JCSE Journal of Computing Science and Engineering

Regular Paper

Journal of Computing Science and Engineering,
Vol. 18, No. 2, June 2024, pp. 80-92

# Sentiment Classification Model of Fresh Agricultural Product Comments based on Semantic Structure-Combined Dictionary Optimization

**Yindong Dong***, **Shundong Shao, Guohua Fan, Deli Chen, Yu Zhang, and Xin Luo**
School of Information and Artificial Intelligence, Anhui Agricultural University, Hefei, China
**dongyindong66@163.com, 873439366@qq.com, fghvan@yeah.net, delichen0711@outlook.com, 2582921962@qq.com, 3154055672@qq.com**

## Abstract
The increasing popularity of various types of online services has led to a growing number of consumers purchasing fresh agricultural products through online platforms. To address the discrepancy between the actual emotional disposition expressed in fresh produce reviews and the sentiment labels that are assigned to them, the current paper proposes a sentiment classification model based on the optimization and integration of a composite semantic structure dictionary. First, the TextRank algorithm is used to extract sentiment-bearing keywords from fresh produce reviews, which are used in the formation of a sentiment dictionary. Sentiment category matching scores are then calculated based on this dictionary. Concurrently, the text vectors processed by bidirectional encoder representations from transformers (BERT) are fed into an enhanced deep pyramid convolutional neural network (DPCNN) and a bidirectional long short-term memory network (BiLSTM), which are tasked with extracting global semantic information and contextual information, respectively. Subsequently, a weighted attention mechanism is used to consolidate these feature representations. Finally, a sentiment classification model, which is designated as Scores-BERT-IDPCNN-BiLSTM-Attention (S-BIDBA), is trained by incorporating sentiment category matching scores. The experimental results obtained herein demonstrate that S-BIDBA significantly enhances the accuracy of sentiment classification for online reviews of fresh agricultural products.

**Category:** Computer Graphics / Image Processing

**Keywords:** Fresh produce reviews; Sentiment classification; Keyword extraction; Improved DPCNN

## I. INTRODUCTION

With the rapid development of the e-commerce industry, fresh agricultural products have come to be delivered to thousands of households through online networks. However, the development of online platforms is to some extent restricted by factors such as short preservation cycles, strict logistics distribution, and large differences in the taste of fresh agricultural products. Due to the diversity associated with various features of fresh agricultural products and differences in consumers' personal preferences, some consumers may leave emotional comments based on their real product experience after purchasing fresh agricultural products online. These reviews include comments on the price, quality, and logistics services of fresh agricultural products. In this process, objective comments about fresh agricultural products are often accompanied by subjective evaluations by consumers,

Open Access http://dx.doi.org/10.5626/JCSE.2024.18.2.80 http://jcse.kiise.org



**Received** 09 February 2024; **Accepted** 10 June 2024
*Corresponding Author


pISSN: 1976-4677 eISSN: 2093-8020

and different evaluations of fresh agricultural products also contain different degrees of emotional information from consumers. Therefore, having a deep understanding of emotional evaluations of specific fresh agricultural products can help consumers understand the real experiences and feelings of other buyers. Such an understanding can help consumers better judge whether the quality, price, and logistics services of fresh agricultural products meet their expectations, which can ultimately support them in making more informed purchasing decisions, which substantially affects consumers' purchase intentions. Meanwhile, for producers, understanding consumers' satisfaction with the product and the value direction of the product itself can help producers make timely adjustments to their production strategies and optimize product quality to meet market demand. For businesses, having a deeper understanding of their consumers' preferences and needs for fresh agricultural products can help them improve their sales methods and lead their product orientation in a way that is more in line with market demand.

In this context, reviews of fresh agricultural products generally consist of two parts: a star rating and text. The emotional tendencies present in the text comments are typically related to the star rating, and people tend to divide evaluation categories according to star ratings, where reviews of more than 3 stars, 3 stars, and fewer than 3 stars are defined as good, medium, and negative reviews, respectively. However, after crawling through a large number of fresh agricultural product reviews, it is found that the star emotion evaluation given in the rating system is not always necessarily consistent with the actual emotional tendency expressed in the text reviews. As an example, a fresh produce review might read, "The apples in this store are delicious!" According to the actual content of the text analysis, the star rating should be more than 3 stars, but due to subjective randomness or other objective reasons, the user might accompany such text comments with a star rating of 3 stars, which will accordingly affect how shopping systems categorize reviews. There will also be similar cases where the actual content of the review is a negative review but the system will classify it as a medium review. Therefore, classifying fresh agricultural product reviews by star rating alone has certain limitations, making it necessary to further consider the semantic and emotional information of customers in reviews for comprehensive analysis. To solve the problem of emotion classification in online reviews of fresh agricultural products, the current paper proposes an emotion classification model that integrates emotion dictionary and semantic structure information, and which can deeply explore the effective emotion information present in reviews of fresh agricultural products.

The main contributions of this paper are as follows:

(1) The IDPCNN algorithm improves upon the deep pyramid convolutional neural network (DPCNN) algorithm by enhancing the model's ability to acquire semantic information from reviews in a multi-scale manner through splicing and average pooling operations.
(2) The IDPCNN and bidirectional long short-term memory network (BiLSTM) are used to extract fresh produce reviews with global information features and contextual information features, respectively. The collaborative combination of these two algorithms allows the model to learn accurate semantic information and language patterns.
(3) The TextRank algorithm is employed to construct a sentiment-emotion lexicon for fresh produce reviews, thus assisting the model in correcting the sentiment information representations. The attention mechanism is then used to further scrutinize the feature confidence and integrate the sentiment keyword scores into the sentiment classification of the review text.

## II. RELATED WORK

This sentiment classification can be understood as a task demanding the categorization of the sentiment tendencies of texts featuring subjective opinions [1]. The mainstream classification methods can be roughly divided into three types: those based on sentiment lexicons [2-6], those based on traditional machine learning [7-10], and those based on deep learning [11, 12].

Sentiment lexicon-based methods mainly rely on a sentiment lexicon to determine the sentiment polarity of a review; this method requires a lot of human intervention, such as in constructing the lexicon and formulating judgment rules. In this method, it is typically necessary to rely on third-party sentiment resource databases to determine sentiment polarity. However, with the emergence of a large number of new and unknown words, it is necessary to continuously expand the sentiment lexicon to meet the needs of sentiment classification, which not only consumes time and resources, but also reduces the efficiency of text classification.

Traditional sentiment classification techniques use traditional machine learning algorithms to train sentiment classifiers by extracting and selecting text features and optimizing classification algorithms to recognize and classify text sentiment. Traditional methods that are commonly used for this purpose include Naive Bayes, support vector machines [13, 14], etc. The feature selection and extraction involved in these methods require a lot of manual involvement, which is time-consuming and labor-intensive, and medium selection and feature extraction also require rich knowledge and experience, as lacking such expertise could cause some important features to be overlooked, ultimately ending in poor classification results.

