International Journal of Information Technologies and Systems Approach
Volume 16 • Issue 2

An Effective Emotional Analysis Method of Consumer Comment Text Based on ALBERT-ATBiFRU-CNN

Mei Yang, Nanchang Institute of Technology, China*
https://orcid.org/0000-0003-0253-7787

ABSTRACT

To address the challenges of insufficient feature extraction for text sentiment analysis in the e-commerce big data environment, the author proposes a deep learning-based emotion analysis method of consumer comment text. Firstly, the author obtained the contextualized word vectors by using a pretrained language model called A Lite Bidirectional Encoder Representations From Transformers (ALBERT). Secondly, the researcher used the bidirectional gate recurrent unit (BiGRU) model to capture the semantic information through the combination of positive and negative directions, measure the emotional polarity information of each text as a whole, and then catch the local characteristic information of the text using the convolutional neural network (CNN) model. Finally, the author calculated the weight distribution through the attention mechanism. The experiments on a publicly available consumer review dataset showed that the recall, precision, and F1-score of the proposed text emotion analysis method were 0.9417, 0.9552, and 0.9484, respectively, which are higher than the existing methods. Therefore, the proposed method is of great significance in capturing the emotions of consumers on e-commerce platforms.

KEYWORDS

E-Commerce Big Data, Bidirectional Gate Recurrent Unit, Text Emotion Analysis, ALBERT Pretraining, Attention Mechanism

INTRODUCTION

With the rapid development of Internet + in recent years, Internet information has shown explosive growth. For example, users can easily express their feelings about life and share their joys and sorrows in real time through social platforms such as Weibo, blogs, and Twitter (Zhou et al., 2019). In the era of Web 2.0, numerous Internet products have enabled people to have a greater voice and influence in the network. For example, after some daily consumption, people can comment and score merchants and dishes on life service platforms such as Meituan, Dianping, and Alipay Koubei. These comments

DOI: 10.4018/IJITSA.324100
*Corresponding Author

and scores directly affect which restaurants other people eat in and what dishes they order (Day & Lin, 2017; Kim et al., 2019; Ruiz-Mafe et al., 2020; Simpson et al., 2011; Xu, 2020; Xu et al, 2019). Among such comments, emotion is the best way to reflect users' attitudes, thoughts, and judgments. Therefore, emotion analysis task emerged, that is using an intelligent computing method to identify the emotional tendencies expressed by users in a paragraph of text (Rúa-Hidalgo et al, 2021; Yu, 2014).

Emotion analysis also known as viewpoint mining (Meng et al., 2021; Murgia et al., 2018; Tosun & Sezgin, 2021; Wu et al, 2019). Emotion analysis is also widely used in industry for predicting market changes based on news comments and emotions in blogs. The government can understand people's needs and feelings through the views they express on the Internet, and thus focus on people's livelihood. Emotion analysis is an indispensable part of the existing recommendation system. The system can understand the user's preferences by analyzing their emotional changes in comments and can recommend more appropriate content to the user (Christodoulides et al., 2021; Park & Han, 2018). Businesses can understand the advantages and disadvantages of their products and consumers' opinions on the products according to this information, improve the shortcomings of the products, and promote the products to more suitable target groups. Consumers also want to buy goods that are suitable for them, and they use other users' reviews to make purchasing decisions. Therefore, comment information is also the focus of consumers. However, e-commerce platforms include a large number of reviews on widely distributed product. It is difficult for consumers and sellers to browse all the review information, and it is even more difficult to find valuable information from the huge amount of reviews. Inappropriate and incomplete comments can lead to regretful shopping decisions by consumers; this directly affects the attitudes of future consumers towards products and market evaluation of products (Han et al., 2018; Kujur & Singh, 2018; Wang et al, 2021).

A growing number of deep learning models have been applied to natural language processing tasks, especially the emergence of word vector technology, which enables text to be represented as low-dimensional and continuous features. The advantages of automatic learning and feature extraction by deep learning can not only overcome the disadvantages of complex feature engineering brought by traditional machine learning to a certain extent, but also achieve better results (Dhaoui et al., 2017; Haavisto & Sandberg, 2015).

The rest of the paper is organized as follows: The second section provides the literature review; the third section describes the text sentiment analysis based on the deep learning model; the fourth section presents the experiment and its analysis; lastly, the fifth section provides the conclusion.

