Sentiment Analysis of Cross-Border E-Commerce Product Reviews Based on Interactive Attention Mechanism

Xu-Dong Cheng[1,*], Wei Li[2], Li-Yan Sun[3]

[1]ChengDu Industry & Trade College, Chengdu 611730, P. R. China
chengxudong0319@163.com

[2]SiChuan Normal University, Chengdu 610066, P. R. China
li_wei233@sina.com

[3]Northwestern University, Chiang Mai 20260, Thailand
ae6572@163.com

*Corresponding author: Xu-Dong Cheng

ABSTRACT. Traditional sentiment analysis methods for e-commerce product reviews mainly focus on within a single domain, making it difficult to effectively deal with cross-domain sentiment analysis tasks, and it is easy to introduce noise when dealing with complex target words, which affects the accuracy. In this paper, we propose a cross-border e-commerce product review sentiment analysis model based on interactive attention mechanism and bi-directional long short-term memory network (BiLSTM). Firstly, through the bilateral interactive attention mechanism, the model is able to better capture the complex dependencies between target words and their contexts, thus improving the generalisation ability in different domains. Second, the improved target word representation effectively solves the problems of target word information dilution and noise introduction. By introducing BiLSTM and internal attention mechanism, the model is able to extract features accurately and performs well in sentiment classification tasks. Experimental results show that the proposed BA-BiLSTM model exhibits high accuracy in several cross-domain sentiment classification tasks, with an average classification accuracy of 83.90%, which is a significant improvement over existing method. This study provides an effective method for cross-border e-commerce sentiment analysis, which significantly improves the accuracy and robustness of sentiment classification, and solves the shortcomings of traditional methods in dealing with cross-domain sentiment analysis and complex target words.
Keywords: Interactive attention mechanism; BiLSTM; Cross-border e-commerce; Sentiment analysis; Target word representation

1. **Introduction.** In the context of globalisation, cross-border e-commerce has become an important part of modern business. In the context of swift expansion of international e-commerce ventures, the number of consumer reviews on cross-border products has increased significantly [1, 2, 3]. These reviews contain rich user sentiment information, which is important for merchants to understand market demand and user satisfaction. However, cross-domain language differences and diversity of expressions make sentiment analysis of these reviews more complex and challenging. Sentiment analysis techniques that are currently utilized typically concentrate on analyzing one particular domain, making it difficult to effectively handle cross-domain sentiment analysis tasks [4, 5]. In addition,

target words in comment texts often contain multiple words, and traditional sentiment analysis methods are prone to information dilution or noise introduction when dealing with these complex target words, affecting the accuracy and robustness of the model.

In the latest era, the swift progress of profound learning algorithms has been a catalyst for change, neural network-based sentiment analysis methods have been widely used. In particular, Long Short-Term Memory (LSTM) networks [6] and Bidirectional Long Short-Term Memory (BiLSTM) networks [7, 8] have shown superior performance in capturing long-distance dependencies in text sequences. Meanwhile, the incorporation of the attention mechanism enables the model to concentrate more effectively on the critical elements for sentiment analysis. However, the traditional unidirectional attention mechanism tends to consider only one direction of dependencies when dealing with the relationship in relation to the focal term and its surrounding text, making it difficult to fully utilise the contextual information.

The aim of this paper is to investigate how to enhance the accuracy of cross-domain cross-border e-commerce product review sentiment analysis using BiLSTM in combination with an improved bilateral interactive attention mechanism. By introducing the bilateral interaction attention mechanism, the complex dependencies between target words and their contexts can be better captured, thus improving the generalisation ability of the model across different domains. Our study not only provides an effective method for cross-border e-commerce sentiment analysis, but also provides a reference for research in related fields.

1.1. **Related work.** E-commerce platforms have generated a large amount of user review data, which can not only reflect the real feelings of users about the products [9], but also provide valuable market feedback for merchants. Therefore, how to effectively analyse and make use of these review data has become a hot research topic [10]. Sentiment analysis, an approach in the realm of natural language engineering, seeks to interpret users' attitudes and opinions by analysing the sentiment tendencies in text data. Many researchers have used various methods to perform sentiment analysis on e-commerce product reviews. Almestekawy and Abdulsalam [11] used a bag-of-words model-based plain Bayesian classification method for sentiment classification of Amazon product reviews, and the study showed that the method has high accuracy when dealing with large-scale data. Torres-Valencia et al. [12] proposed a sentiment analysis method based on Support Vector Machine (SVM), which improved the classification performance of the model by optimising the kernel function. Li et al. [13] used Convolutional Neural Network (CNN) to perform sentiment analysis on reviews of e-commerce platforms, and it was found that CNN has obvious advantages in capturing local features of text. Wang et al. [14] used BiLSTM combined with the attention mechanism to modify the effect of sentiment analysis, and the results showed that this method performed well in capturing long-distance dependencies in the reviews. In addition, Wang and Tong [15] explored the Transformer-based sentiment analysis model, which significantly improves the accuracy and efficiency of sentiment classification through the self-attention mechanism.