In recent years, deep learning-based methods have come to be widely used in the field of sentiment classification due to their ability to automatically extract features and their powerful representation capabilities. Studies have shown that deep learning methods outperform traditional feature-based methods in many sentiment classification tasks. In deep learning methods, text data is first vectorized to represent it. Commonly used text vectorization methods are Word2Vec, Glove, BERT (bidirectional encoder representations from transformers), etc. Mikolov et al. [15] first proposed the Word2Vec algorithm and used it for text categorization tasks, which solved the sparsity and dimensionality catastrophe problems involved in the traditional methods. Pennington et al. [16] proposed the Glove model to represent text sequences with features. However, the word vectors generated by the above methods are static, as they ignore the changes of word meanings in different contexts, and they cannot solve the problem that occurs with words that have multiple meanings. For example, in the two phrases "watermelon is delicious" and "watermelon is nice to look at", watermelon does not have the same meaning, but if the above methods are used, the word vectors of "watermelon" will be the same vector. To address this problem, Peters et al. [17] proposed the GPT and ELMO model, which—to some extent—solved the problem wherein words have multiple meanings. On this basis, to better obtain text semantic information, Devlin et al. [18] proposed a BERT pre-training model based on the transformer. The neural networks that are commonly used in extracting text sentence features are convolutional neural network (CNN) [19, 20], recurrent neural network (RNN) [21, 22], long short-term memory network (LSTM) [23-25], etc. Qorich and El Ouazzani [26] used a CNN model to classify the textual sentiment of the Amazon review dataset, and this model achieved a high accuracy rate. Yao [27] proposed a sentiment classification method that combines CNN and LSTM due to the fact that CNN ignores the correlation between discrete words, and they achieved good results on the hotel review dataset. However, the LSTM model can only consider the relationship above, while BiLSTM [28] has the property of bidirectional propagation, which can simultaneously capture the semantic information of the preceding and following texts; this precisely solves the deficiency of LSTM. Guo et al. [29] proposed an online review sentiment classification method combining CNN and BiLSTM, and this method achieved better results. However, traditional CNN models are prone to the problem of discrepancy in features due to the size or location of different regions. DPCNN [30] exhibits better performance and processing capability than CNN models. Bahdanau et al. [31] introduced an attention mechanism in text processing that can be used to weight different scales of information in the input, which allows the trained model to focus on important information. Chang et al. [32] proposed a Two-channel hierarchical short text classification model based on the attention mechanism, and this model achieved excellent performance on an open-source dataset. Zhang et al. [33] proposed a directed sentiment text categorization model that combined the attention mechanism with double BERT, and this model was shown to outperform its comparative model on relevant datasets. Su et al. [34] proposed an LMAEB-CNN model combining BiLSTM and CNN with a multi-head attention mechanism for comments in a microblogging community, which not only solves the overfitting problem but also improves the accuracy of sentiment polarity classification.

In this paper, for review text of fresh produce, the BERT pre-training model is first used to vectorize the produce review text, after which the vectorized text is input into IDPCNN and BiLSTM to obtain sentence-level semantic information and contextual information, respectively. Then, the attention mechanism helps focus on the valid sentiment information in the reviews, and the softmax function is also used to classify the sentiment of the reviews. Further, to enhance the expression of the semantic structure information in terms of sentiment in the learning process, the sentiment matching score is inputted into the model as a degree of validity information, which serves to help the model better understand the sentiment information of the comments and further enhance the stability of the model.

## III. CONSTRUCTION OF AN EMOTION LEXICON AND AN EMOTION CLASSIFICATION MODEL

### *A. Construction of an Emotional Lexicon*

TextRank is a graph-based model sorting algorithm that can be used for tasks such as keyword extraction, text summarization, and importance ranking of single or multiple documents [35]. It represents an improvement upon the PageRank algorithm (a graph ranking algorithm that is widely used in search engines) and is particularly suited to text data processing. The principle of TextRank can be summarized in these steps: First, the input text is divided into several sentences, and each sentence is partitioned and lexically annotated, e.g., assuming that the text is $T$, and the partitioned text is $T = [S1, S2, ..., Sn]$; next, an undirected graph $G = (V, E)$ is constructed in which each node $V$ represents a word, and each edge $E$ represents the strength of semantic connection between nodes and nodes; then, the PageRank algorithm is used to sort the nodes in the graph and get the weight of each node, which represents the degree of importance of the word in the text; and finally, according to the node's weight of the keywords, the first $K$ words are selected as keywords. The principle of the algorithm is shown in Eq. (1):

**Table 1.** Selected keyword examples

| Keyword | Category |
|---|---|
| Love, Special, Careful, Food, Satisfaction, Genuine, Like, Delicious, Good.... | Good review |
| Understand, Average, Pity, Celebrate, Buy with caution, Forget it, Make do, Okay, Not bad.... | Medium review |
| Bamboozled, Returned, Failed, Deceived, Disappointed, Complained, Refunded, Poor, Difficult to eat.... | Negative review |

$$WS(V_i) = (1-d) + d \times \sum_{V_j \in In(V_i)} \frac{w_{ji}}{\sum_{V_k \in Out(V_j)} w_{jk}}, \qquad (1)$$

where $V_j$, $V_i$ are the keywords, $WS(V_i)$ is the weight of node $V_i$, $In(V_i)$ denotes the set of in-degree nodes, $Out(V_j)$ denotes the set of out-degree nodes, $w_{ji}$ is the weight of the edge between nodes $V_i$ and $V_j$, and $d$ is the damping coefficient, which typically has a value of 0.85. In practice, TextRank algorithms usually make certain correlations between the nodes adjustments to overcome some common problems, such as the weight bias between long and short sentences as well as the problem of synonyms and near-synonyms.

The process in this paper begins by using the TextRank algorithm to extract keywords from a large number of sentiment-tagged online reviews of fresh produce. According to the sentiment tags of the reviews, the keywords are added to the lists of "good review keywords", "medium review keywords" and "negative review keywords"; after screening and processing, some irrelevant words—such as deactivated words and common words—are removed to retain the keywords that are related to the sentiment categories of the reviews; the filtered keywords are retained and used in the construction of a complete emotion dictionary containing good reviews, medium reviews and negative reviews. This sentiment dictionary is used for the subsequent fusion of feature information to enhance the classification effect of the model. Table 1 presents some sample keywords.

Here, a function "calculate_keyword_scores (CKS)" is defined for calculating keyword scores, which accepts the list of datasets and the created sentiment dictionary as inputs, and—for each review—it iterates over each sentiment category (good, medium, negative) in the sentiment dictionary that matches whether the comment contains the corresponding keyword. If it does, the score for that sentiment category is incremented. The keyword match scores for positive, neutral, and negative reviews are eventually stored in an array. The keyword matching scores will assist the model in better understanding the sentiment tendencies of fresh produce reviews.

