

Research Article

Sentiment Classification for Financial Texts Based on Deep Learning

Shanshan Dong¹ and Chang Liu² 

¹Department of Economics in Engineering and Technology College, Hubei University of Technology, Wuhan, Hubei 432200, China

²Department of Qualitative Economics and Mathematics, School of Statistics and Mathematics, Zhongnan University of Economics and Law, Wuhan, Hubei 430073, China

Correspondence should be addressed to Chang Liu; z0004574@zuel.edu.cn

Received 31 August 2021; Revised 14 September 2021; Accepted 18 September 2021; Published 11 October 2021

Academic Editor: Bai Yuan Ding

Copyright © 2021 Shanshan Dong and Chang Liu. 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.

Sentiment classification for financial texts is of great importance for predicting stock markets and financial crises. At present, with the popularity of applications in the field of natural language processing (NLP) adopting deep learning, the application of automatic text classification and text-based sentiment classification has become more and more extensive. However, in the field of financial text-based sentiment classification, due to a lack of labeled samples, such applications are limited. A domain-adaptation-based financial text sentiment classification method is proposed in this paper, which can adopt source domain (SD) text data with sentiment labels and a large amount of unlabeled target domain (TD) financial text data as training samples for the proposed neural network. The proposed method is a cross-domain transfer-learning-based method. The domain classification subnetwork is added to the original neural network, and the domain classification loss function is also added to the original training loss function. Therefore, the network can simultaneously adapt to the target domain and then accomplish the classification task. The experiment of the proposed sentiment classification transfer learning method is carried out through an open-source dataset. The proposed method in this paper uses the reviews of Amazon Books, DVDs, electronics, and kitchen appliances as the source domain for cross-domain learning, and the classification accuracy rates can reach 65.0%, 61.2%, 61.6%, and 66.3%, respectively. Compared with nontransfer learning, the classification accuracy rate has improved by 11.0%, 7.6%, 11.4%, and 13.4%, respectively.

1. Introduction

With the rapid development of technologies such as big data, artificial intelligence, and deep learning, how to adopt these technologies to the financial field has become a research hotspot. The sentiment classification of financial texts based on machine learning is of great significance for predicting stock markets and financial crises. Compared with traditional methods based on financial text sentiment classification, it has the following advantages: (1) the influence of human emotional factors can be reduced and (2) a large amount of data samples can be adopted for analysis. As a matter of fact, sentiment prediction of financial text based on machine learning has become a research hotspot. Through

the analysis of current financial texts from the Internet, a preliminary judgment can be made about current economic fundamentals.

Currently, there has been in-depth research in the field of text classification. In terms of text representation, Collobert [1] firstly proposed to represent the text through the word vectors. References [2, 3] proposed a word2vec representation method. In this method, words not only can be represented by vectors, but the representation also embeds the word vectors to another space so that the vector distance in this space can express semantic and syntactic similarity. The representation method of word2vec has adopted the prior information obtained in the corpus, which is trained through the CBOW model [4] or skip-gram [5] model to

obtain the vector representation of words. Generally speaking, since the corpus covers a wide range of texts, the mentioned word representation method is also representative. Reference [6] has proposed the doc2vec model by extending words to paragraphs, where the granularity of the text representation is changed. In terms of text-based classification, it can be divided into traditional machine-learning-based and deep-learning-based methods. The authors in [7] proposed a method of hierarchical support vector machine (SVM) to classify texts. The authors in [8] proposed to improve the pruning strategy of the traditional decision tree model to classify the text, which has achieved good results. The authors in [9] proposed to use the Naive Bayes model to classify text. Since the deep-learning-based methods can extract more abstract and higher-level features from the samples, they generally have better classification performances than traditional methods. The authors in [10] firstly introduced deep learning into the field of natural language processing (NLP). The authors in [11] combine word vectors and convolutional neural networks for the sentiment classification of texts and has achieved good results.

The topic of this paper is sentiment classification for financial texts. Judging whether the financial text is positive through the classification network can be of great significance for judging the current economic situations or investment enthusiasm. It has great significance for predicting the stock market and financial crisis. However, in the field of financial text sentiment classification, the biggest difference from text sentiment classification in other domains is the lack of a large number of labeled samples. Therefore, the application of sentiment classification in the financial text has been greatly restricted. In order to adopt deep learning methods for performing sentiment analysis on financial texts, a straightforward idea is to use cross-domain learning methods to transfer sentiment classification knowledge from other domains to the domain of financial texts. Generally speaking, the methods of transfer learning in this field can be divided into three categories. (1) The first category is the parameter fine-tuning-based method. In [12], by pretraining the convolutional neural network using the source domain data and then adopting the target domain data to fine-tune the parameters of the neural network, the original neural network can be transferred to the target domain. One disadvantage of this method is that it requires a large amount of labeled target domain data. Since the amount of labeled data in financial text is not sufficient, this method is not suitable for the application of financial text classification. The authors in [13] propose to freeze shallow network parameters in the process of parameter fine-tuning and only change higher-layer network parameters. The principle for the method is that the shallow features of cross-domain samples are the same, but the abstract features in the high level are different. Therefore, only high-level features need to be transferred. This method can reduce the number of labeled samples needed in the target domains but is still not suitable for the application in our paper. (2) The second category is feature representation learning methods. The authors in [14, 15] propose that cross-domain representation learning can be