Despite the many advances in the theory and practice of e-commerce product review sentiment analysis, it still faces some challenges and problems. Current research focuses on the following directions: first, how to improve the accuracy and robustness of sentiment analysis models is still an important topic. Yan et al. [16] improved the accuracy of sentiment analysis by fusing multimodal data (e.g., text, image, and audio), and found that the multimodal approach has a better performance in dealing with complex sentiment expressions. However, the high cost of acquiring and processing multimodal data limits its application. Secondly, the issue of interpretability of sentiment analysis models

has also received much attention. Lieto et al. [17] proposed an interpretable sentiment analysis model on the attention mechanism, which explains the decision-making process by visualising the attention weights, and although this method improves the transparency of the model to some extent, its complexity increases the computational overhead of the model. In addition, the problem of data imbalance is one of the common challenges in sentiment analysis. Ma et al. [18] employed a sample balancing technique based on Generative Adversarial Networks (GANs) to effectively mitigate the impact of data imbalance on model performance. However, the training process of GAN is unstable and can easily lead to model crash. Finally, cross-domain sentiment analysis is also a popular research direction. Luna-Jiménez et al. [19] investigated how to migrate sentiment analysis models between different e-commerce platforms and different languages, and the results showed that transfer learning has potential in cross-domain sentiment analysis, but the problems of model adaptability and generalisation ability need to be addressed. In summary, although sentiment analysis techniques show a broad application prospect in e-commerce product reviews, further in-depth research is needed to improve model performance, enhance interpretability, deal with data imbalance, and cross-domain applications.

1.2. **Motivation and contribution.** Existing approaches to sentiment analysis usually focus on a single domain, making it difficult to effectively handle cross-domain sentiment analysis tasks. Comments from different domains differ significantly in their linguistic expressions and sentiment tendencies, making models often perform poorly on new domains. Target words in review texts often contain multiple words, and traditional sentiment analysis methods are prone to information dilution or noise introduction when dealing with these complex target words. To address the aforementioned issues, we propose a BiLSTM model based on interactive attention mechanism to improve the effectiveness of sentiment analysis of cross-border e-commerce product reviews. The main innovations and contributions of this work include:

(1) To address the challenging problem of cross-domain sentiment analysis, the bilateral interactive attention mechanism proposed in this paper is able to better capture the complex dependencies between target words and their contexts, thus improving the generalisation ability of the model across different domains. This improvement is particularly significant in cross-domain sentiment analysis.

(2) Aiming at the complexity problem of target words, the improved target word representation proposed in this paper converts target words composed of multiple words into one word vector representation by weighted average method, which avoids the problems of information dilution and noise introduction, and thus improves the accuracy of sentiment classification.

(3) For the problem of capturing contextual dependencies, the bilateral interaction attention mechanism introduced in this paper enhances the ability to capture sentiment information by simultaneously considering the interaction between the target word and its left and right contextual words, so that the model can understand the semantic information of the target word more comprehensively.

2. **Related technical studies.**

2.1. **Overview of sentiment analysis.** Sentiment analysis is an important branch of natural language processing (NLP) [20, 21] that aims to identify and extract sentiment information by analysing textual data. Based on the level of detail in the subject matter being analyzed, sentiment analysis is categorized into analyses at the word, sentence, chapter, and aspect levels.

(1) **A rule-based approach to sentiment analysis.**
Rule-based methods are one of the earliest approaches in the field of sentiment analysis [22]. These methods rely on predefined rules and lexicons to determine the sentiment tendency of a text. These rules are usually hand-written by domain experts and contain a large number of sentiment words and their corresponding sentiment polarity. Despite their simplicity and ease of implementation, rule-based methods rely on the completeness of the lexicon and the accuracy of the rules, and have a weak generalisation ability that makes it difficult to cope with the diversity and complexity of languages.

(2) **Machine Learning Based Sentiment Analysis Methods.**
Machine learning-based methods are widely used in sentiment analysis [23]. Common machine learning algorithms include plain Bayes, maximum entropy model and support vector machine. These methods extract text features through feature engineering and train classifiers for sentiment classification. Compared with rule-based methods, machine learning methods have better generalisation ability, but their performance depends on the selection of features and the quality of training data.