RELATED WORK

In face of complex and changing language phenomena, the deep learning method captures more comprehensive and in-depth text features. Collobert et al. (2011) introduced a convolutional neural network (CNN) into the sequence annotation task and obtained good results. Kim (2014) proposed text CNN, which uses features similar to n-gram, adds position information of words to feature extraction, and uses multiple window CNN networks to encode sentences. Kalchbrenner et al. (2014) proposed a dynamic CNN by combining lexicalization technology K-MAX. Wu and Ong (2021) proposed a text emotion analysis model based on QACGBert. The model uses a multihead attention memory network similar to the BERT model and introduces an attention mechanism to consider the relationship between different words. The disadvantage of this method is that the temporal relationship is not used. Zhao et al. (2018) applied the capsule network to the emotion analysis task to reduce the loss of characteristic information of CNN in the pooling layer and achieved better classification performance than that of the traditional CNN. Rao et al. (2018) proposed the SR-LSTM model and cleaned up the parts with the weak emotional tendency in the dataset, which achieved good results in long text classification. Rhanoui et al. (2019) combined Doc2Vec embedding, CNN, and BiLSTM to build a long text analysis model that could effectively process text about specific products. Jamal et al. (2021) utilized the hybrid model of term frequency reverse document frequency and deep learning to conduct emotional analysis. Yu et al. (2018) suggested a framework for extracting aspect and view terms using BiLSTM and multilayer attention networks. Tang et al. (2016) stored information by adding an external memory network and utilized the attention mechanism to capture each context word related to the target word.

However, various challenges often hinder the application of the above-mentioned methods in the text emotion analysis task; they include low text utilization rate, ineffective information extraction, and insufficient feature extraction. In this paper, the author suggests a deep learning-based emotion analysis method for consumer comments in the e-commerce big data environment to overcome the above-described problems. The innovation of the proposed method lies in the following features:

1. The “a lite bidirectional encoder representations from transformers” (ALBERT) pretraining language model obtains the contextualized word vector, the semantic information is simultaneously captured through the combination of the positive and negative directions through the bidirectional gate recurrent unit (BiGRU) model, and then the CNN model captures the local feature information from the text. The utilization rate of the text data is effectively improved through the effective combination of the two models.
2. The attention mechanism improves the accuracy of classification by calculating the weight distribution of the input to the generated emotion vector.

TEXT EMOTION ANALYSIS BASED ON THE ALBERT-ATBIGRU-CNN MODEL

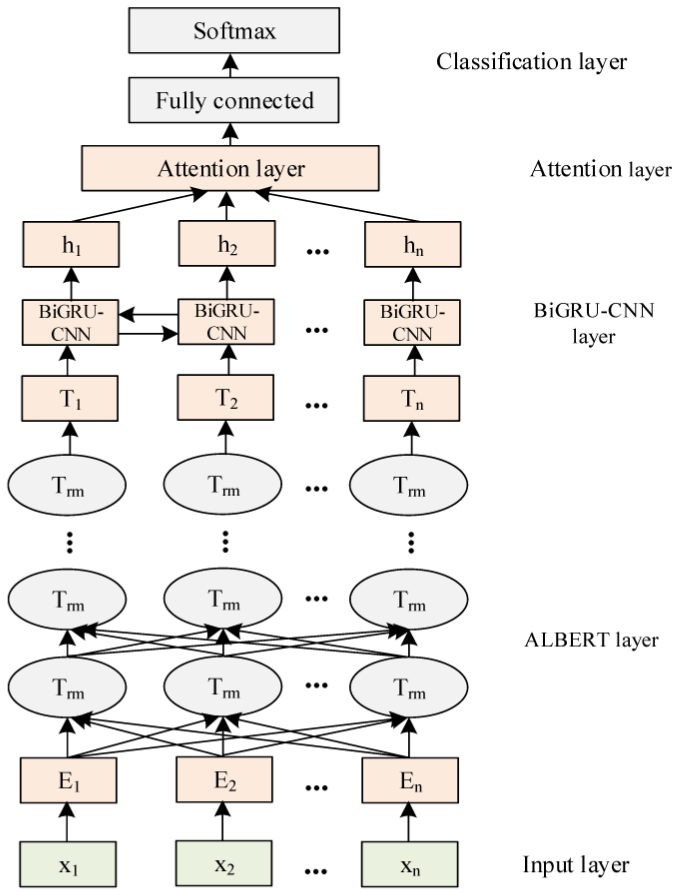
To solve the problems in the fine-grained emotion analysis task, the author proposes an ATBiGRU-CNN fine-grained emotion analysis model, which ueses the ALBERT model to obtain word embedding. The proposed model includes four main components: ALBERT word embedding layer, BiGRU-CNN (bidirectional gate recurrent unit-convolutional neural network) layer for generating sentence representation, attention layer for focusing on context representation, and emotion classification layer. Figure 1 shows the system architecture of the proposed model. The contextualized word vector can be obtained by the proposed method from the ALBERT pretrained language model. The BiGRU model is used to capture the semantic information through the combination of the negative and positive directions and the emotional polarity information of each text is measured as a whole. Then, the local characteristic information of the text is captured using the CNN model. Through the effective combination of the two models, the proposed method can obtain more feature information that contributes to text emotion analysis. Finally, weight distribution is calculated through the attention mechanism and important words are assigned more weight than others.