performed through stacked denoising autoencoder (SDA). After learning, the feature extraction neural network can be obtained. The network can extract shared features between different data domains. By adopting the extracted shared feature vectors, the problem of cross-domain text sentiment classification can be solved through the SVM classifier. In this method, no labeled target domain samples are required. The authors in [16, 17] propose a method for financial text sentiment classification based on a generative confrontation network (GAN), which combines the random noise generated in the generative network with text representation vectors. Then a discriminant network module is adopted to distinguish true source domain samples, generated samples, and the sentiment. This method can achieve good results, but the training of the network need large-volume datasets. The authors in [18, 19] propose to adopt the active learning method, which can effectively reduce the number of labeled samples in the target domain. The authors in [20, 21] propose a method of adversarial learning; by sentiment classification learning and domain discrimination learning, the purpose of domain adaptation can be achieved.

A financial text sentiment classification method based on unsupervised domain adaptation (DA) is proposed in this paper. The proposed method can transfer the sentiment classification knowledge of the source domain to the target domain of the financial text. Although the distribution of the source domain and the target domain are different, the shared features can still be learned for classification. The proposed method has the following advantages: (1) it does not require labeled samples in financial texts, so the method is unsupervised and suitable for the actual situation of lack of sentiment labels in financial texts and (2) the network structure adopted in the proposed method is similar to the original network structure, where the only difference is that the domain classification subnetwork needs to be added. This can make the proposed network structure less complex. The proposed method is a cross-domain neural network transfer learning method. The domain classification subnetwork is added to the traditional sentiment classification network, and the domain classification cost is added to the original training cost so that the network can simultaneously adapt to the target domain as well as the source domain. The experiment of emotion classification migration learning is carried out through the open-source dataset. The experiment of the proposed sentiment classification transfer learning method is carried out through the open-source dataset. The proposed method in this paper uses the reviews of Amazon Books, DVDs, electronics, and kitchen appliances as the source domain for cross-domain learning, and the classification accuracy rates have reached 65.0%, 61.2%, 61.6%, and 66.3%. Compared with nontransfer learning, the classification accuracy rate has improved by 11.0%, 7.6%, 11.4%, and 13.4%, respectively.

2. Methods

The problem to be studied in this paper can be expressed as follows: after training with labeled text data in the source domain and unlabeled data in the target domain, emotion

classification is performed on the target domain, that is, financial text. Among them, the vector of the text sample of the source domain is represented as \mathbf{x}_i , where i represents the vector representation of the i -th sample, and the corresponding label can be represented as s_i , where i represents the label of the i -th sample. For the target domain sample, it is represented as \mathbf{y}_i , and the corresponding domain label is d_i . In the application of this article, the number of classifications required is 2, which are positive evaluation labels and negative evaluation labels. In this method, the model is trained according to the source domain samples and part of the target domain sample set, and then the target domain sample test set is used to test the accuracy of the model. In this section, the model is introduced according to the following parts. First, the overall model of domain adaptation in this article is introduced. Second, the cost function of the model in this article is introduced. Finally, the training process of the proposed model is introduced.

2.1. The Structure of the Proposed Network. The proposed model in this paper is shown in Figure 1. It can be seen that the model in this paper has added a domain classification subnetwork to the traditional classification model. The idea of domain transfer learning herein is: through the addition of domain classification subnetwork, the extracted features cannot distinguish the data domains. As a result, the extracted features from different data domains are more similar, thus achieving the transfer learning purpose.