(3) **Sentiment analysis methods based on deep learning.**
Rule-based methods are among the earliest approaches in the field of sentiment analysis [24, 25]. These approaches depend on established guidelines and dictionaries to ascertain the emotional inclination of the text. These rules are usually hand-written by domain experts and contain a large number of sentiment words and their corresponding sentiment polarities. Despite their simplicity and ease of implementation, rule-based methods rely on the completeness of the lexicon and the accuracy of the rules, and have a weak generalisation ability, making it difficult to cope with the diversity and complexity of languages.

CNN extracts the local semantic information of a sentence through convolutional operations, and is the earliest deep network model that has been used to solve the problem of sentiment classification. By extracting features layer by layer from the lexical level to the sentence level, the final sentiment classification is performed.

The Recurrent Neural Networks (RNN) and its variant LSTM can effectively capture long-range dependencies in sequential data and are powerful tools for processing textual sequential data. In the field of sentiment analysis, RNN and LSTM have demonstrated superior performance.

The BiLSTM improves the contextual understanding of the model by taking into account both forward and backward information of the sequence, and excels in sentiment analysis tasks.

Sentiment analysis methods based on deep learning not only improve classification accuracy, but also handle complex linguistic structures and diverse textual content.

2.2. **Word vectors.** Word vectors are a method of representing words in terms of vectors, which is a fundamental technique in NLP. Word vector technology enables computers to understand and process semantic information in natural language text by mapping words into a continuous, high-dimensional vector space. The methods for generating word vectors [26] mainly include those based on One-Hot Encoding and those based on distributed representation (e.g., Word2Vec, GloVe, BERT, etc.)

One-Hot Encoding is one of the simplest word representations, which represents each word as a high-dimensional vector, where the corresponding position of the word is 1 and the rest of the positions are 0. For example, for a vocabulary list containing 5 different words {"I", "like", "eat", "apple", "fruit"}, its unique heat representation is shown in Table 1:

Table 1. Example of Unique Heat Representation

Term	Unique heat vector
I	[1, 0, 0, 0, 0, 0]
like	[0, 1, 0, 0, 0]
eat	[0, 0, 1, 0, 0]
apple	[0, 0, 0, 0, 1, 0]
fruit	[0, 0, 0, 0, 0, 1]

Despite its simplicity and intuition, the solo thermal representation has obvious drawbacks: 1) Dimensionality disaster. When the vocabulary is large, the dimension of the vector increases, resulting in huge consumption of computational and storage resources. 2) Semantic gap. Solitary heat vectors cannot capture the semantic relationship between words, for example, "apple" and "fruit" are semantically related, but they are completely independent in the solo heat representation.

Distributed representations overcome the drawbacks of solo-hot representations by mapping words into a low-dimensional continuous vector space. Commonly used distributed representations include Word2Vec, GloVe, and BERT. Word2Vec is a commonly used word vector generation method, which is trained to obtain word vectors by two model architectures (CBOW and Skip-Gram). CBOW infers target words through contextual vocabulary. Skip-Gram uses the target word to infer its context. CBOW model illustrates the training process as follows. The training process of CBOW model is as follows:

- Input layer: unique heat vectors for context words.
- Hidden layer: mapping the unique heat vector to a low-dimensional space through a weight matrix.
- Output layer: calculating the probability patterns for the central term.

The CBOW model works as shown below:

$$h = \frac{1}{C}\mathbf{W}^T(x_1 + x_2 + \ldots + x_C) \tag{1}$$

where $\mathbf{W}$ is the weight matrix; x_i is the unique heat vector of context words; and C is the number of context words.

GloVe is a method for generating word vectors based on a global co-occurrence matrix, which learns word vectors by counting the frequency of co-occurrence of words in the whole corpus. The basic idea of GloVe is that, for the words i and j, the probability of their co-occurrence in the corpus can be represented by the inner product of their word vectors.

$$P_{ij} = \frac{X_{ij}}{X_i} \tag{2}$$

where X_{ij} is the number of times the words i and j co-occur; X_i is the total number of times the word i occurs.

The definition of the objective function is shown below:

$$J = \sum_{i,j} f(X_{ij}) \left(w_i^T w_j + b_i + b_j - \log(X_{ij}) \right)^2 \tag{3}$$

where w_i and w_j are word vectors for words i and j; b_i and b_j are bias terms; f is a weighting function.