## B. Construction of the Model

The research framework of Scores-BERT-IDPCNN-BiLSTM-Att (S-BIDBA) is shown in Fig. 1, where it can be seen that it consists of four modules: input layer, feature extraction layer, feature fusion layer, and output layer. First, this model adopts the BERT pre-training model to transform the text of fresh produce reviews into

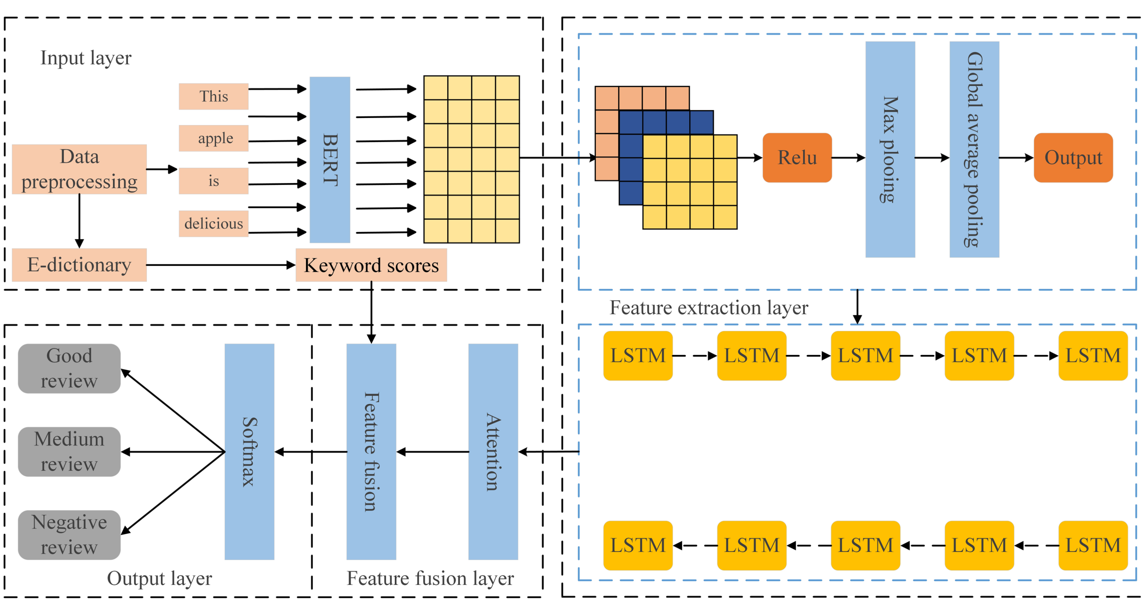



**Fig. 1.** Research framework.

a word vector matrix and then inputs the word vectors into IDPCNN and BiLSTM to obtain the feature information. IDPCNN is responsible for extracting the local feature information and global feature information, while BiLSTM is responsible for extracting the contextual feature information; the attention mechanism is then connected to make the model pay more attention to the input sequences of the key information and to improve the performance and generalization ability of the model, and the sentiment feature information extracted from IDPCNN and BiLSTM is finally fused with the keyword matching score vector for feature fusion, at which point the fused feature vector is inputted into the softmax layer to obtain the sentiment classification result.

***1) Input Layer***

The input layer is the first layer, and it is where the model receives information and contains the index of the text sequence, the text sequence template, the sequence of keyword matching scores for the text sentence, and operations such as normalizing the keyword scores by MinMaxScaler along with vectorizing the input text sequence to obtain a vector for each word that subsequently composes a vector representation of the entire sentence.

Special tags are used to represent the beginning and end of a text sentence. Next sentence prediction (NSP) and masked language model (MLM) are two unsupervised pre-training tasks in the BERT model that are used to learn semantic and contextual information in text, thereby generating high-quality word vector representations (Fig. 2). The Mask strategy involved in this process is shown in Fig. 3. When the input sentence is "This apple is big", BERT will treat the word "apple" as one unit and carry out Mask in a unified manner. This strategy enables BERT to better understand the context information in the text, thus achieving better performance in downstream tasks, and ultimately solving the polysemous problem of the word in the text.

Since the BERT pre-training model has been trained on a large amount of corpus, the output word vector can

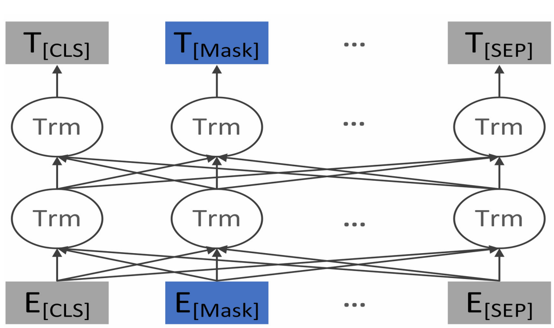



**Fig. 2.** BERT pre-trained model.

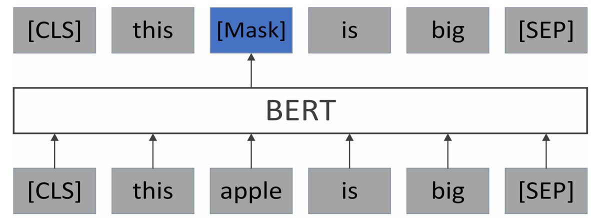



**Fig. 3.** Mask strategy.

contain most of the information of the text sequence, and it therefore has strong generalization ability due to the requirements of fresh produce in terms of the storage, transportation, and intrinsic quality of the product, which will make the review content in the structure and the semantics of the review more complex, and the BERT pre-training model has the advantage of both language comprehension and semantic representation, so in this paper, we use the BERT pre-training to map the review into the word vector space to obtain a vector representation of the text sentence of the review of the fresh produce.

***2) Feature Extraction Layer***

***a) IDPCNN model:*** In the feature extraction layer, the model is by IDPCNN as part of the feature extraction, and DPCNN introduces a deep pyramid pooling structure compared to CNN, which enhances the model's ability to extract global information to a certain extent, with the ultimate intention of further enhancing the ability of DPCNN to obtain global semantic information. In this way, this paper makes an improved DPCNN, which is done by splicing multiple pooling layer outputs to form a connected channel, which in turn obtains multi-scale semantic information, and then performs a global average pooling operation; the obtained outputs with the fully connected layer are used as global semantic information. The structure of IDPCNN is shown in Fig. 4.