The network structure proposed in this paper is composed of three substructures: feature extraction subnetwork, sentiment prediction subnetwork, and domain classification subnetwork. The details of the three subnetworks are introduced as follows:

- (1) The detailed structure of the feature extraction subnetwork is shown in Figure 2. The output is the extracted feature vector, which can be denoted as \mathbf{v}_i . The feature extraction network in this paper can have two structures. As shown in the figure, the first one is a common convolutional network structure, including three layers of convolutional network layers. The feature extraction subnetwork is similar to the traditional convolutional network layers, which extracts from shallow to deep features in the samples for classification. The second one is the feature extraction structure from a residual network. Since the residual network has direct links between convolutional layers that are not adjacent, it is usually easier to train. In our implementation, the residual network is adopted as the structure for the feature extraction subnetwork. In traditional convolutional neural networks, since there is only the cost function for classification, the extracted features can be used to distinguish sample labels. However, herein, the sentiment classification cost and the domain classification cost are both adopted. Through training, the feature extraction subnetwork can extract features that meet the following two conditions: firstly, the features can be effectively adopted for sentiment

differentiation, and secondly, it is impossible to distinguish which domain the input sample comes from according to the extracted features, for example, the data domain of financial text or ordinary product review text.

- (2) The structure of the sentiment classification subnetwork is shown in Figure 3, which consists of two fully connected layers plus a softmax layer for the final sentiment prediction. The output features of the previous feature extraction subnetwork can be adopted as the input, and the output denotes the probability of whether the sample belongs to the positive evaluation or not. The feature extraction subnetwork and sentiment classification subnetwork mentioned above can form a traditional sentiment classification neural network. Therefore, in this paper, the existing neural network structure of sentiment classification can be directly used. Herein, a cross-domain sentiment classification network can be obtained with improvement on the existing sentiment classification network. The following equation denotes the softmax layer operation, where N represents the number of categories and e^i represents the feature input of the dimension i .

$$O_i = \frac{e^i}{\sum_{j=1}^N e^j}. \quad (1)$$

- (3) The structure of the domain classification network is shown in Figure 4. For the domain classification subnetwork, its main purpose is to be able to distinguish whether the samples come from the source domain or the target domain (in this paper, the samples come from financial texts or other texts). As mentioned earlier, for the feature extraction subnetwork, we hope to extract features that cannot distinguish the sample data domain so as to achieve the purpose of transfer learning. However, this purpose is contradictory to the purpose of domain classification. In order to solve this problem, as shown in Figure 4, a reverse gradient layer is added to the traditional classification network. The reverse gradient layer can be expressed as follows:

$$\begin{aligned} \mathbf{RG}(\mathbf{x}) &= \mathbf{x}, \\ \frac{\partial \mathbf{RG}(\mathbf{x})}{\partial \mathbf{x}} &= -\tau \mathbf{I}. \end{aligned} \quad (2)$$

Among them, $\mathbf{RG}(\cdot)$ represents the function of the forward propagation of the reverse gradient layer. During the forward propagation, this layer does not have any influence on the propagation of the data. When the gradient is backpropagated, the layer adds a minus sign to the original gradient. Designing the reverse gradient layer in this way can enable the domain classification subnetwork trainable and enable the feature extraction subnetwork to obtain features that cannot distinguish domains.

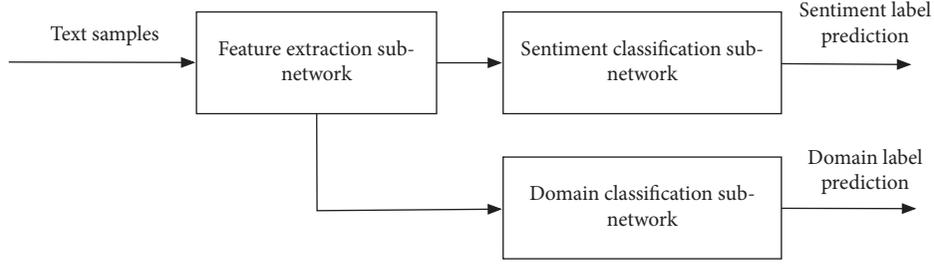


FIGURE 1: The network model of the proposed method.