BERT is a bi-directional Transformer-based language model that achieves a comprehensive understanding of contextual information through pre-training and fine-tuning. BERT employs a self-attention mechanism that is capable of capturing long-range dependencies. and is trained through two stages:

(1) The basic idea of Masked Language Model (MLM) is to randomly replace some of the words with [MASK] tokens and the model predicts the masked words by context.

$$\text{Loss}_{MLM} = -\sum_t \log P(x_t|x_{\backslash t}) \qquad (4)$$

where x_t is the masked word; $x_{\backslash t}$ is the context.

(2) Next Sentence Prediction (NSP) learns relationships between sentences by predicting whether two sentences are connected.

$$\text{Loss}_{NSP} = -\log P(y|s_1, s_2) \qquad (5)$$

where y is the label for whether the sentences are connected; s_1 and s_2 are the sentences.

3. Attention mechanisms.

3.1. Fundamentals.

Attention Mechanism (AM) is a technique that has been widely used in the field of deep learning in recent years, initially used in machine translation tasks. Its core idea is to pick out the primary parts for the current task by assigning different weights to better capture contextual dependencies.

The AM works by calculating the weights of the parts of the structure the input sequence to guide the model's attention towards the critical data points. In the realm of NLP, the attention mechanism typically facilitates the alignment of a focal word with its surrounding textual environment. The main process is as follows:

(1) Input Representation: Suppose the input sequence is $\{x_1, x_2, \ldots, x_n\}$, which is represented as a vector $\{h_1, h_2, \ldots, h_n\}$ through the embedding layer.

(2) Weight calculation: Calculate the attentional weight of each input vector h_i on the target word h_t. The weight e_{it} is calculated as follow:

$$e_{it} = h_i^{\mathrm{T}} W_a h_t \qquad (6)$$

where W_a is a trainable parameter matrix.

(3) Weight normalisation: use Softmax function to normalise the weights to get the attention weights α_{it}:

$$\alpha_{it} = \frac{\exp(e_{it})}{\sum_{j=1}^n \exp(e_{jt})} \qquad (7)$$

(4) Weighted Summing: weight the input vectors according to the attention weights to obtain the context vector c_t:

$$c_t = \sum_{i=1}^n \alpha_{it} h_i \qquad (8)$$

Currently, the mainstream text analysis tasks are modelled using a sequence-to-sequence model. A commonly used sequence-to-sequence model is the "encode-decode" structure, shown in Figure 1.

3.2. Application examples.

In cross-domain cross-border e-commerce product review sentiment analysis, the attention mechanism can help the model better capture the parts of the review that are crucial to the sentiment polarity judgement, so as to improve the accuracy of sentiment classification. Here is a simple example: suppose there is a review "The battery life is amazing but the camera quality is poor", the model needs to judge the sentiment polarity of "battery life" and "camera quality". Through the attention mechanism, the model can pay more attention to the modifiers "amazing" and "poor" to accurately determine their sentiment tendencies.
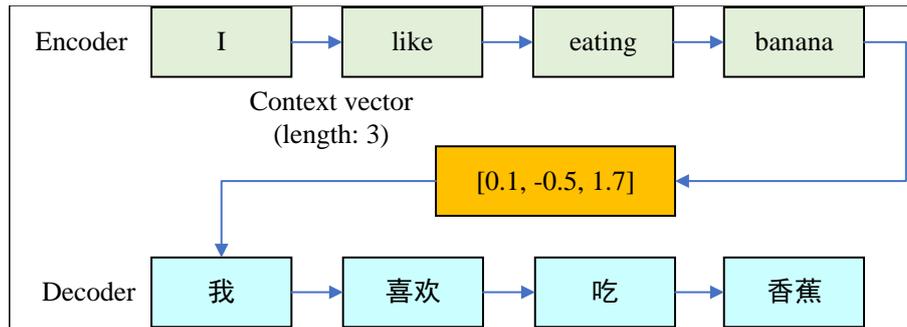
Figure 1. Encoder-decoder structure. The left part is the encoder, which processes the sequence input and compresses the information into a fixed-length context vector. The right part is the decoder, which is initialised by the context vector and produces a transcoded output at each moment.

In the model of this paper, the input comments will be processed by BiLSTM combined with the attention mechanism, and the performance of the model will be further improved by the bilateral interaction attention mechanism.

4. Model of this paper.

4.1. **Overall model architecture.** The overall architecture of the model proposed in this paper combines BiLSTM and the improved bilateral interactive attention mechanism to represent the input text through preprocessing and word vector embedding, capture the bidirectional dependency among the focal word and its surrounding textual elements using BiLSTM, and model the target word and the context at a fine-grained level through the internal AM and the bilateral interactive AM, respectively. Finally, the sentiment analysis of cross-domain cross-border e-commerce product reviews is realised by classifying the sentiment polarity through the fully connected layer and Softmax function. The overall structure of the proposed Bilateral Interactive Attention Mechanisms BiLSTM (BA-BiLSTM) model is shown in Figure 2.