First, BERT is used to convert the text into word

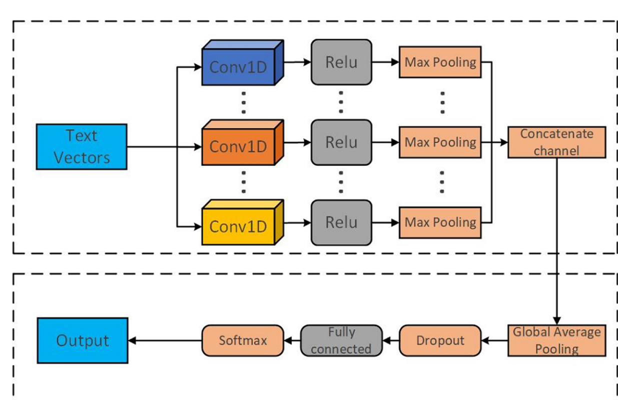



**Fig. 4.** Improved DPCNN model.

vectors, and the local features of the textual information are extracted by different sizes of convolutional kernels, while the convolutional manipulation is as shown in Eq. (2) as follows:

$$c_i = relu(Wx_{i:i+k-1} + b)\,, \tag{2}$$

where *relu* is the activation function, $x_{i:i+k-1}$ is the input local features, and $W$ and $b$ denote the weights and bias terms, respectively. The local features are then pooled using maximum pooling, and finally, the output of the maximum pooling layer is spliced to form a connected channel, while the pooling operation is performed on the output of the connected channel through the global average pooling layer. The global average pooling operation is shown in Eq. (3):

$$gap_i = \frac{\Sigma_{j=1}^{N} \mathrm{CC}_{[i,j]}}{N}\,, \tag{3}$$

where $N$ is the length of the sequence and $\mathrm{CC}_{[i,j]}$ denotes the value of the data point at position $j$ on a particular channel $i$. Finally, the output values are concatenated with the fully connected layer to obtain the IDPCNN output vector, which is computed as shown in Eqs. (4) and (5):

$$v_i = gap(CC_i)\,, \tag{4}$$

$$D = \{v_1,\ v_2,\ \ldots,\ v_n\} \tag{5}$$

where $v_i$ denotes the global average pooling result for the ith sample, and after which the features are aggregated to obtain the final feature vector $D$.

***b) BiLSTM model:*** IDPCNN only extracts the local and global features of the textual information whereas BiLSTM can effectively extract the contextual information of textual features; the combination of the two can enhance the acquisition of effective information of the model in terms of semantic and structural aspects. Fig. 5 shows the structure of the BiLSTM model, which consists of two mutually independent LSTM layers: one is a forward LSTM and the other is a reverse LSTM [28]. The forward LSTM starts from time step 0 and processes the information of each moment in the input sequence in the order of increasing time steps, with an output of $\overrightarrow{h_t}$, while the reverse LSTM starts from the end of the time step and processes the information of each moment in the input sequence in the order of decreasing time steps, with an output of $\overleftarrow{h_t}$. The outputs of forward LSTM and reverse LSTM are spliced at each time step to form the output of BiLSTM $h_i = concat(\overrightarrow{h_t},\ \overleftarrow{h_t})$, after which all the outputs can be represented by Eq. (6):

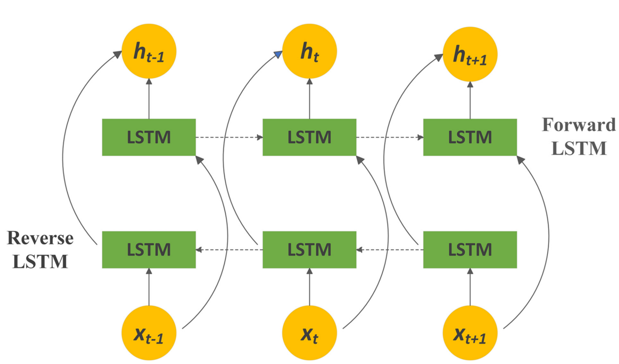



**Fig. 5.** BiLSTM model structure.

$$G = h_1 + h_2 \ldots + h_n\,. \tag{6}$$

Compared to LSTM, BiLSTM can simultaneously account for the past and future information in a sequence, which allows it to understand the context more comprehensively, and then better capture the context information in the sequence and extract richer feature representations. Therefore, this paper chooses BiLSTM as a part of the feature extraction module.

***3) Feature Fusion Layer***

In this layer, the model will pay attention to the semantic and structural information through the attention mechanism, which enhances the target word by taking the target word as a query while taking the other words in its context as keys, and then calculating the similarity between them by summing them up with a certain weight to combine the value of the context word with the original value of the target word, ultimately enhancing the target word's semantic representation [31]. The input to the attention mechanism consists of semantic vector representations of the target word and the individual words of the context, which are linearly transformed to obtain the corresponding Query, Key and Value vector representations. The weights are determined by calculating the similarity between the Query vector and the individual Key vectors, while the Value vectors of the context words are weighted and fused into the Value vector of the target word to obtain the enhanced semantic vector representation of the target word as the output of the attention. This calculation process is as shown in Eq. (7):

$$Attention(Q,\ K,\ V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V\,, \tag{7}$$

where $Q$, $K$, and $V$, respectively represent the query, key, and value, and $\sqrt{d_k}$ is the dimension of the key, which is used to scale the result of the dot product to prevent the value from being too large. In this paper, we use the attention mechanism to weight the different features of the input text sequence so that, during training, the model can focus on the important information to enhance the accuracy of model training. The role of this layer is to fuse the feature vector extracted by IDPCNN, the feature vector extracted by BiLSTM, the output vector of the attention mechanism, and the vector of the keyword matching score to form the output of feature information. First, the feature vector extracted by DPCNN, the feature vector extracted by BiLSTM, and the output vector of the attention mechanism are spliced and then connected with the keyword-matching score vectors of different categories of comment texts through concatenation to obtain the final text representation vector Z. The equations representing these processes are shown in (8)-(9):

$$H = D + G + A\,, \tag{8}$$

$$Z = concat[H,\ KS]\,. \tag{9}$$

***4) Output Layer***

In the output layer, the model inputs the final text feature vector to the fully connected layer for classification and adopts softmax as the activation function to obtain the text classification results; the calculation process is shown in Eq. (10):

$$y = softmax(WZ + b)\,, \tag{10}$$

where $y$ denotes the result of text classification, $W$ is the weight, and $b$ is the bias term. The cross-entropy loss function is used for the loss function during training, and the calculation process used here is shown in Eq. (11):

$$L = -\Sigma y\_i \log(p\_i)\,, \tag{11}$$

where $L$ is the value of the cross-entropy loss function, $y\_i$ is the ith element of the true label, which indicates the actual category of the sample, and $p\_i$ is the ith element of the predictive probability of the model, which indicates the predictive probability of the model for the samples belonging to the respective category. The training is concluded to obtain the sentiment classification result of the input sequence.