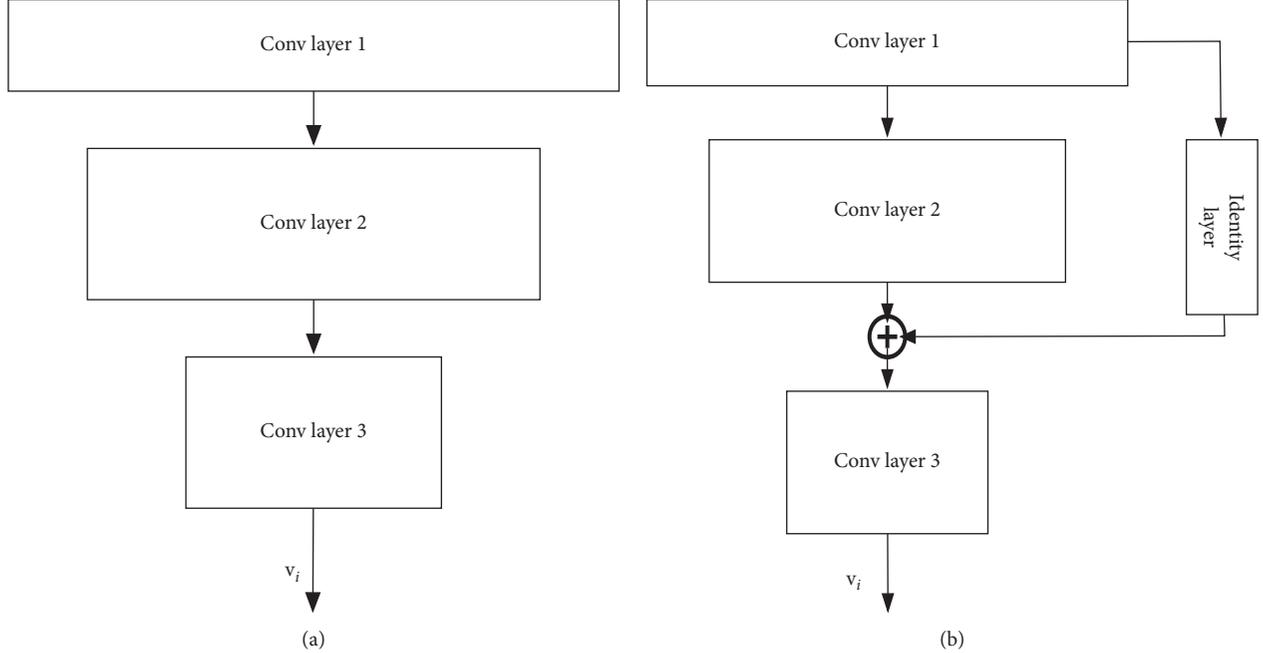


FIGURE 2: The detailed structure of the feature extraction subnetwork: (a) convolutional network and (b) residual network.

2.2. *The Loss Function for the Proposed Network.* According to the above descriptions, Table 1 shows the symbols of functions and parameters in different subnetworks.

According to the above symbols, it can be seen that there are two cost functions in the network structure:

- (1) Cost function for sentiment classification: the cost can be expressed by C_s , which can be written as follows:

$$C_s(\theta_f, \theta_s) = \sum_{i=1}^M C_{s,i}(F_s(F_f(\mathbf{x}_i; \theta_f); \theta_s), s_i), \quad (3)$$

here i denotes the index of the sample, M represents the number of samples, s_i represents the sentiment label of the corresponding sample, \mathbf{x}_i represents the sample of the input source domain, and $C_{s,i}(\cdot)$ represents the corresponding cross-entropy function, which can be written as follows:

$$C_{s,i}(\cdot) = - \sum s_i \log O_i, \quad (4)$$

where O_i represents the output of the softmax layer and s_i represents the number of one-hot encoding based on the actual sentiment classification label.

- (2) Domain classification cost function: the cost can be written as follows:

$$C_d(\theta_f, \theta_d) = \sum_{i=1}^M C_{d,i}(F_d(F_f(\mathbf{x}_i; \theta_f); \theta_d), d_i). \quad (5)$$

Among them, C_d represents the cost function, i represents the index of the sample, M represents the number of samples, d_i represents the domain label of the corresponding sample, \mathbf{x}_i represents the sample of the input source or target domain, and $C_{d,i}(\cdot)$ represents the domain classification cross-entropy:

$$C_{d,i}(\cdot) = - \sum d_i \log O_i, \quad (6)$$

where O_i represents the output of the softmax layer of the domain classification subnetwork and d_i represents the number of domain sources after one-hot encoding.

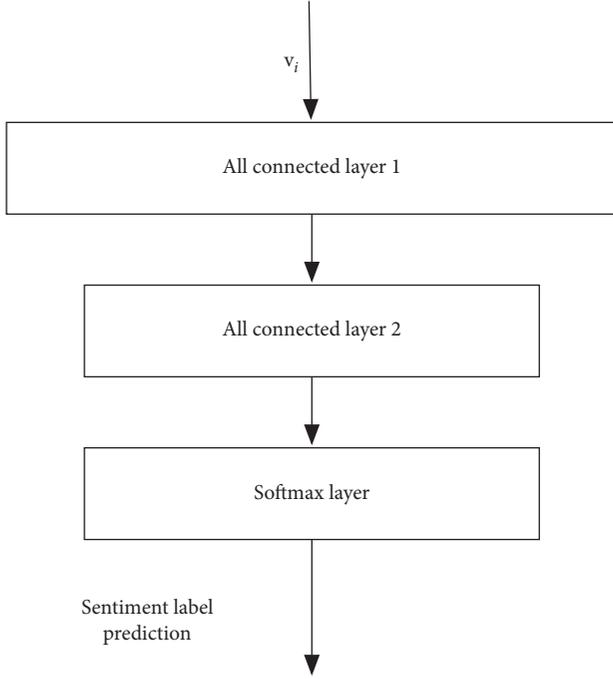


FIGURE 3: The sentiment classification subnetwork structure.

