Learning Model

Linyu Xiao [1] **and Xiaoyi Liao** [2]

^{1}*Hunan Agricultural University, Institute of Marxism, Changsha 410128, Hunan, China*
^{2}*Hunan Agricultural University, College of Horticulture, Changsha 410128, Hunan, China*

Correspondence should be addressed to Xiaoyi Liao; liaoxiaoyi@hunau.edu.cn

At present, there is a lack of research on Marx's idea of "combining education and productive labor" and its guiding significance for youth labor education, and no effective teaching model has been formed. In response to this problem, this study proposes a semi-supervised deep learning model based on u-wordMixup (SD-uwM). When there is a shortage of labeled samples, semi-supervised learning uses a large number of unlabeled samples to solve the problem of labeling bottlenecks. However, since the unlabeled samples and labeled samples come from different fields, there may be quality problems in the unlabeled samples, which makes the generalization ability of the model worse., resulting in a decrease in classification accuracy. The model uses the u-wordMixup method to perform data augmentation on unlabeled samples. Under the constraints of supervised cross-entropy and unsupervised consistency loss, it can improve the quality of unlabeled samples and reduce overfitting. The comparative experimental results on the AGNews, THUCNews, and 20Newsgroups data sets show that the proposed method can improve the generalization ability of the model and also effectively improve the time performance. The study found that the SD-uwM model uses the u-wordMixup method to enhance the unlabeled samples and combines the idea of the Mean Teacher model, which can significantly improve the text classification performance. The SD-uwM model can improve the generalization ability and time performance of the model, respectively, 86.4 ± 1.3 and 90.5 ± 1.3. Therefore, the use of SD-uwM in Marx's program is of great practical significance for the guidance process of youth labor education.

1. Introduction

From the standpoint of the proletariat, Marx opposed the exploitation and oppression of workers by the bourgeoisie, combined with the actual development of the modern machinery industry to construct a theoretical system combining productive labor and education, and demonstrated its inevitability in the process of social development [1]. This theory also had a profound impact on Marx's educational thought and provided important guidance on how the proletariat should carry out educational activities. In our country, it is very necessary to excavate the connotation of Marx's thought of "combining education with productive labor," which has theoretical value and practical significance. This thought of Marx can guide our country to educate and cultivate laborers with professional quality and promote the improvement of productivity, thereby promoting social development [2]. Since the 18th National Congress of the Communist Party of China, General Secretary Xi Jinping has delivered many important speeches on the topic of labor based on Marx's discussion on labor. He highly recognized the important position of labor in socialist countries, and also clarified the importance of cultivating young people's correct labor values and loving physical labor in the new era. General Secretary Xi Jinping clearly stated at the 2018 National Education Conference: "We must educate and guide students to advocate and respect labor." [3] In today's society, we must attach importance to labor

education and give full play to the educating function of physical labor and use Marxist labor Ideology arm the minds of young people, guide them to love labor, and train young people into laborers who meet the needs of social development in practice, so that they can better serve my country's modernization drive. At the same time, the "Opinions on Comprehensively Strengthening Labor Education in Colleges and Elementary Schools in the New Era" issued by my country pointed out that labor education is an indispensable part and an important part of education in the education system of socialist countries. We must pay attention to the unique educational value of labor and strengthen modern labor education [4]. Through classroom teaching, labor practice, etc., effective measures are taken to change the attitude of young people towards labor and to curb the phenomena of extravagance, lack of progress, and greed for pleasure in the campus. In this context, it is more necessary to study Marx's thought of "combining education and productive labor" and its guiding significance for youth labor education.

The famous Russian educator Ushinsky believes that education is an important hand to guide students to respect and love labor, and it is also an indispensable part of cultivating students' labor habits [5, 6] Lenin attached great importance to the combination of education and productive labor. He inherited and developed Marx's idea and pointed out that the ideal realization of the future society cannot be separated from the combination of education and productive labor of the new generation. In a socialist country, labor courses should be an indispensable course in school education. Students can master labor skills by participating in voluntary labor and combine practice and theory to help themselves better master scientific knowledge and become socialist laborers [7]. Makarenko believes that labor has an important impact on a person's future development. Individual labor can improve one's living standard and enhance happiness; in collective labor, people help each other and can establish good interpersonal relationships [8]. In the book "Dedicate the whole heart to the child," Sukhomlinsky said that the beauty of human beings is the brightest in labor, so an important task of education is to make the child's surroundings full of the natural world and through labor a world created and built. Labor has the function of aesthetic education, and schools should make labor a need for students' spiritual life and become a powerful educational force on this basis. Lenin, Makarenko, and Sukhomlinsky all stood from the perspective of the proletariat and believed that the ultimate goal of labor education was to train laborers for a communist society. The discussion of this topic by British scholars can be traced back to the early utopian socialist Thomas More who required children to learn agricultural knowledge in school, go to work in the fields, and required every youth to learn at least one kind of handicraft. Owen also mentioned that people will form their own characters during labor, so labor education should be paid attention to in childhood [9]. Locke believes that labor can make people physically and mentally healthy and overcome many vices. Students can choose one of crafts, gardening, or agriculture to practice [10]. German educator Comenius said in "The Great Teaching Theory" that when

realizing the ideal of education through labor, attention should be paid to adapting to the natural order, adapting to human nature and age characteristics. Pestalozzi, a Swiss educator, was the first educator to put the idea of combining education and labor into practice in the history of Western education. He believed that this method could coordinate and develop people's abilities in all aspects and advocated the integration of education and labor [11].