ALBERT Pretrained Model

The improvement of the ALBERT mainly includes the following four points, and Figure 2 shows its structure:

1. **Decomposition of the Embedded Layer Parameter Factor:** The ALBERT considers that the output value of the hidden layer not only contains context information, but also the original meaning of the word.
2. **Cross-Layer Parameter Sharing:** Due to the high similarity of features learned in each layer, parameter sharing operations are carried out within sub modules, thereby reducing the total number of parameters in the attention feed forward module.
3. **Loss of Coherence Between Sentences:** The next sentence prediction task in the BERT model includes topic prediction, two subtasks, and coherence prediction.
4. **Removal of Dropout:** Preventing model overfitting using dropout.

Figure 1. System architecture



The BiGRU-CNN

The feature extraction layer is composed of the BiGRU model and CNN model in series. First, the BiGRU model is used to capture semantic information through the combination of negative and positive directions to measure the emotional duality information of each text as a whole. Then, the local characteristic information of the text is captured using the CNN model. The effective combination of the two models provides rich feature information that improves the accuracy of the text emotion analysis. The output layer delivers the extracted features to the sigmoid function to obtain the emotion tags (Figure 3).

Feature extraction is the core step of text emotion analysis. The feature extraction layer employs BiGRU and CNN models to obtain the global and local semantic information of the text, respectively.

Global Semantic Information Extraction

The reset gate of the gate control module multiplies the information from both the previous time step and the current time step by a weight linear transformation and transfers the weight to the update gate. The weight of the updated gate is multiplied by the sigmoid function to obtain a value between [0, 1], and then spliced with the current input and transmitted to the tanh function to reduce the data to [- 1, 1]. At the same time, the hidden state calculated by the reset gate is obtained, and the hidden state of the reset gate is selectively added to the current hidden state to remember the state

Figure 2. The ALBERT structure

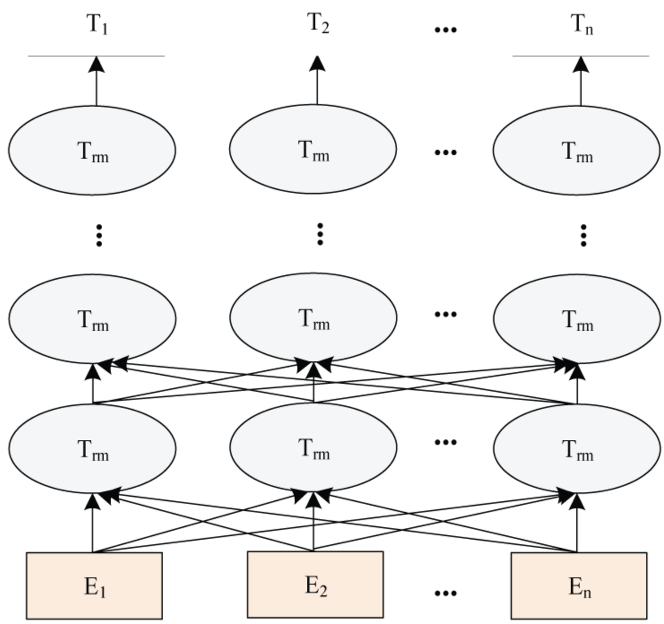
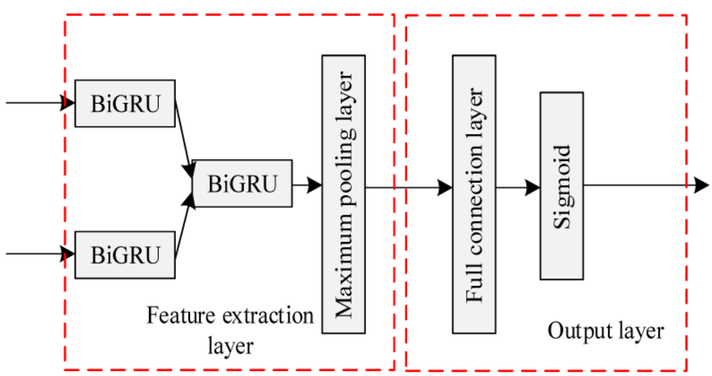