## IV. EXPERIMENT AND RESULT ANALYSIS

### A. Data Sources

To assess the efficacy of the model proposed in this article, online reviews of agricultural products from JD Fresh and Amazon Fresh were collected, and these were used to establish a JD Fresh and Amazon Fresh dataset comprising reviews of fresh agricultural products such as fruits, eggs, meat, soy products, seafood, etc. The JD Fresh dataset comprises 8,000 positive reviews, 5,399 positive reviews, and 7,213 negative reviews, resulting in a total of 20,612 online comments. The Amazon Fresh dataset comprises 5,000 positive reviews, 2,043 positive reviews, and 3,014 negative reviews, for a total of 10,057. For a detailed overview of the data, please refer to Table 2. The data was subjected to manual analysis, and the obtained results were found to be consistent with the expected situation. The actual emotional tendencies of positive and negative reviews did not align with the star rating accounting for only 1.05% and 0.71%, respectively. The actual emotional tendencies of Medium reviews did not align with the star rating either, accounting for 6.54%.

**Table 2.** Data details

| Emotional categories | Quantities | Number of stars not met | Percentage |
|---|---|---|---|
| Good review | 13,000 | 136 | 1.05 |
| Medium review | 7,442 | 487 | 6.54 |
| Negative review | 10,227 | 73 | 0.71 |

The data format is presented in Table 3. The dataset contains two contents: comments and labels. A review is a comment sentence, while a label is an emotion category that includes three categories: 0–2, where 0 represents a negative comment, 1 represents a medium comment, and 2 represents a positive comment. In the experiment, the dataset is divided into a training set, a verification set, and a test set according to an 8:1:1 ratio through random extraction.

**Table 3.** Data format

| Review | Label |
|---|---|
| Favorable sentences | 2 |
| Median sentence | 1 |
| Bad sentence | 0 |

### B. Data Preprocessing

Due to the large number of online review data of fresh agricultural products, which involves user reviews from different regions, a mix of casual form and serious oral expression, and a large number of irrelevant data point, it is necessary to pre-process the experimental text data before starting the experiment. First, all punctuation marks and special characters are removed such, and only information with semantic and emotional value is retained. Secondly, the Jieba word segmentation tool is used to segment the text; finally, a de stop word operation is performed on the text.

### C. Evaluation Indicators

In this paper, four evaluation metrics are used to measure the performance of the model: accuracy, precision, recall, and F1 value. Accuracy indicates the ratio of the number of correctly classified samples to the total number of samples. The precision rate indicates the ratio of the number of correctly classified samples to the number of samples that are classified as positive examples among all samples that are classified as positive examples. Recall rate indicates the proportion of the number of correctly classified samples to the total number of samples that are classified as positive cases among all samples that are classified as positive cases. The value of F1 is the reconciled mean of the precision rate and the recall rate. Their specific calculation processes are shown in Eqs. (12)-(15):

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN}, \tag{12}$$

$$Precision = \frac{TP}{TP+FP}, \tag{13}$$

$$Recall = \frac{TP}{TP+FN}, \tag{14}$$

$$F1 = \frac{2Precision * Recall}{Precision+Recall}. \tag{15}$$

In the formula, TP represents true cases, that is, the number of samples that the classifier correctly predicts to be positive cases which are indeed positive cases; TN represents true negative cases, that is, the number of samples that the classifier correctly predicts to be negative cases that are indeed negative cases; FP represents false positive examples, that is, the number of false samples that the classifier predicts to be positive examples which are actually negative examples; and FN represents false negative cases, that is, the number of error samples that the classifier predicts to be negative cases which are actually positive cases.

**Table 4.** Model parameters

| Parameter name | Value |
|---|---|
| Word vector dimension | 768 |
| Activation function | ReLU |
| Batch size | 128 |
| Epoch | 7 |
| Optimizer | Adam |
| Dropout | 0.3 |
| Learning rate | 2e-5 |
| Loss function | sparse_categorical_crossentropy |

### D. Experimental Environment and Parameter Settings

The experiments in this paper were conducted on the Ubuntu 22.04.2 platform, the development tool was Visual Studio Code, the software environment used was Python 3.7.16, and the deep learning framework used was TensorFlow 2.13.0 along with third-party libraries such as Keras 2.13.1, Jieba 0.42.1, numpy 1.24.3, etc. The hardware environment is as follows: Intel i7-12700 CPU; NVIDIA GeForce RTX 4070 GPU; and 12 GB of RAM. Table 4 lists the optimal parameters and settings under multiple experiments.

### E. Comparison of Experimental Models and Analysis of Results

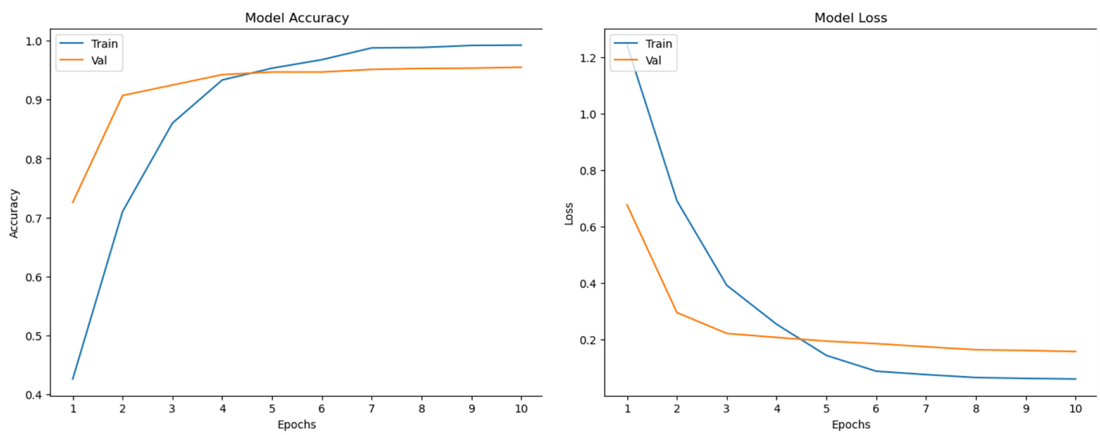
For model verification, baseline models LSTM, BiLSTM, CNN, DPCNN, and BERT, and the more advanced emotion classification models CNN-LSTM [36], Bert-CNN-LSTM [37], and S-BIDBA, are selected for the comparison of performance indicators on the test set. All models utilize the same dataset and the optimal parameters. Tables 5 and 6 present the comparative results obtained when using the model on the Jingdong Fresh and Amazon

**Table 5.** Comparison of evaluation indicators of the model on the Jingdong Fresh dataset (unit: %)

| Model name | Accuracy | Precision | Recall | F1 |
|---|---|---|---|---|
| LSTM | 87.70 | 87.36 | 87.51 | 87.41 |
| BiLSTM | 89.34 | 89.33 | 89.27 | 89.19 |
| CNN | 90.16 | 90.09 | 90.01 | 90.04 |
| DPCNN | 91.80 | 91.74 | 91.68 | 91.61 |
| CNN-LSTM | 93.05 | 93.08 | 92.43 | 92.70 |
| BERT | 93.53 | 93.45 | 93.23 | 93.25 |
| BERT-CNN-LSTM | 94.70 | 95.12 | 94.04 | 94.46 |
| S-BIDBA | 96.91 | 96.95 | 96.20 | 96.55 |