Through the consistency training framework [12], based on the wordMixup data enhancement method, this paper proposes a u-wordMixup (unlabeled sample word mixture) data enhancement method for unlabeled samples. Under the constraint of loss, the enhancement generates high-quality additional training samples and reduces overfitting. Based on the u-wordMixup method, a new semi-supervised deep learning model (SD-uwM) is proposed using the Mean Teacher model for consistent training. The u-wordMixup data augmentation method enhances unlabeled samples with the goal of reducing unsupervised consistency loss, constrains the quality of unlabeled training samples, and reduces model overfitting. The objective loss function combines supervised cross-entropy loss and unsupervised consistency loss and uses the MeanTeacher method for consistency training to improve the generalization ability of the model. Optimize the key words of the "Combination of Education and Productive Labor" program to achieve the optimal application of youth education.

2. Principles and Methods

2.1. wordMixup Data Enhancement. wordMixup is a data enhancement method for labeled samples. The idea is to interpolate the word embedding vectors of two samples to generate a new sample word embedding matrix as an enhanced sample [13]. Given a pair of labeled samples (x_i, y_i) and (x_j, y_j), perform word embedding to get (x_i, y_i) and (x_j, y_j), where $x_i \in \text{RN} \times d$, $x_j \in \text{RN} \times d$ is the word embedding matrix of text x_i and x_j, N is the number of words, d is the word vector dimension, and y_i and y_j are the corresponding class labels. Then perform interpolation according to formulas (1) and (2) to obtain a new sample (x_{ij}, y_{ij}), where x_{ij} is the word embedding matrix of the enhanced sample, and y_{ij} is its class label.

$$\tilde{\mathbf{x}}_{ij}^k = \lambda \mathbf{x}_i^k + (1 - \lambda)\mathbf{x}_j^k, \quad k = 1 \dots N, \tag{1}$$

$$\tilde{\mathbf{y}}_{ij} = \lambda \mathbf{y}_i + (1 - \lambda)\mathbf{y}_j. \tag{2}$$

Among them, x_i^k and x_j^k represent the word vector of the kth word in the text x_i and x_j, respectively, $\lambda \in [0, 1]$ is the interpolation weight factor, and x_{ij}^k is the word vector of the kth word of the new sample generated by interpolation. Perform word vector interpolation on each word in (x_i, y_i) and (x_j, y_j) one-to-one to obtain the embedding matrix x_{ij} of the new sample, where y_{ij} is the class label of x_{ij}, and (x_{ij}, y_{ij}) is the enhanced additional training samples.

The wordMixup method achieves good results in supervised text classification. But unlabeled in semi-supervised learning has no labels, how to generate pseudo labels for its

interpolation? To this end, based on wordMixup, this paper proposes an improved data enhancement method u-word-Mixup for unlabeled samples, which will be introduced in the following chapters [14].

2.2. SD-uwM Semi-Supervised Deep Learning Model

2.2.1. SD-uwM Model. The framework of the semi-supervised deep learning model (SD-uwM) is shown in Figure 1 [15]. It uses the idea of the Mean Teacher model to construct a teacher model T and a student model S, using labeled samples and unlabeled samples, based on supervised cross-entropy loss and unsupervised consistency loss objective functions for semi-supervised deep learning [16].

As shown in Figure 1, $DL = \{(xli, yli)\}Nl\,i = 1$ represents the labeled training sample set, $DU = \{(xui)\}Nu\,i = 1$ represents the unlabeled training sample set, Nl represents the labeled sample set number, and Nu represents the number of unlabeled samples. T is the teacher model, S is the student model, and T and S have the same network structure. LS stands for supervised loss, LUS stands for unsupervised loss, and L stands for consistency objective loss function [17].

SD-uwM trains and learns simultaneously on labeled samples and unlabeled samples based on a consistent target loss function L [18]. As shown in the left half of Figure 1, the student model S computes a supervised cross-entropy loss LS on DL. At the same time, as shown in the right half of Figure 1, the student model S uses the u-wordMixup method to enhance the unlabeled samples. According to the prediction of the enhanced samples by the student model S and the prediction of the unlabeled samples by the teacher model T, the calculation is unsupervised and consistent. The performance loss LUS, LS and LUS together constitute the consistency target loss function L of the SD-uwM model. After many iterations, the SD-uwM model is trained to output the parameters of the student model S as the parameters of the final classification model [19].

2.2.2. u-wordMixup Data Enhancement. Using the Mean Teacher model idea and consistency training [20], the target loss function L of SD-uwM takes into account both supervised cross-entropy loss and unsupervised consistency loss and is defined as shown in

$$L = L_S + \beta L_{US}, \tag{3}$$

where LS is the supervised cross-entropy loss on the labeled samples DL, LUS is the unsupervised consistency loss on the unlabeled samples DU, and β is the scale coefficient. LS is the supervised loss of the student model S on the labeled sample set DL, which is calculated as

$$L_S = E_{\mathbf{x}_i^l, y_i^l \in D_L}\left[-y_i^l \log\, p_{\theta'}\left(\mathbf{x}_i^l\right)\right]. \tag{4}$$

Among them, y_i^l is the true label of the labeled sample x_i^l, θ' represents the parameter of the student model S, and $p_\theta'(x_i^l)$ represents the predicted pseudo-label of the sample x_i^l by the student model S, that is, y_i^l.

Based on wordMixup, a u-wordMixup data augmentation method for unlabeled samples is proposed as part of the SD-uwM model [21]. Different from the wordMixup method, the interpolation operation object of the u-word-Mixup method has no real class label. The u-wordMixup method is shown in Figure 2 [22].