Figure 3. The structure of the BiGRU-CNN model



at the current time. Next is the "update memory" stage, which performs two steps of forgetting and selecting memory at the same time. The BiGRU model composed of two opposite GRUs is utilized to obtain the semantic information of the text. This enables the model to extract more comprehensive text features that are subsequently passed to the convolution layer.

Extraction of Local Semantic Information

In the text emotion analysis task, the convolution layer automatically extracts the features of the word vector matrix by setting different convolution kernel sizes and aggregates the feature information and sends it to the pooling layer. The pooling layer can effectively reduce the matrix size, reducing the dimension of feature information. Thus, the parameters of the fully-connected layer are reduced. The maximum pooling method retains the maximum value of all feature maps calculated by convolution. The pooling layer can not only accelerate the calculation, but also prevent overfitting. CNN is highly sensitive to local small text features. It extracts the n-gram local emotional features of local text through convolution operation, which is more effective for deep-seated emotional information mining of text sentences, and the weight sharing significantly reduces the number of parameters trained by the neural network.

To optimize the weight space parameters, the output layer uses the sigmoid function and produces the final text emotion analysis results.

Attention Mechanism

Figure 4 shows the model structure of the attention mechanism.

The calculation process of the attention mechanism is as follows:

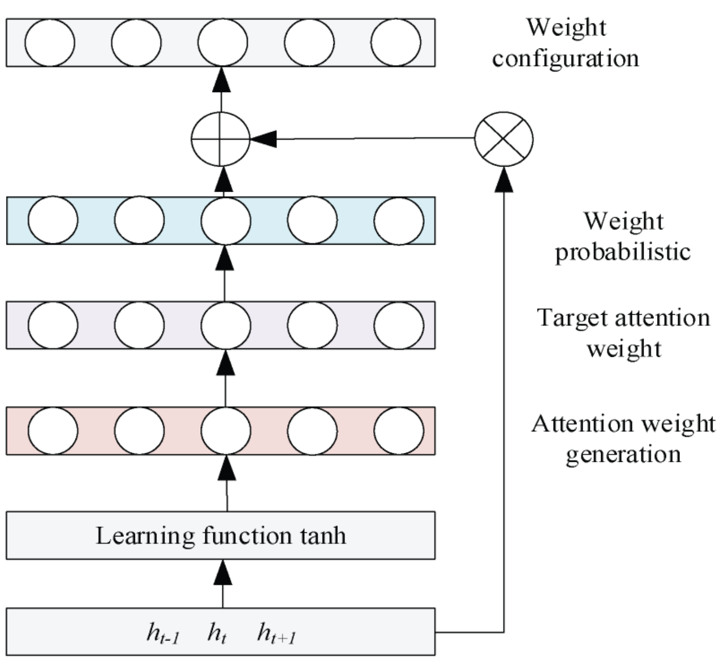
1. The features extracted by Text-CNN and BiGRU are represented as h_T *and* y_B, respectively, and the input features are recorded as h_A. Equation 1 shows their relationship:

$$h_A = h_T + y_B \tag{1}$$

2. Calculate the target attention weight u_t^i, where the weight matrix is W_u and the bias is b_u, as Equation 2 shows:

$$u_t^i = v_i \tanh(W_u h_A + b_u) \tag{2}$$

Figure 4. Model structure of the attention mechanism



3. The weight coefficient a_t^i is obtained by normalizing the attention weight with the SoftMax function to highlight the weight of important words, as Equation 3 shows:

$$a_t^i = soft\max(u_t^i) \tag{3}$$

4. The weighted sum of the fused feature h_A and the weight coefficient a_t^i provide the text vector c^i of each word in the text, as Equation 4 shows:

$$c^i = \sum_{i=1}^{n} a_t^i h_A \tag{4}$$

EXPERIMENT AND ANALYSIS

Experimental Setup

The author applied the proposed model and conducted all the experiments below under the PyTorch deep learning framework. Table 1 shows the experimental environment configuration.

The experimental dataset represented the user comments on digital products (https://github.com/SophonPlus/ChineseNlpCorpus/blob/master/datasets/online_shopping_10_cats/intro.ipynb). This dataset contains over 60,000 comments, with approximately 30,000 positive and 30,000 negative comments each. 12,000 comments were selected from the original data, including 6,000 positive, and 6,000 negative comments. The entire dataset was divided into training, testing, and validation sets using a ratio of 6:2:2.

Evaluating Indicator

The author evaluated the proposed method using four dimensions: Accuracy, recall, precision, and F1-score. The four indicators are defined as follows:

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

Table 1. Experimental environment configuration

Name	Configuration
Operating system	WIN10
CPU	CPU Intel→ Core TM i7-7700 CPU@3.60 GHz
Hard disk	2 TB
Development framework	PyTorch 1.2.0
Development language	Python 3.7
Graphics card	GTX 1080 Ti
Memory	16 GB

$$F = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{7}$$

$$Accuracy = \frac{TP + TN}{TP + FN + FP + FN} \tag{8}$$

where TP, FP, TN, and FN represent the number of elements of the real case, false positive case, positive negative case, and false negative case, respectively.

Model Training

The author used the deep learning-based emotion analysis method for analyzing emotions in the consumer comment text and trained the model on the emotion classification task of online goods comment text. Figure 5 shows the process diagram accuracy of the training process of 70 epochs. When the epoch is 40, the accuracy of the proposed model is basically close to the highest value 95.38%, which means that the proposed model has reached a convergence state.

Comparison and Analysis

Figure 6 shows the performance of the loss function of the ALBERT-ATBiGRU-CNN method under different learning rates. The experimental results show that the loss function does not converge to the global minimum when a higher learning rate is selected for training. The network weight adjusted by training a batch is small when the learning rate is too small. In this way, the possibility of skipping the global minimum is avoided to a certain extent and the reliability of training is enhanced, but the time cost of training the model is significantly increased. Thus, the learning rate in this experiment is 1×10-4. The number of layers in the neural network is large, and the network structure of the ALBERT model is complex. Small changes in the learning rate will cause network oscillation with large amplitude, and directly impacts the rate at which the model accumulates information. Deviation from the optimal value will prevent the model from achieving the best accuracy.

To demonstrate the effectiveness of the proposed text emotion analysis method, the author compared it with BiGRU-CNN and ALBERT text emotion analysis methods under the same experimental conditions. The researcher measured the text emotion analysis effect of each model using recall, precision, and F1-score. Table 2 and Figure 7 show the results. Recall, precision, and F1-score of the proposed text emotion analysis method in the text dataset were 0.9417, 0.9552, and 0.9484, respectively, which