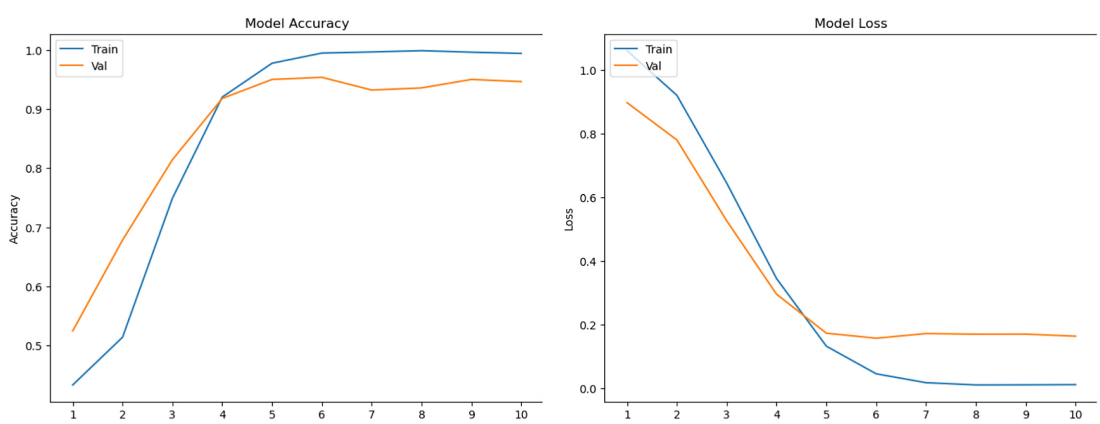
**Table 6.** Comparison of evaluation indicators of the model on the Amazon Fresh dataset (unit: %)

| Model name | Accuracy | Precision | Recall | F1 |
|---|---|---|---|---|
| LSTM | 83.91 | 83.59 | 83.73 | 83.63 |
| BiLSTM | 85.48 | 85.47 | 85.41 | 85.34 |
| DPCNN | 89.03 | 89.06 | 88.44 | 88.70 |
| CNN | 90.23 | 90.15 | 89.94 | 89.96 |
| CNN-LSTM | 92.06 | 92.10 | 91.39 | 91.72 |
| BERT | 92.72 | 92.76 | 92.04 | 92.38 |
| BERT-CNN-LSTM | 93.49 | 93.53 | 92.80 | 93.14 |
| S-BIDBA | 94.64 | 94.49 | 93.75 | 94.07 |




(a) (b)

**Fig. 6.** Trends of accuracy and loss rate on the Jingdong Fresh dataset: (a) accuracy and (b) loss rate.




(a) (b)

**Fig. 7.** Trends of accuracy and loss rate on Amazon Fresh dataset: (a) accuracy and (b) loss rate.

Fresh datasets, respectively. Figs. 6 and 7 illustrate the evolving patterns of accuracy and loss rate observed in the training set and validation set of the S-BIDBA on the Jingdong Fresh and Amazon Fresh datasets, respectively.

As can be seen in Figs. 6 and 7, the accuracy rate of the training and validation sets of the S-BIDBA model starts to show a gradually increasing trend that stabilizes after the 7th round; the loss rate shows a decreasing trend at the beginning, and it gradually stabilizes with an increase in epochs. The accuracy and loss rate of the training and validation sets increase and decrease at the same time, and they also tend to stabilize, which shows that the model has converged to a better state and has a certain degree of generalizability.

Tables 5 and 6 respectively demonstrate that the S-BIDBA model exhibits the highest performance indicators on the Chinese Jingdong Fresh dataset and the English Amazon Fresh dataset, outperforming other comparison models. This indicates that the S-BIDBA is effective for datasets with different language and cultural backgrounds, suggesting that it has robust performance and a certain universality.

The LSTM model and BiLSTM model use their respective network structures for feature extraction, after which a dropout layer and a fully connected layer are used for text sentiment categorization, and the activation function taken by the fully connected layer is softmax. Since the LSTM model is very sensitive to the input order, if the input data order is not reasonable, the stability of the model will be affected; meanwhile, BiLSTM has a bi-directionality, which allows it to take into account both forward and backward information, ultimately better extracting the features in the sequence, which helps solve the order-sensitive problem that exists in the LSTM model. Therefore, BiLSTM performs better than LSTM. The CNN model utilizes a convolutional kernel to extract text features, and the same softmax activation function is taken on the fully connected layer to classify the text for the sentiment. LSTM and BiLSTM are both affected by the input order and sequence length, resulting in a weaker extraction of local features than CNN. DPCNN exhibits a better performance than CNN in the text sentiment classification task due to the use of multilayer pooling and convolutional operations. However, the performance of DPCNN in the Amazon Fresh dataset is slightly inferior to that of CNN. This may be attributed to either the limited number of datasets used or the absence of complex characteristics. BERT uses a pre-trained model that is connected to the softmax layer to classify the sentiment of the text. BERT outperforms the four benchmark models in the model evaluation metrics because it has strong representativeness stemming from its basis upon the Transformer architecture.

The CNN-LSTM model also enables the extraction of local features from text, such as phrases and word groups, which is highly beneficial for the identification of emotional cues in comments. In the BERT-CNN-LSTM model, the contextual semantics of the text are extracted by BERT, combined with the local features extracted by CNN and the sequence dependence captured by LSTM. This ultimately allows the model to more comprehensively understand the sentiment of the comment text, resulting in improved performance. The evaluation indexes of the S-BIDBA model proposed in this paper are higher than those of the current advanced CNN-LSTM model and BERT-CNN-LSTM model. This is attributed to the fact that the S-BIDBA combines the deep semantic understanding of BERT, the global and local feature extraction of IDPCNN, the sequence modeling of BiLSTM, and the weighting of key features of attention. The text category matching score is also fused on this basis. This score not only provides additional feature information but also enhances the stability and comprehension of the model's emotional responses. Altogether, the results of the analysis indicate that the S-BIDBA exhibits certain advanced characteristics.

## F. Ablation Experiment

To verify the effectiveness of the main modules in the S-BIDBA, we carry out a comparative analysis through ablation experiments, and the comparative results of the ablation experiments in each index are given in Table 7.

(1) BDBA: The keyword matching score module is deleted, the original DPCNN should be adopted, and the remaining modules should be kept unchanged.
(2) S-BDBA: The IDPCNN feature extraction module is deleted, the original DPCNN is employed, and the remainder of the modules remain unaltered.
(3) S-BBA: The IDPCNN feature extraction module should be deleted whereas the remaining modules should be retained.
(4) S-BIDA: The BiLSTM feature extraction module is deleted, and IDPCNN is adopted instead. The remaining modules remain unchanged.