As shown in Figure 2, x_i^u and x_j^u represent two unlabeled samples, where $x_i^u \in \mathrm{RN} \times d$, $x_j^u \in \mathrm{RN} \times d$, N is the number of words, and d is the word vector dimension. The feature interpolation of the word vector is performed for each word in x_i^u and x_j^u one-to-one, and the embedding matrix $\tilde{x}_{ij}^u$ of a new unlabeled sample is obtained as an additional training sample [23]. The teacher model T predicts the unlabeled samples x_i^u and x_j^u to generate pseudo-labels $\hat{y}_i^u$ and $\hat{y}_j^u$ and performs pseudo-label interpolation on $\hat{y}_i^u$ and $\hat{y}_j^u$ to obtain $\tilde{y}_{ij}^u$, which is used as the pseudo-label of the enhanced sample $\tilde{x}_{ij}^u$. Then the student model S predicts the enhanced sample $\tilde{x}_{ij}^u$ to get the predicted label $\hat{y}_{ij}^u$. Among them, the calculation of feature interpolation and pseudo-label interpolation are as follows:

$$\tilde{\mathbf{x}}_{ij}^{uk} = q_\lambda\left(\mathbf{x}_i^u, \mathbf{x}_j^u\right) = \lambda\mathbf{x}_i^{uk} + (1 - \lambda)\mathbf{x}_j^{uk}, \quad k = 1\ldots N, \tag{5}$$

$$\tilde{\mathbf{y}}_{ij}^u = q_\lambda\left(\hat{\mathbf{y}}_i^u, \hat{\mathbf{y}}_j^u\right) = \lambda\hat{\mathbf{y}}_i^u + (1 - \lambda)\hat{\mathbf{y}}_j^u. \tag{6}$$

Among them, $q\lambda\,(x_i^u, x_j^u)$ represents the u-wordMixup data enhancement transformation, x_{ki}^u and x_{kj}^u represent the word vector of the kth word of the samples x_i^u and x_j^u, respectively, $\lambda \in [0, 1]$ is the interpolation weight factor, and $\tilde{x}_{kij}^u$ is the interpolation. The word vector of the kth word of the generated augmented sample. $\hat{y}_i^u$ is the predicted pseudo-label of x_i^u by the teacher model T, $\hat{y}_j^u$ is the predicted pseudo-label of x_j^u by the teacher model T, and $\tilde{y}_{ij}^u$ is the pseudo-label generated by interpolation, that is, the pseudo-label of $\tilde{x}_{ij}^u$.

Based on consistent training, the pseudo-label predicted by the student model S for the enhanced unlabeled sample $\tilde{x}_{ij}^u$ is $\hat{y}_{ij}^u$, which should be as consistent as possible with the pseudo-label $\tilde{y}_{ij}^u$ generated by interpolation, that is, $\tilde{y}^u{}_{ij} \approx \hat{y}^u{}_{ij}$, and ideally the two are equal. Therefore, the unsupervised consistency loss LUS is computed as

$$L_{US} = E_{\mathbf{x}_i^u, \mathbf{x}_j^u \in D_U} E_{\tilde{\mathbf{x}}_{ij}^u \sim q_\lambda\left(\mathbf{x}_i^u, \mathbf{x}_j^u\right), \lambda \in [0,1]}$$
$$\left[-\left(\lambda p_\theta\left(\mathbf{x}_i^u\right) + (1 - \lambda)p_\theta\left(\mathbf{x}_j^u\right)\right)\log\, p_{\theta'}\left(\tilde{\mathbf{x}}_{ij}^u\right)\right], \tag{7}$$