According to the sentiment classification cost function and domain classification cost function, the overall cost function can be written as follows:

$$C_{all}(\theta_f, \theta_s, \theta_d) = C_s(\cdot) - \tau C_d(\cdot), \quad (7)$$

where C_{all} represents the overall cost function, and the parameter τ denotes the ratio of the two cost functions. Noting that here, a minus sign is added before the domain classification cost function, which is equivalent to the operation of the gradient backpropagation layer mentioned above. After the addition of the minus sign, the training process is equivalent to maximizing the cost of the domain classification subnetwork while minimizing the cost of the sentiment classification subnetwork. In this way, the features extracted by the feature extraction subnetwork can be both effective for distinguishing sentiment labels and at the same time insensitive for distinguishing domains.

2.3. The Training Processes. According to the overall cost function mentioned above, during training, different subnetwork parameters can be updated as follows:

$$\begin{aligned} \theta_f &\longrightarrow \theta_f - \mu \left(\frac{\partial C_s}{\partial \theta_f} - \tau \frac{\partial C_d}{\partial \theta_f} \right), \\ \theta_s &\longrightarrow \theta_s - \mu \frac{\partial C_s}{\partial \theta_s}, \\ \theta_d &\longrightarrow \theta_d - \mu \frac{\partial C_d}{\partial \theta_d}. \end{aligned} \quad (8)$$

The parameter μ indicates the update rate. The value of the learning rate μ in our implementation is set to 0.0005. The value is a typical value for the learning rate.

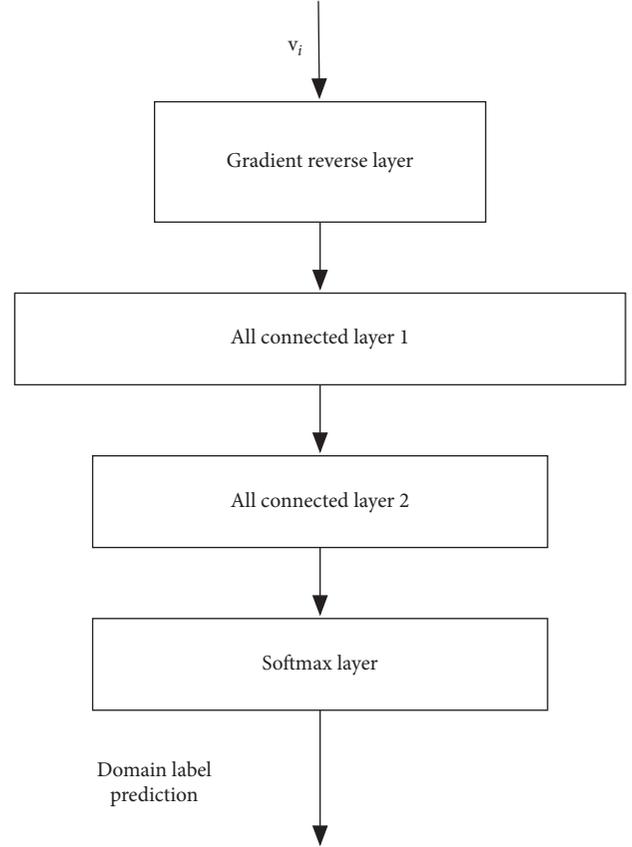


FIGURE 4: The domain classification subnetwork structure.

In actual training process, training is carried out in batches. In order to optimize all cost function items in each batch, the batch adopted here needs to include the samples of the source domain and the target domain. As a result, during optimization, there is a certain confrontational relationship between minimizing the cost of sentiment classification and maximizing the cost of domain classification. Here, the parameter τ_p determines the proportion of the two cost functions. During the training process, the hyperparameter τ_p is set according to the following equation:

$$\tau_p = \frac{2}{1 + \exp(-\alpha p)} - 1. \quad (9)$$

To set the hyperparameter according to the above formula, the advantages are as follows. (1) At the early stage of training, the recognition rate is low and the value of p is close to 0. Setting p close to 0 can make the training network become a pure sentiment classification network without considering the problem of transfer learning. (2) When p increases, τ_p also increases. It means that with the convergence of the feature extraction and sentiment classification subnetwork during training, the cost of transfer learning should be taken into consideration. The parameter of α determines the conversion speed for the two situations. In one situation, only the recognition rate is considered. In the other, only the domain classification accuracy is considered. In the above formula, the value of α is set to 25 according to subsequent experiments.