In Table 7, it can be seen that the absence of the IDPCNN feature extraction module leads to the lowest indicators, because IDPCNN extracts local and global features in text through multilayer convolution and pooling operations, which can capture the semantic information in sentences more effectively. On the other hand, the absence of a BiLSTM feature extraction module affects the structural acquisition of contextual information as well as the effective capture of both temporal information and semantic relations. From the experimental results of the S-BIDA model, it can be seen that fusing the two models together is the optimal choice.

It can be seen that the S-BIDBA model proposed in this paper achieves a reasonable improvement in the

**Table 7.** Results of ablation experiments (unit: %)

| Model | Precision | Recall | F1 |
|---|---|---|---|
| BDBA | 93.77 | 93.80 | 93.75 |
| S-BDBA | 95.75 | 95.32 | 95.34 |
| S-BBA | 93.40 | 93.54 | 93.45 |
| S-BIDA | 94.92 | 94.65 | 94.64 |
| S-BIDBA | 96.95 | 96.20 | 96.55 |

performance of sentiment classification, and merchants can understand the emotional tendency of consumers more effectively by tagging the review text with the S-BIDBA model, which facilitates the optimization and improvement of the industry chain of fresh products.

## V. CONCLUSION

In this paper, to solve the problem of inconsistency between actual emotional tendency and emotional markers in online reviews of fresh produce, we propose an emotion classification model based on S-BIDBA that enables merchants to more accurately grasp the emotional tendency of consumers, ultimately helping the fresh produce production chain better understand consumer demand, optimize product quality and service, and achieve the goal of improving the balance between supply and demand in the fresh produce market while also increasing consumer satisfaction. The experimental results show that the performance indices used in this paper's model for sentiment classification on the fresh produce review dataset are better than those of the comparison model, thus verifying the reasonableness and effectiveness of the model.

This paper still has some room for improvement and research content. Directions for follow-up research include the following:

(1) A network of entities is constructed to study the emotional correspondence between various characteristics of fresh produce, including freshness, logistics, transportation, packaging, shape, and size. This network is used to analyze the relationship between these characteristics and consumer perceptions of similar fresh produce. A knowledge graph is constructed by considering the ternary storage of "entity-relationship-entity" and "entity-attribute-attribute-value." The purpose of this is to assist merchants in comprehending customer satisfaction with fresh produce and subsequently enhance the quality and service level of fresh produce.

(2) Research in this area can be expanded to the transferability of S-BIDBA in the categorization of online reviews for other practical goods.

## CONFLICT OF INTEREST

The authors have declared that no competing interests exist.

## ACKNOWLEDGEMENTS

This research was funded by Anhui Province Education Department Research Foundation (No.2022AH050896), Anhui Natural Science foundation (No. 2108085MF209), and Anhui Province Beidou Precision Agriculture Information Engineering Laboratory Open Fund Project (No. AHBD201904).

## REFERENCES

1. H. Liu, I. Chatterjee, M. Zhou, X. S. Lu, and A. Abusorrah, "Aspect-based sentiment analysis: a survey of deep learning methods," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 6, pp. 1358-1375, 2020. https://doi.org/10.1109/TCSS.2020.3033302
2. S. Shojaeilangari, W. Y. Yau, K. Nandakumar, J. Li, and E. K. Teoh, "Robust representation and recognition of facial emotions using extreme sparse learning," *IEEE Transactions on Image Processing*, vol. 24, no. 7, pp. 2140-2152, 2015. https://doi.org/10.1109/TIP.2015.2416634
3. G. Xu, Z. Yu, H. Yao, F. Li, Y. Meng, and X. Wu, "Chinese text sentiment analysis based on extended sentiment dictionary," *IEEE Access*, vol. 7, pp. 43749-43762, 2019. https://doi.org/10.1109/ACCESS.2019.2907772
4. S. Zhang, Z. Wei, Y. Wang, and T. Liao, "Sentiment analysis of Chinese micro-blog text based on extended sentiment dictionary," *Future Generation Computer Systems*, vol. 81, pp. 395-403, 2018. https://doi.org/10.1016/j.future.2017.09.048
5. J. Wu, K. Lu, S. Su, and S. Wang, "Chinese micro-blog sentiment analysis based on multiple sentiment dictionaries and semantic rule sets," *IEEE Access*, vol. 7, pp. 183924-183939, 2019. https://doi.org/10.1109/ACCESS.2019.2960655
6. S. M. Sarsam and H. Al-Samarraie, "A lexicon-based method for detecting eye diseases on microblogs," *Applied Artificial Intelligence*, vol. 36, no. 1, article no. 1993003, 2022. https://doi.org/10.1080/08839514.2021.1993003
7. T. Baltrusaitis, C. Ahuja, and L. P. Morency, "Multimodal machine learning: a survey and taxonomy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 2, pp. 423-443, 2019. https://doi.org/10.1109/TPAMI.2018.2798607
8. K. X. Han, W. Chien, C. C. Chiu, and Y. T. Cheng, "Application of support vector machine (SVM) in the sentiment analysis of twitter dataset," *Applied Sciences*, vol. 10, no. 3, article no. 1125, 2020. https://doi.org/10.3390/app10031125
9. D. Wang, A. J. Adam, Y. Xiao, and Y. Gao, "Effective application of Naive Bayesian classifier for personal online learning networks," *International Journal of Wireless Information Networks*, vol. 26, pp. 174-182, 2019. https://doi.org/10.1007/s10776-019-00436-9
10. A. Yudhana, R. Umar, and S. Saputra, "Fish freshness identification using machine learning: performance comparison of k-NN and Naïve Bayes classifier," *Journal of Computing Science and Engineering*, vol. 16, no. 3, pp. 153-164, 2022. http://dx.doi.org/10.5626/JCSE.2022.16.3.153
11. N. M. H. Alamri, M. Packianather, and S. Bigot, "Deep learning: parameter optimization using proposed novel hybrid bees Bayesian convolutional neural network," *Applied Artificial Intelligence*, vol. 36, no. 1, article no. 2031815, 2022. https://doi.org/10.1080/08839514.2022.2031815