where θ denotes the parameters of the teacher model T, θ' denotes the parameters of the student model S, and θ is the moving average of θ'. $p_\theta(x_i^u)$ represents the predicted pseudo-label of the sample x_i^u by the teacher model T, $p_\theta(x_j^u)$ represents the predicted pseudo-label of the sample x_j^u by the teacher model T, and $p_{\theta'}(\tilde{x}_{ij}^u)$ represents the prediction of the newly generated sample $\tilde{x}_{ij}^u$ by the student model S Pseudo tags. $\lambda \in [0, 1]$ is the interpolation weight factor, and $q\lambda(x_i^u, x_j^u)$ represents the u-wordMixup data augmentation transformation [24].
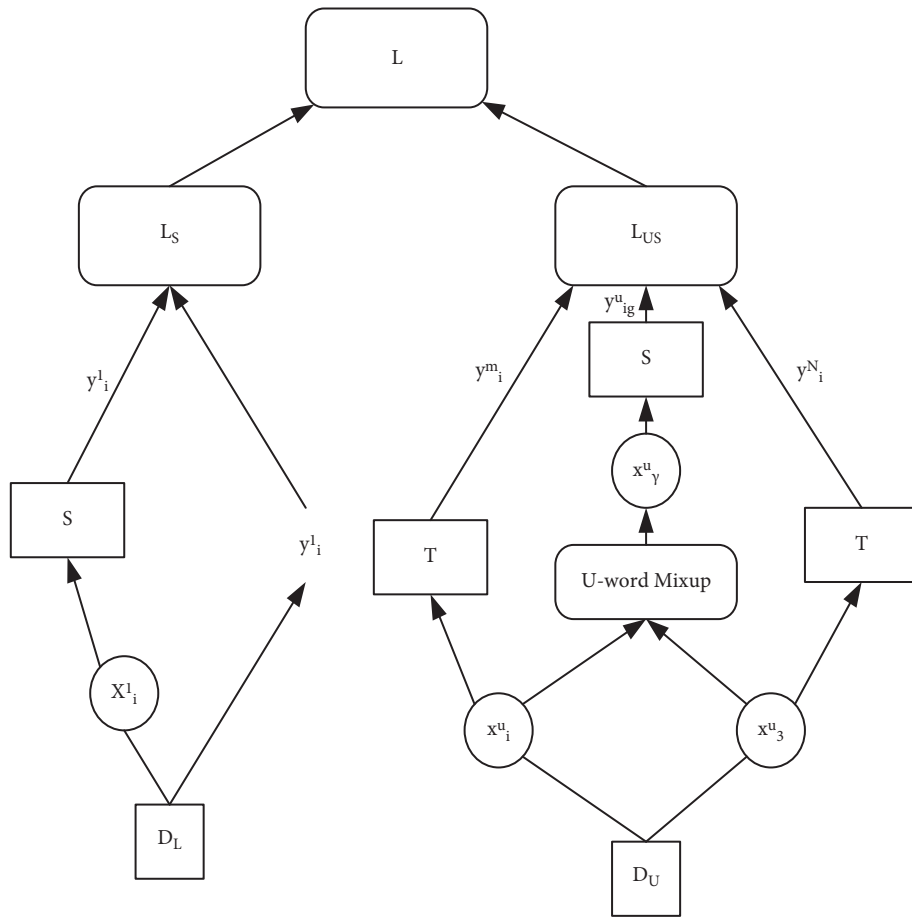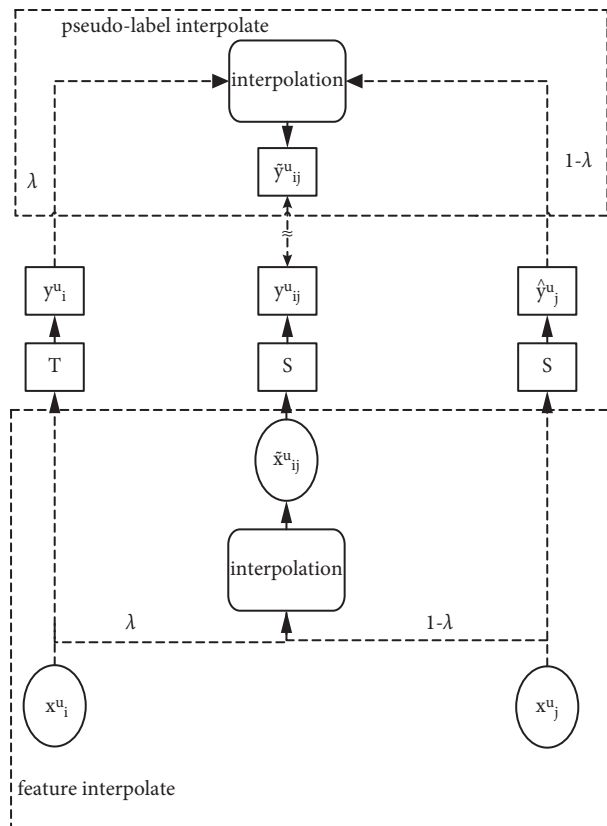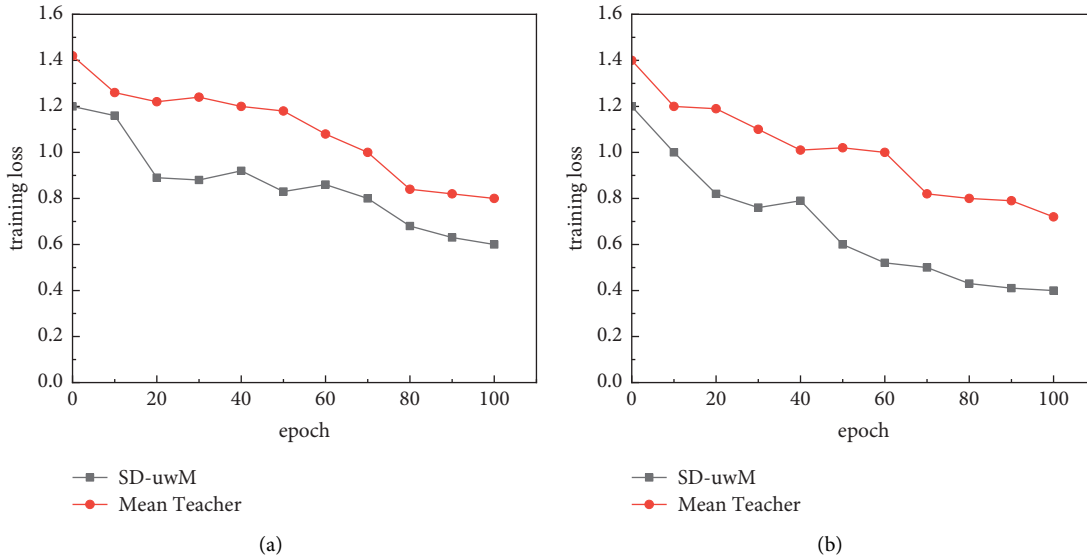
Figure 1: SD-uwM model.

Figure 2: u-wordMixup method.

FIGURE 3: Comparison of training loss between SD-uwM and Mean Teacher when TextCNN is selected: (a) Nl = 200, Nu = 3000 (THUCNews data set); (b) Nl = 300, Nu = 5000 (AGNews data set).