3. Results and Discussion

3.1. The Adopting Dataset. In order to verify the proposed cross-domain sentiment classification method, the classic Amazon dataset is adopted. The dataset contains more than 340,000 reviews, which cover 22 different products. Because the dataset contains too many types, the size of the set is large, and the positive and negative reviews are uneven; herein, the corresponding reduced Amazon dataset is adopted for experimentation. The basics of the reduced dataset are shown in Table 2. It can be seen that the dataset contains four different domains of texts: books, DVDs, electronics, and kitchen appliances. Each category contains 2000 tagged reviews, of which 1000 are positive reviews and 1000 are negative reviews, each accounting for 50%. In addition, they each have a few thousand unlabeled reviews. The reduced dataset can be adopted to evaluate the domain adaptation method proposed in this paper. In order to verify the effect of sentiment classification in the financial text field, the Microblog dataset in Stocktwits is also adopted, which contains 2000 unlabeled samples and 500 samples with labels. In this experiment, this dataset can be used as the target domain for research.

3.2. The Performance of the Proposed Method. According to the dataset mentioned above, the proposed method can be verified. The reviews of the four products in the reduced Amazon dataset are used as the source domains, and the Microblog dataset is used as the target domain for training and testing. The test results are shown in Table 3, and the different source and target domains corresponding to different situations in the table are shown in Table 4. It can be seen that the proposed method can effectively perform cross-domain sentiment classification, and its correction rate can reach 65.0%, 61.2%, 61.6%, and 66.3% under the situation of A, B, C, and D, respectively; compared with the results without transfer learning, the recognition rates have been improved by 11.0%, 7.6%, 11.4%, and 13.4%, respectively. For the average recognition rate, it is improved by 10.9%. The method without transfer learning is the method without the domain classification subnetwork. For this subsection, the main purpose is to validate the transfer learning method. Therefore, the method to compare has the same structure as the proposed method, where the only difference is that there is no domain classification subnetwork.

4. Method Comparisons

The proposed method is compared with the previously mentioned SVM-based method, the parameter fine-tuning method, and the SDA-based transfer method. Herein, the SVM-based method is used as the baseline method. Note that in the baseline method, transfer learning is not performed. The purpose of adding the baseline method is to

evaluate the effect of the transfer learning method according to the following metrics. In order to better compare the transfer effect, inspired by [14], the transfer ratio index R_t and in-domain ratio index R_i are defined. Among them, the transfer ratio index R_t can be expressed as follows:

$$R_t = \frac{1}{N} \sum_{(s,t)} \frac{e(s,t)}{e_b(t,t)}, \quad (10)$$

where s and t represent the source domain and the target domain respectively; N represents the number of pairs from the source domain to the target domain; $e(s,t)$ represents the transfer error, which denotes the test error of transfer learning; and $e_b(t,t)$ represents the in-domain error according to the baseline method. The transfer ratio index reflects the transfer learning effect of the corresponding method, and its value is positively correlated with the performance of the transfer learning method. The in-domain ratio index can be expressed as follows:

$$R_i = \frac{1}{M} \sum_{(s)} \frac{e(t,t)}{e_b(t,t)}. \quad (11)$$

Among them, M represents the number of all domains; $e(t,t)$ represents the in-domain error of the corresponding method, that is, the test error under the condition that both training and testing are in the source domain, and $e_b(t,t)$ represents the in-domain error of the baseline method. The in-domain error can be used to represent the sentiment classification effect of the method under the conditions of the same data domain for training and testing. The smaller the value is, the better the classification performance is. Figure 5 shows the comparisons of the transfer ratio and in-domain ratio of the proposed method, the baseline method, the parameter fine-tuning method, and the SDA method. It can be seen that the values of transfer and in-domain errors of the proposed method in this paper are both the smallest, which shows that the proposed method has the best classification performance not only under the same domain but also under the condition of cross-domains. For the adopted dataset, the data volume is not large enough for the GAN-based network. Therefore, it is not compared in our implementation. This also shows the disadvantage of the GAN-based method, which is hard to be trained. The performance of the GAN-based method will be studied in the future.

Figure 6 shows the bar plot of the cross-domain recognition rates, including the proposed method, the baseline method, the parameter fine-tuning method, and the SDA method. Table 5 shows the comparisons of the cross-domain average recognition rate of the four methods. It can be seen that the cross-domain recognition rate of the proposed method has reached 63.5%. Compared with the baseline method, the parameter fine-tuning method, and the SDA method, the average recognition rate has improved by 11.1%, 5.1%, and 1.8%, respectively.

TABLE 1: Symbols for the subnetworks.