Figure 5. The accuracy results in model training

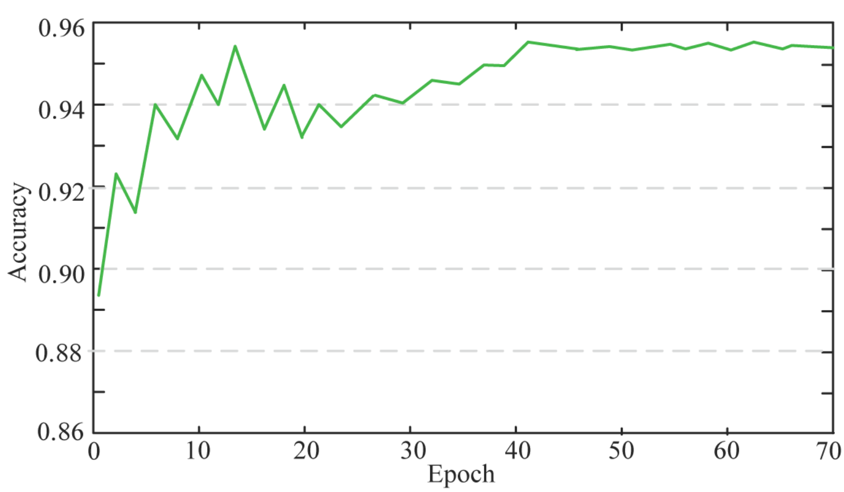


Figure 6. Experimental results of different learning rates

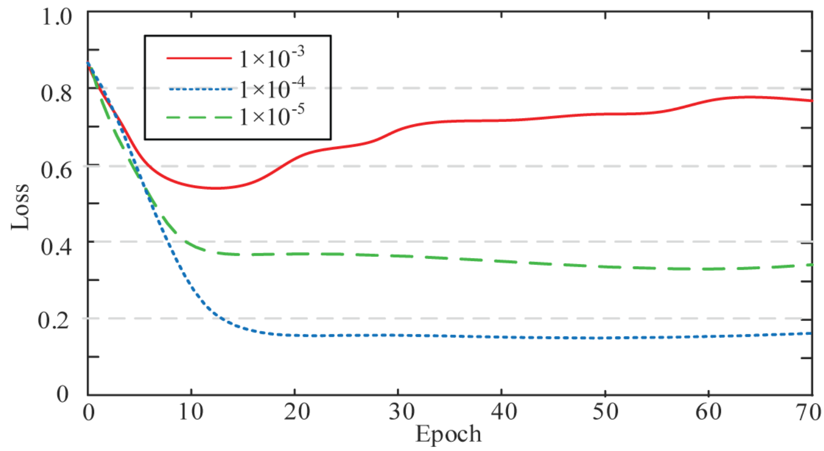
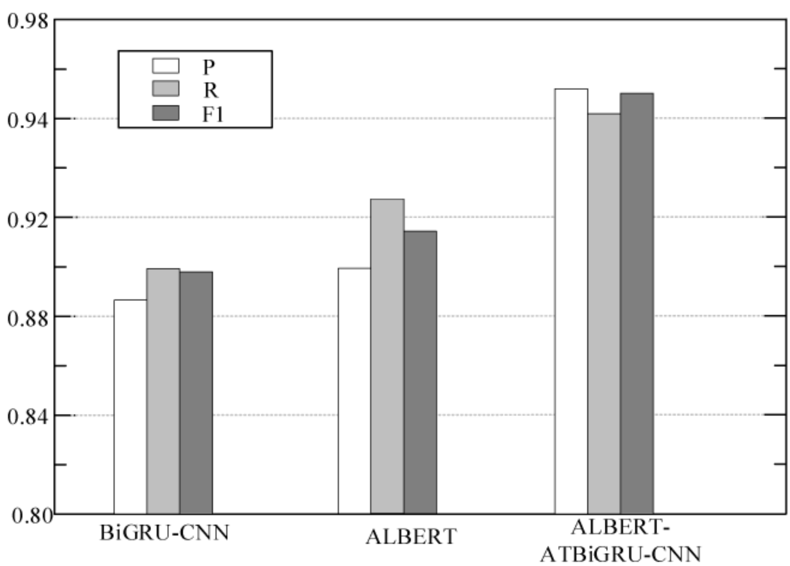


Table 2. Experimental results of three models

Model	P	R	F1
BiGRU-CNN	0.8841	0.9093	0.8965
ALBERT	0.9012	0.9245	0.9127
ALBERT-ATBiGRU-CNN	0.9552	0.9417	0.9484

Figure 7. Comparison of evaluation indexes of different methods



are higher than those of the comparative methods. This is because the proposed method obtains the contextualized word vector through the ALBERT pretrained language model, captures the semantic information through the combination of the positive and negative directions through the BiGRU model, and then captures the local characteristic information of the text using the CNN model. The effective combination of these two models improves the utilization rate of the text data.