The deep semi-supervised learning SD-uwM model uses the u-wordMixup method to enhance the unlabeled samples. The unsupervised consistency loss LUS aims to reduce the consistency loss and constrain the quality of the enhanced unlabeled samples. Combined with the Mean Teacher model, the teacher model T and the student model S are constructed, and the labeled samples and unlabeled samples are trained. The weighted summation of LS and LUS is used as the objective function L of the model SD-uwM [25].

2.2.3. SD-uwM Model Application. The algorithm description of SD-uwM is shown in Algorithm 1 [26]. The objective function L of the algorithm takes into account the supervised cross-entropy loss LS and the unsupervised consistency loss LUS, and the constraint enhancement generates unlabeled training samples [27]. In each iteration, according to the objective function L, the parameter θ' of the student model S is optimized. After many iterations, the parameter θ' of the optimal student model S is finally obtained.

3. Test Analysis

3.1. Data Set. This article selects three data sets: AGNews, 20Newsgroups and THUCNews. AGNews selects four categories of "world," "politics," "education" and "labor," and 20Newsgroups selects "alt.atheism," "soc.religion.Christian," "comp.graphics," and "sci.med" 4 categories, THUCNews selects 4 categories of "Education," "Labor," "Program," and "Technology." The scaling factor β is set to 1 in Algorithm 1.

3.2. Analysis of Experimental Results. The comparison method in the experiment is as follows:

(1) SD-uwM: the semi-supervised deep learning model based on u-wordMixup data augmentation proposed in this paper

(2) wM-SL: a supervised text classification method based on wordMixup data augmentation method in the literature

(3) SL: supervised text classification method without data augmentation

(4) Mean Teacher: semi-supervised method applied to image classification in the literature, modified for text classification tasks

3.2.1. Comparison of Target Loss between SD-uwM and Mean Teacher. In order to verify the effectiveness of the u-wordMixup method, a comparative experiment was conducted on SD-uwM and Mean Teacher models, and the training loss changes are shown in Figure 3.

As can be seen from Figure 3, compared to the Mean Teacher model, the SD-uwM model using the u-wordMixup data augmentation method has lower training loss. This is because the SD-uwM model objective loss function is more realistic, in which LUS is combined with the u-wordMixup method, aiming to reduce the unsupervised consistency loss, which can improve the quality of unlabeled samples, thereby improving the performance of the model.

3.2.2. Classification Comparison between SD-uwM Model and Other Methods

(1) Comparison of the Accuracy of SD-uwM Model with Other Methods. SD-uwM model and Mean Teacher, wM-SL, SL model on AGNews(Nl = 300, Nu = 5000),

TABLE 1: Classification comparison of four models on three classification results.

Model	Network structure	Accuracy		
		AGNews	THUCNews	20Newsgroups
SL		75.4 ± 1.1	77.5 ± 1.3	71.5 ± 1.3
wM-SL	LSTM	80.4 ± 1.3	83.2 ± 1.2	75.4 ± 1.3
Mean Teacher		82.1 ± 1.3	86.1 ± 1.3	77.5 ± 1.1
SD-uwM		90.4 ± 1.2	91.4 ± 1.3	85.4 ± 1.2
SL		76.4 ± 1.2	78.4 ± 1.4	71.2 ± 1.2
wM-SL	TextCNN	80.5 ± 1.2	84.5 ± 1.3	75.3 ± 1.2
Mean Teacher		83.6 ± 1.1	86.1 ± 1.5	78.1 ± 1.3
SD-uwM		91.2 ± 1.3	92.2 ± 1.3	86.2 ± 1.1

THUCNews(Nl = 300, Nu = 5000), and 20 Newsgroups (Nl = 200, Nu = 2000) The comparative experimental results are shown in Table 1.

As can be seen from Table 1, on the three data sets, regardless of whether the network structure is LSTM or TextCNN, the classification accuracy of the SD-uwM model is better than that of the SL, wM-SL, and Mean Teacher models, up to 90.4. The SL model is a supervised learning method, which requires a large number of labeled samples to achieve better performance. The wM-SL model only enhances the labeled samples, and the Mean Teacher model does not use the u-wordMixup method to enhance the samples. The SD-uwM model uses the u-wordMixup method to enhance the data of unlabeled samples and uses the unsupervised consistency loss constraint to improve the generalization ability of the model.

(2) Comparison of Classification Performance of SD-uwM Model with Increasing Iterations. The number of labeled samples and the number of unlabeled samples are fixed. As the number of iterations increases, the changes in the Macro-F1 value of the SD-uwM model and the SL, wM-SL, and Mean Teacher models are compared and analyzed. On AGNews (Nl = 300, Nu = 5000), THUCNews (Nl = 300, Nu = 5000), and 20Newsgroups (Nl = 200, Nu = 2000), the experimental results are shown in Figures 4 and 5.

It can be seen from Figures 4 and 5 that with the increase of the number of iterations, although the indicators of the SL, wM-SL, Mean Teacher, and SD-uwM models generally show an upward trend and converge to a certain upper limit, the SD-uwM model classification. The performance is significantly better than SL, wM-SL, and Mean Teacher models. As shown in Figure 5(a), using LSTM on AGNews, compared with Mean Teacher, wM-SL, and SL, the Macro-F1 of SD-uwM reaches 90.3%, which is increased by 8%, 9.9%, and 14.5%, respectively. This is because SD-uwM uses the u-wordMixup method to target unsupervised consistency loss for unlabeled sample enhancement, which can reduce overfitting and improve classification performance.