Subnetwork	Network symbol	Parameter symbol
Feature extraction	$F_f(.)$	θ_f
Sentiment classification	$F_s(.)$	θ_s
Domain classification	$F_d(.)$	θ_d

TABLE 2: The basics of the reduced Amazon dataset.

Products	Labeled number	Unlabeled number
Books	2000	4465
DVDs	2000	3586
Electronics	2000	5681
Kitchen	2000	5945

TABLE 3: The test results of the proposed method.

Situation	Proposed (%)	Without transfer learning (%)
A	65.0	54.0
B	61.2	53.6
C	61.6	49.8
D	66.3	52.9

TABLE 4: Correspondences between source and target domains in different situations.

Situation	Source domain	Target domain
A	Books	Microblog
B	DVDs	Microblog
C	Electronics	Microblog
D	Kitchen	Microblog

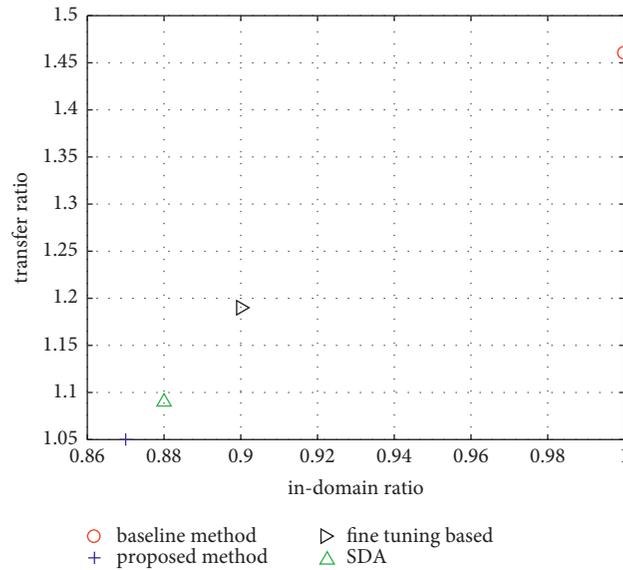


FIGURE 5: Comparisons of transfer ratio and in-domain ratio of different methods.

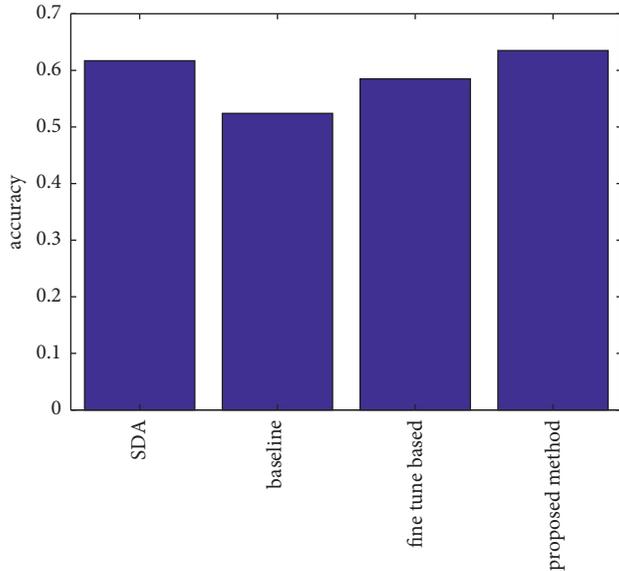


FIGURE 6: Bar plot of average recognition rate of different methods.

TABLE 5: Comparisons of average recognition rate of different methods.

Method	Accuracy (%)	Improvement
Baseline	52.4	11.1%
Proposed method	62.5	—
Fine-tune-based	58.5	5.1%
SDA	61.7	1.8%

5. Conclusions

For financial-text-based sentiment classification, it is often difficult to obtain a high recognition rate due to the lack of labeled samples for financial texts. In order to solve the problem, a transfer learning method based on domain adaptation is proposed in this paper. In the proposed method, a large amount of source domain text samples is adopted for sentiment classification. The proposed method is a cross-domain transfer-learning-based method. The domain classification subnetwork is added to the original sentiment classification network, and the domain classification loss function is also added to the original training loss function. Therefore, the network can simultaneously adapt to the target domain and accomplish the classification task. The experiment of the proposed sentiment classification transfer learning method is carried out through the open-source dataset. The proposed method in this paper uses Amazon books, DVDs, electronics, and kitchen appliances as the source domain for cross-domain learning, and the classification accuracy rates have reached 65.0%, 61.2%, 61.6%, and 66.3%, respectively. Compared with nontransfer learning, the classification accuracy rate has improved by 11.0%, 7.6%, 11.4%, and 13.4%, respectively. Compared with the SVM-based baseline method, the CNN-based parameter fine-tuning transfer learning method, and the SDA-based method, the average recognition accuracy rate has been improved by 11.1%, 5.1%, and 1.8%, respectively.