CONCLUSION

In this paper, the author proposed a deep learning-based emotional analysis method of consumer comment text in the e-commerce big data environment to solve the problems, such as difficult information extraction, and insufficient feature extraction in the existing emotional analysis methods. The conclusions drawn from the experimental results are as follows: 1) The ALBERT pretrained language model obtains the contextualized word vector; 2) the BiGRU model is used to capture semantic information by combining the positive and negative directions, and then the CNN model is used to capture the local characteristic information of the text; 3) the accuracy of results is improved through the attention mechanism. The experimental results demonstrated that the proposed method is better than the comparative methods in the emotional analysis of consumer comments.

The text information is diverse, but in this study the author did not perform fine-grained identification and mining of text emotions, such as “happy, angry, sad, happy” and other categories, which will be more detailed in future research. Moreover, with the rapid development of e-commerce platforms, comments can contain a growing amount of information. In addition to the comment text, comments may also contain multimodal forms, such as emoticons, pictures, and videos. In the follow-up research of emotion analysis, the author will fuse multimodal information such as text and images with features to improve the emotional analysis performance of consumer comment information.

DATA AVAILABILITY

The paper includes the data the author used to support the findings of this study.

CONFLICTS OF INTEREST

The author declares that there is no conflict of interest regarding the publication of this paper.

FUNDING STATEMENT

This work was supported by the Science and Technology Research Project of Jiangxi Provincial Department of Education (GJJ212115).

REFERENCES

Christodoulides, G., Gerrath, M. H. E. E., & Siamagka, N. T. (2021). Don't be rude! The effect of content moderation on consumer-brand forgiveness. *Psychology and Marketing*, *38*(10), 1686–1699. doi:10.1002/mar.21458

Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., & Kuksa, P. (2011). Natural language processing (almost) from scratch. *Journal of Machine Learning Research*, *12*(1), 2493–2537.

Day, M. Y., & Lin, Y. D. (2017). Deep learning for sentiment analysis on google play consumer review. In *Proceedings of the 2017 IEEE International Conference on Information Reuse and Integration (IRI)* (pp. 382—388). IEEE. doi:10.1109/IRI.2017.79

Dhaoui, C., Webster, C., Tan, L. P., Norberg, P., & Fortin, D. (2017). Social media sentiment analysis: Lexicon versus machine learning. *Journal of Consumer Marketing*, *3*(10), 200–212. doi:10.1108/JCM-03-2017-2141

Haavisto, P., & Sandberg, B. (2015). "Man, this frustrates me:" Change of consumer emotions in online discussions. *Journal of Research in Interactive Marketing*, *5*(3), 1116–1123. doi:10.1108/JRIM-03-2014-0018

Han, Y., Lei, Y., & Chen, G. (2018). Consumption experience model and identification based on IWOM and emotional computing. *Cluster Computing*, *21*(1), 1023–1031. doi:10.1007/s10586-017-1017-y

Jamal, N., Chen, X., Al-Turjman, F. M., & Ullah, F. (2021). A deep learning–based approach for emotions classification in big corpus of imbalanced tweets. *Transactions on Asian and Low-Resource Language Information Processing*, *20*(3), 1–16. doi:10.1145/3410570

Kalchbrenner, N., Grefenstette, E., & Blunsom, P. (2014). *A convolutional neural network for modelling sentences.* ar Xiv preprint ar Xiv.

Kim, T., Kim, H., & Kim, Y. (2019). How do brands' Facebook posts induce consumers' e-word-of-mouth behavior? Informational versus emotional message strategy: A computational analysis. *Journal of Advertising Research*, *59*(4), 402–413. doi:10.2501/JAR-2019-027

Kim, Y. (2014). *Convolutional neural networks for sentence classification.* ar Xiv preprint ar Xiv.

Kujur, F., & Singh, S. (2018). Emotions as predictor for consumer engagement in YouTube advertisement. *Journal of Advances in Management Research*, *5*(1), 500–509. doi:10.1108/JAMR-05-2017-0065

Meng, L. M., Duan, S., Zhao, Y., Kevin, L., & Chen, S. (2021). The impact of online celebrity in livestreaming e-commerce on purchase intention from the perspective of emotional contagion. *Journal of Retailing and Consumer Services*, *6*(3), 102733. doi:10.1016/j.jretconser.2021.102733