As can be seen from the above Figure 5, in order to verify the influence of unlabeled samples on the SD-uwM model, the number of labeled samples Nl = 300 is fixed on AGNews and THUCNews, the number of labeled samples Nl = 200 is fixed on 20Newsgroups, and the number of unlabeled samples is constantly increasing. Compare the classification results of SD-uwM model with SL, wM-SL, and Mean Teacher models. It can be seen that the indicators of SD-uwM and Mean Teacher show an upward trend with the increase of unlabeled samples, but the classification results of SD-uwM are significantly better than that of Mean Teacher, wM-SL, and SL. Using LSTM on THUCNews, compared with Mean Teacher, wM-SL, and SL, the Macro-F1 of SD-uwM has reached 91.4%, an increase of 5.3%, 8.2%, and 13.9%, respectively. It can be seen that since the SD-uwM model uses the u-wordMixup method to enhance the unlabeled samples and combines the idea of the Mean Teacher model, it can improve the text classification performance.

3.2.3. Time Performance Analysis of the SD-uwM Model. When selecting unlabeled samples, the usual semi-supervised learning algorithm needs to calculate the similarity matrix between the unlabeled samples and the labeled samples, which will increase the time complexity. The SD-uwM model in this paper is random sampling, and there is no need to calculate the similarity between the two samples. The temporal performance comparison between SD-uwM model and typical semi-supervised learning method Co-training is shown in Table 2.

As can be seen from Table 2, the classification accuracy of the SD-uwM model is 86.4% ± 1.3 and 90.5% ± 1.3 in the two different data sets, respectively, and the classification accuracy of the SD-uwM model is significantly higher than that of the co-training model. The classification accuracy was 83.3% ± 1.2 and 88.4% ± 1.2, respectively. At the same time, the time performance of SD-uwM model is significantly better than that of co-training, in which SD-uwM model time is basically maintained at 0.01–0.02 s, while co-training model has exceeded 30 s. The reason is that when selecting unlabeled training samples, the SD-uwM model is random sampling, and the
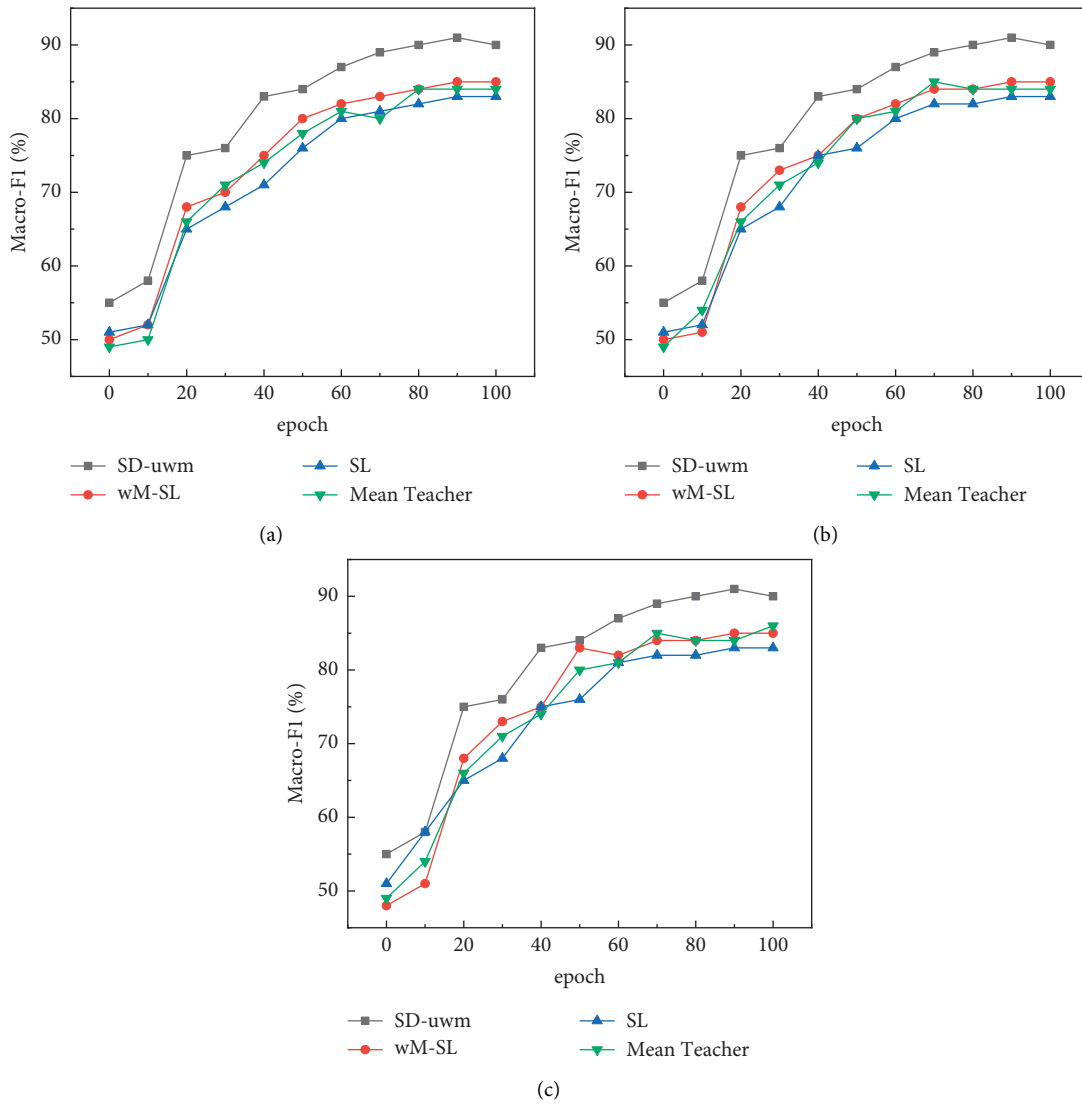
(a)