12. S. Ma, H. Cheng, and H. Lee, “A Practical approach to indoor path loss modeling based on deep learning,” *Journal of Computing Science and Engineering*, vol. 15, no. 2, pp. 84-95, 2021. http://dx.doi.org/10.5626/JCSE.2021.15.2.84
13. W. Gao, C. Shao, and Q. Gao, “An optimization algorithm with novel flexible grid: applications to parameter decision in LS-SVM,” *Journal of Computing Science and Engineering*, vol. 9, no. 2, pp. 39-50, 2015. https://doi.org/10.5626/JCSE.2015.9.2.39
14. X. Zhou, F. Wang, M. Liu, and D. Wang, “Prediction of mechanical properties of hot pressed hybrid material plates using PSO-SVR model,” *Journal of Northeast Agricultural University (English Edition)*, vol. 49, no. 2, pp. 87-96, 2018. https://doi.org/10.3969/j.issn.1005-9369.2018.02.010
15. T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient estimation of word representations in vector space,” 2013 [Online]. Available: https://arxiv.org/abs/1301.3781.
16. J. Pennington, R. Socher, and C. D. Manning, “GloVe: global vectors for word representation” in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, 2014, pp. 1532-1543. https://aclanthology.org/D14-1162
17. M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer, “Deep contextualized word representations,” in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, New Orleans, LA, USA, 2018, pp. 2227-2237.
18. J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, “BERT: pre-training of deep bidirectional transformers for language understanding,” 2018 [Online]. Available: https://arxiv.org/abs/1810.04805.
19. R. Xin, J. Zhang, and Y. Shao, “Complex network classification with convolutional neural network,” *Tsinghua Science and technology*, vol. 25, no. 4, pp. 447-457, 2020. https://doi.org/10.26599/TST.2019.9010055
20. Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, “A survey of convolutional neural networks: analysis, applications, and prospects,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 12, pp. 6999-7019, 2022. https://doi.org/10.1109/TNNLS.2021.3084827
21. O. Habimana, Y. Li, R. Li, X. Gu, and G. Yu, “Sentiment analysis using deep learning approaches: an overview,” *Science China Information Sciences*, vol. 63, article no. 111102, 2020. https://doi.org/10.1007/s11432-018-9941-6
22. J. Du, C. M. Vong, and C. P. Chen, “Novel efficient RNN and LSTM-like architectures: Recurrent and gated broad learning systems and their applications for text classification,” *IEEE Transactions on Cybernetics*, vol. 51, no. 3, pp. 1586-1597, 2021. https://doi.org/10.1109/TCYB.2020.2969705
23. S. Hochreiter and S. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997. https://doi.org/10.1162/neco.1997.9.8.1735
24. N. M. Vural, S. Ergut, and S. S. Kozat, “An efficient and effective second-order training algorithm for LSTM-based adaptive learning,” *IEEE Transactions on Signal Processing*, vol. 69, pp. 2541-2554, 2021. https://doi.org/10.1109/TSP.2021.3071566
25. X. Yan, W. Weihan, and M. Chang, “Research on financial assets transaction prediction model based on LSTM neural network,” *Neural Computing and Applications*, vol. 33, pp. 257-270, 2021. https://doi.org/10.1007/s00521-020-04992-7
26. M. Qorich and R. El Ouazzani, “Text sentiment classification of Amazon reviews using word embeddings and convolutional neural networks,” *The Journal of Supercomputing*, vol. 79, pp. 11029-11054, 2023. https://doi.org/10.1007/s11227-023-05094-6
27. L. Yao, “Sentiment analysis based on CNN-LSTM hotel reviews,” *Journal of Physics: Conference Series*, vol. 2330, no. 1, article no. 012018, 2022. https://doi.org/10.1088/1742-6596/2330/1/012018
28. M. Schuster and K. K. Paliwal, “Bidirectional recurrent neural networks,” *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2673-2681, 1997. https://doi.org/10.1109/78.650093
29. X. Guo, N. Zhao, and S. Cui, “Sentiment analysis of consumer online reviews based on CNN-BiLSTM,” *Systems Engineering – Theory and Practice*, vol. 40, no. 3, pp. 653-663, 2020. https://doi.org/10.12011/1000-6788-2018-1890-11
30. R. Johnson and T. Zhang, “Deep pyramid convolutional neural networks for text categorization,” in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Vancouver, Canada, 2017, pp. 562-570.
31. D. Bahdanau, K. Cho, and Y. Bengio, “Neural machine translation by jointly learning to align and translate,” 2014 [Online]. Available: https://arxiv.org/abs/1409.0473v1.
32. G. Chang, S. Hu, and H. Huang, “Two-channel hierarchical attention mechanism model for short text classification,” *The Journal of Supercomputing*, vol. 79, pp. 6991-7013, 2023. https://doi.org/10.1007/s11227-022-04950-1
33. M. Zhang, H. Zhou, and J. Cao, “Research on dual BERT directed sentiment text classification based on attention mechanism,” *CAAI Transactions on Intelligent Systems*, vol. 17, no. 6, pp. 1220-1227, 2022. https://doi.org/10.11992/tis.202112038
34. Y. J. Su, W. C. Hu, J. H. Jiang, and R. Y. Su, “A novel LMAEB-CNN model for Chinese microblog sentiment analysis,” *The Journal of Supercomputing*, vol. 76, pp. 9127-9141, 2020. https://doi.org/10.1007/s11227-020-03198-x
35. R. Mihalcea and P. Tarau, “TextRank: bringing order into text,” in *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, Barcelona, Spain, 2004, pp. 404-411.
36. K. K. Mohbey, G. Meena, S. Kumar, and K. Lokesh, “A CNN-LSTM-based hybrid deep learning approach for sentiment analysis on Monkeypox tweets,” *New Generation Computing*, vol. 42, pp. 89-107, 2024. https://doi.org/10.1007/s00354-023-00227-0
37. H. Murfi, T. Gowandi, G. Ardaneswari, and S. Nurrohmah, “BERT-based combination of convolutional and recurrent neural network for Indonesian sentiment analysis,” *Applied Soft Computing*, vol. 151, article no. 111112, 2024. https://doi.org/10.1016/j.asoc.2023.111112

**Yindong Dong** https://orcid.org/0000-0003-0390-4761

Yindong Dong received Ph.D. degree in computer application technology from Hefei University of Technology, China in 2021. He works for the School of Information and Artificial Intelligence, Anhui Agricultural University since June 2021. His research interests are affective computing, natural language processing, and intelligent recognition based on physiological signals.

**Shundong Shao** https://orcid.org/0009-0003-6629-9412

Shundong Shao received his bachelor's degree in management from Hefei University of Economics in 2022. He is currently pursuing a master's degree at Anhui Agricultural University. His research focuses on recommendation systems and natural language processing.

**Guohua Fan** https://orcid.org/0009-0003-0999-7978

Guohua Fan received Master's degree in computer science from the University of Hefei University of Technology, China in 2007. He joined in the School of Information and Artificial Intelligence, Anhui Agricultural University of China in September 2007. His research interests are in game design, virtual reality, and image recognition.

**Deli Chen** https://orcid.org/0009-0001-0937-1671

Deli Chen received his bachelor's degree in computer science and technology from Anhui Agriculture University in 2023. He is currently pursuing a master's degree at Anhui Agricultural University. His research focuses on affective computing and natural language processing.

**Yu Zhang** https://orcid.org/0009-0006-4833-4861

Yu Zhang received his bachelor's degree in management from BoZhou University in 2022. He is currently pursuing a master's degree at Anhui Agricultural University. His research focuses on recommendation systems and natural language processing.

**Xin Luo** https://orcid.org/0009-0007-2818-2060

Xin Luo received his bachelor's degree in management from Suzhou University in 2021. He is currently pursuing a master's degree at Anhui Agricultural University. His research focuses on recommendation systems and natural language processing.