Murgia, A., Ortu, M., Tourani, P., Adams, B., & Demeyer, S. (2018). An exploratory qualitative and quantitative analysis of emotions in issue report comments of open source systems. *Empirical Software Engineering*, *23*(1), 521–564. doi:10.1007/s10664-017-9526-0

Park, S., & Han, W. P. (2020). A webometric network analysis of electronic word of mouth characteristics and machine learning approach to consumer comments during a crisis. *El Profesional de la Información*, *29*(5), 200–210. doi:10.3145/epi.2020.sep.16

Rao, G., Huang, W., Feng, Z., & Cong, Q. (2018). LSTM with sentence representations for document-level sentiment classification. *Neurocomputing*, *30*(8), 49–57. doi:10.1016/j.neucom.2018.04.045

Rhanoui, M., Mikram, M., Yousfi, S., & Barzali, S. (2019). A CNN-BiLSTM model for document-level sentiment analysis. *Machine Learning and Knowledge Extraction*, *1*(3), 832–847. doi:10.3390/make1030048

Rúa-Hidalgo, I., Galmes-Cerezo, M., Cristofol-Rodríguez, C., & Aliagas, I. (2021). Understanding the emotional impact of GIFs on Instagram through consumer neuroscience. *Behavioral Sciences (Basel, Switzerland)*, *11*(8), 108–113. doi:10.3390/bs11080108 PMID:34436098

Ruiz-Mafe, C., Bigné-Alcañiz, E., & Currás-Pérez, R. (2020). The effect of emotions, eWOM quality, and online review sequence on consumer intention to follow advice obtained from digital services. *Journal of Service Management*, *31*(3), 465–487. doi:10.1108/JOSM-11-2018-0349

Simpson, L., Taylor, L., O'Rourke, K., & Shaw, K. (2011). An analysis of consumer behavior on Black Friday. *American International Journal of Contemporary Research*, *2*(4), 200–209.

Tang, D., Qin, B., & Liu, T. (2016). *Aspect level sentiment classification with deep memory network.* arXiv preprint arXiv.

Tosun, P., & Sezgin, S. (2021). Voluntary simplicity: A content analysis of consumer comments. *Journal of Consumer Marketing*, *38*(5), 484–494. doi:10.1108/JCM-04-2020-3749

Wang, M., Zhou, J., & Lin, H. (2021). A sentiment analysis of the influence of service attributes on consumer satisfaction. *Journal of Intelligent & Fuzzy Systems*, *40*(6), 10507–10522. doi:10.3233/JIFS-201207

Wu, H. H., Tipgomut, P., Chung, H. F. L., & Chu, W. K. (2019). The mechanism of positive emotions linking consumer review consistency to brand attitudes: A moderated mediation analysis. *Asia Pacific Journal of Marketing and Logistics*, *2*(4), 120–126. doi:10.1108/APJML-03-2019-0224

Wu, Z., & Ong, D. C. (2021). Context-guided bert for targeted aspect-based sentiment analysis. In *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 14094—14102). doi:10.1609/aaai.v35i16.17659

Xu, A., Ying, P., & Wang, F. (2019). Research on the consumer reviews of JD bookstore based on text mining technology. In *Proceedings of the 2019 4th International Conference on Big Data and Computing* (pp. 209—214). doi:10.1145/3335484.3335512

Xu, X. (2020). Examining the role of emotion in online consumer reviews of various attributes in the surprise box shopping model. *Decision Support Systems*, *13*(6), 1133–1144. doi:10.1016/j.dss.2020.113344

Yu, J. (2014). We look for social, not promotion: Brand post strategy, consumer emotions, and engagement. *International Journal of Media & Communication*, *1*(2), 28–37.

Yu, J., Jiang, J., & Xia, R. (2018). Global inference for aspect and opinion terms co-extraction based on multi-task neural networks. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *27*(1), 168–177. doi:10.1109/TASLP.2018.2875170

Zhao, W., Ye, J., Yang, M., Lei, Z., Zhang, S., & Zhao, Z. (2018). Investigating capsule networks with dynamic routing for text classification. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing* (pp. 3110—3119). Academic Press.

Zhou, Q., Xu, Z., & Yen, N. Y. (2019). User sentiment analysis based on social network information and its application in consumer reconstruction intention. *Computers in Human Behavior*, *10*(1), 177–183. doi:10.1016/j.chb.2018.07.006