(b)

(c)

FIGURE 4: Comparison of macro-F1 value of each model with iteration times on three data sets using LSTM: (a) AGNews data set; (b) THUCNews data set; (c) 20Newsgroups data set.
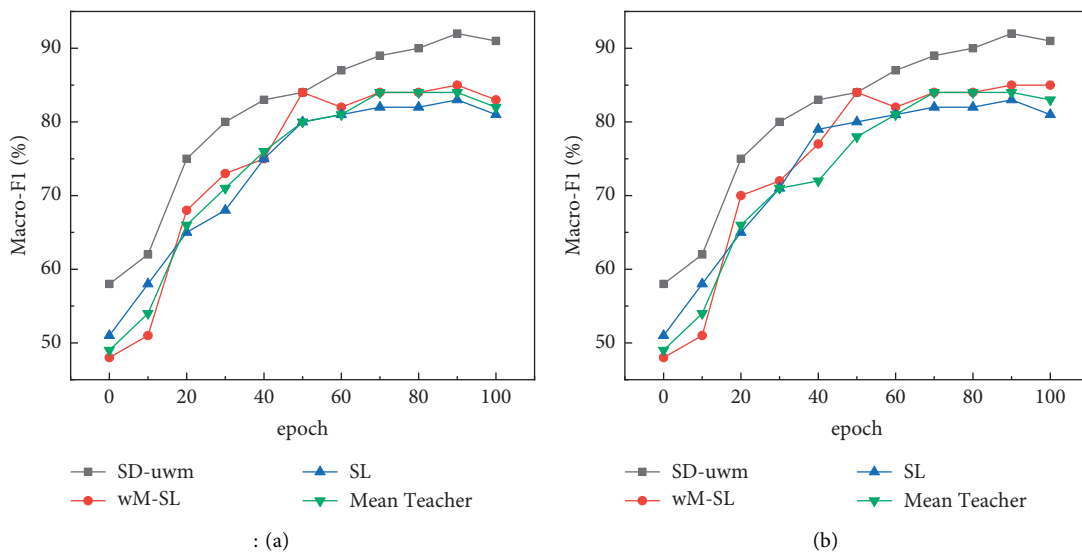

: (a)

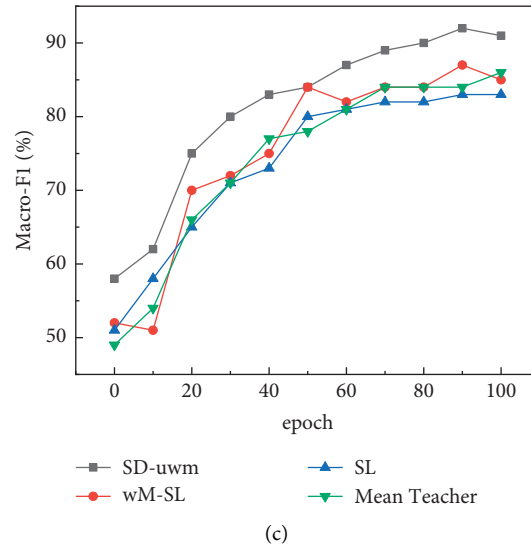(b)

FIGURE 5: Continued.

(c)

FIGURE 5: Comparison of the macro-F1 value of each model with the number of iterations on three data sets using TextCNN: (a) AGNews data set; (b) THUCNews data set; (c) 20 Newsgroups data set.

TABLE 2: Time performance comparison of SN-uwM and co-training.

Dataset	Model	Accuracy (%)	Time (s)	Ratio of train time (SD-uwM/co-training)
20 Newsgroups	SD-uwM	86.4 ± 1.3	0.01	
Nl = 200, Nu = 2000	Co-training	83.3 ± 1.2	30	1/3000
THUCNews	SD-uwM	90.5 ± 1.3	0.02	
Nl = 300, Nu = 4000	Co-training	88.4 ± 1.2	44	1/2200

time complexity is O(1), while the co-training method needs to calculate the sample similarity matrix, and the time complexity is O(Nl ∗ Nu).

4. Conclusion

It is extremely necessary to study Marx's thought of "combining education with productive labor" and its guiding significance to youth labor education. This paper proposes a u-wordMixup method for the data augmentation of unlabeled samples and combines the consistent training framework and the Mean Teacher method to propose a semi-supervised deep learning model SD-uwM. The model uses the u-wordMixup method to enhance the data of unlabeled samples and takes into account the supervised cross-entropy loss and unsupervised consistency loss to construct a new objective function, so as to realize the teaching optimization of the combination of education and production labor. The findings of the study show that the experimental days are as follows:

(1) On the three data sets, regardless of whether the network structure is LSTM or TextCNN, the classification accuracy of the SD-uwM model is better than that of the SL, wM-SL, and Mean Teacher models

(2) SD-uwM model can improve the generalization ability and time performance of the model, which are 86.4 ± 1.3 and 90.5 ± 1.3, respectively

(3) Since the SD-uwM model uses the u-wordMixup method to enhance the unlabeled samples and combines the idea of the Mean Teacher model, it can improve the performance of text classification

Data Availability

The data set can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] G. S. Kostiuk, "Psychological Problems in the Combination of School Learning and Productive Work," *Russian Education & Society*, pp. 33–45, 2014.