For the adopted dataset, the data volume is not large enough for the GAN-based network. Therefore, it is not compared in our implementation. This also shows the disadvantage of the GAN-based method, which is hard to be trained. The performance of the GAN-based method will be studied in the future.

Data Availability

The Amazon dataset is publicly available.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This paper was approved as a research project of humanities and social sciences by Hubei Provincial Department of Education (19G009).

References

- [1] R. Collobert and J. Weston, "A unified architecture for natural language processing: deep neural networks with multitask learning," in *Proceedings of the 25th International Conference on Machine Learning*, Helsinki, Finland, July 2008.
- [2] T. Mikolov, "Distributed representations of words and phrases and their compositionality," 2013, <https://arxiv.org/abs/1310.4546>.
- [3] T. Mikolov, "Efficient estimation of word representations in vector space," 2013, <https://arxiv.org/abs/1301.3781>.
- [4] T. Kenter, A. Borisov, and M. De Rijke, "Siamese CBOW: optimizing word embeddings for sentence representations," 2016, <https://arxiv.org/abs/1606.04640>.
- [5] D. Guthrie, "A closer look at skip-gram modelling," *LREC*, vol. 6, 2006.
- [6] Q. Le and T. Mikolov, "Distributed representations of sentences and documents," in *Proceedings of the International Conference on Machine Learning*, Beijing, China, June 2014.
- [7] R. Moraes, J. F. Valiati, and W. P. Gavião Neto, "Document-level sentiment classification: an empirical comparison between SVM and ANN," *Expert Systems with Applications*, vol. 40, no. 2, pp. 621–633, 2013.
- [8] T. Pranckevičius and V. Marcinkevičius, "Comparison of Naive Bayes, random forest, decision tree, support vector machines, and logistic regression classifiers for text reviews classification," *Baltic Journal of Modern Computing*, vol. 5, p. 221, 2017.
- [9] J. Chen, H. Huang, S. Tian, and Y. Qu, "Feature selection for text classification with Naïve Bayes," *Expert Systems with Applications*, vol. 36, no. 3, pp. 5432–5435, 2009.
- [10] R. Collobert, "Natural language processing (almost) from scratch," *Journal of Machine Learning Research*, vol. 12, pp. 2493–2537, 2011.
- [11] R. Socher, "Parsing natural scenes and natural language with recursive neural networks," in *Proceedings of the 28th International Conference on Machine Learning*, Bellevue, WA, USA, June 2011.
- [12] D. M. Diab and K. M. El Hindi, "Using differential evolution for fine tuning naïve Bayesian classifiers and its application for text classification," *Applied Soft Computing*, vol. 54, pp. 183–199, 2017.

- [13] K. M. El Hindi, R. R. Aljulaidan, H. AlSalman, and H. AlSalman, "Lazy fine-tuning algorithms for naïve Bayesian text classification," *Applied Soft Computing*, vol. 96, p. 106652, 2020.
- [14] X. Glorot, B. Antoine, and Y. Bengio, "Domain adaptation for large-scale sentiment classification: a deep learning approach," in *Proceedings of the 28th International Conference on International Conference on Machine Learning*, Bellevue, Was, USA, June 2011.
- [15] H. Sagha, N. Cummins, and B. Schuller, "Stacked denoising autoencoders for sentiment analysis: a review," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 7, no. 5, p. e1212, 2017.
- [16] Z. Zhang, Z. Wang, C. Gan, and P. Zhang, "A double auction scheme of resource allocation with social ties and sentiment classification for Device-to-Device communications," *Computer Networks*, vol. 155, pp. 62–71, 2019.
- [17] Y. Zhang, D. Miao, and J. Wang, "Hierarchical attention generative adversarial networks for cross-domain sentiment classification," 2019, <https://arxiv.org/abs/1903.11334>.
- [18] S. Li, "Active learning for cross-domain sentiment classification," in *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, Beijing, China, August 2013.
- [19] S. Otálora, O. Perdomo, F. González, and H. Müller, "Training deep convolutional neural networks with active learning for exudate classification in eye fundus images," *Lecture Notes in Computer Science*, vol. 10552, pp. 146–154, 2017.
- [20] Z. Li, "End-to-End adversarial memory network for cross-domain sentiment classification," in *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, Melbourne, Australia, August 2017.
- [21] Y.-T. Tsai, M.-C. Yang, and H.-Y. Chen, "Adversarial attack on sentiment classification," in *Proceedings of the 2019 ACL Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*, Florence, Italy, 2019.