[2] P. P. K. Chan, K. Chen, L. Xu, X. Hu, and D. S. Yeung, "Weakly supervised semantic segmentation with patch-based metric learning enhancement," in *Proceedings of the 30th International Conference on Artificial Neural Networks*, vol. 12893, pp. 471–482, Bratislava, Slovakia, September 2021.

[3] P. R. Agenor, "INFRASTRUCTURE, public education and growth with congestion costs," *Bulletin of Economic Research*, vol. 64, no. 4, pp. 449–469, 2012.

[4] S. Cunningham and T. D. Kendall, "Prostitution, hours, job amenities and education," *Review of Economics of the Household*, vol. 15, no. 4, pp. 1055–1080, 2017.

[5] A. İşçan, A. Ece, N. Vurgun, and F. Özcan, "Breast feeding habits of the labour-families in aydın," *The Journal of Tepecik Education and Research Hospital*, vol. 6, no. 1, pp. 97–103, 1996.

[6] L. Gori and M. Sodini, "Local and global bifurcations in an economic growth model with endogenous labour supply and multiplicative external habits," *Chaos*, vol. 24, no. 1, pp. 013122–013142, 2014.

[7] Z. Gao, "From student-centered pedagogy to student labor: Chinese education's transnational entanglements during the cold war," pp. 255–283, 2021.

[8] D. T. Yang, "Education in production: measuring labor quality and management," *American Journal of Agricultural Economics*, vol. 79, no. 3, pp. 764–772, 1997.

[9] A. Tarvainen and H. Valpola, "Mean teachers are better role models: weight-averaged consistency targets improve semi-supervised deep learning results," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, CA, USA, December 2017.

[10] T. Miyato, S. I. Maeda, M. Koyama, and S. Ishii, "Virtual adversarial training: a regularization method for supervised and semi-supervised learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 8, pp. 1979–1993, 2019.

[11] Y. Wang, "Implications of blended teaching based on theory of semantic wave for teaching English writing in high school," *Journal of Higher Education Research*, vol. 3, no. 2, pp. 166–168, 2022.

[12] Y. Tokozume, Y. Ushiku, and T. Harada, "Between-class learning for image classification," 2017, https://arxiv.org/abs/1711.10284.

[13] J. Miao, Z. Wang, X. Ning, N. Xiao, W. Cai, and R. Liu, "Practical and secure multifactor authentication protocol for autonomous vehicles in 5G. Software: practice and Experience," *Concurrency and Computation: Practice & Experience*, vol. 26, no. 4, pp. 987–1019, 2014.

[14] W. Chen, W. Xiang, A. Jz et al., "Uncertainty estimation for stereo matching based on evidential deep learning," *Pattern Recognition*, vol. 124, Article ID 108498, 2022.

[15] N. Papernot, M. Abadi, L. Erlingsson, I. Goodfellow, and K. Talwar, "Semi-supervised knowledge transfer for deep learning from private training data," 2016, https://arxiv.org/abs/1610.05755.

[16] Z. Yu, S. Li, L. N. U. Sun, L. Liu, and W. Haining, "Multi-distribution noise quantisation: an extreme compression scheme for transformer according to parameter distribution," *Connection Science*, .

[17] X. Ning, K. Gong, W. Li, and L. Zhang, "JWSAA: joint weak saliency and attention aware for person re-identification," *Neurocomputing*, vol. 453, pp. 801–811, 2021.

[18] C. Li, J. Zhu, and B. Zhang, "Max-margin Deep Generative Models for (Semi-)Supervised Learning," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 1, no. 99, 2016.

[19] J. Song, L. Gao, F. Zou, Y. Yan, and N. Sebe, "Deep and fast: deep learning hashing with semi-supervised graph construction," *Image and Vision Computing*, vol. 55, pp. 101–108, 2016.

[20] G. Yu, T. Xie, C. Xu et al., "Accurate recognition of colorectal cancer with semi-supervised deep learning on pathological images," *Cold Spring Harbor Laboratory*, vol. 12, 2020.

[21] W. A. Kamakura, M. Wedel, F. de Rosa, and J. A. Mazzon, "Cross-selling through database marketing: a mixed data factor analyzer for data augmentation and prediction," *International Journal of Research in Marketing*, vol. 20, no. 1, pp. 45–65, 2003.

[22] S. Frühwirth-Schnatter, "Data augmentation and dynamic linear models," *Journal of Time Series Analysis*, vol. 15, no. 2, pp. 183–202, 1994.

[23] L. Ying, Z. Qian Nan, W. Fu Ping et al., L. Nam, Adaptive weights learning in CNN feature fusion for crime scene investigation image classification, *Connection Science*, vol. 33, no. 3, pp. 719–734, 2021.

[24] J. A. Royle and R. M. Dorazio, "Parameter-expanded data augmentation for Bayesian analysis of capture–recapture models," *Journal of Ornithology*, vol. 152, no. S2, pp. 521–537, 2012.

[25] H. Guo, Y. Mao, and R. Zhang, "Augmenting data with mixup for sentence classification: an empirical study," 2019, https://arxiv.org/abs/1905.08941.

[26] N. G. Polson and S. L. Scott, "Rejoinder: data augmentation for support vector machines," *Bayesian Analysis*, vol. 6, no. 1, pp. 43–47, 2011.

[27] J. Gordon and J. M. Hernández-Lobato, "Combining deep generative and discriminative models for Bayesian semi-supervised learning," *Pattern Recognition*, vol. 100, Article ID 107156, 2020.