4.2. **Input construction based on improved target word representation.** In traditional sentiment analysis models, target words may consist of multiple words, such as "battery life" or "delivery speed". These target words themselves do not directly contain sentiment information, which is more determined by the context words around them. Inputting multiple words of these target words directly into the model may introduce noise and lose important information, thus adversely affecting the model performance. Specific issues include:

(1) Dilution of target word information: when the target word consists of multiple words, simply averaging its vectors leads to semantic dilution, making it difficult for the model to capture critical sentiment information.

(2) Noise introduction: irrelevant target word components may introduce noise that interferes with the model's extraction of important sentiment information.

(3) Lack of context dependency: existing methods fail to take full advantage of the dependency between target words and their contexts, resulting in lower accuracy of sentiment classification.

To tackle the previously mentioned challenges, an improved target word representation is proposed in this paper. The target word vector representation is performed first. For a sentence $S = \{c_1, \ldots, c_{\tau+1}, \ldots, c_{\tau+m}, \ldots, c_n\}$ containing n words, where the target word consists of m words with subscripts starting from $\tau+1$, such as "battery life". Each word c_i
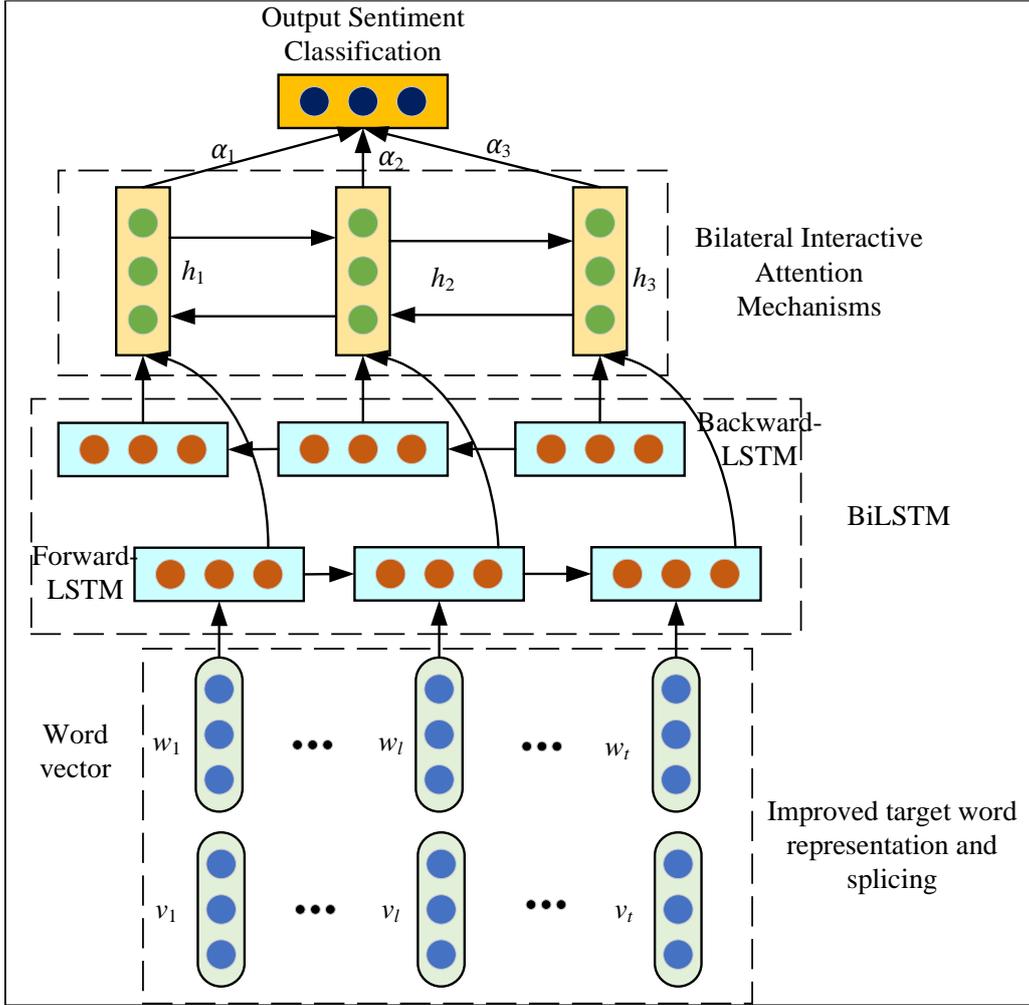
Figure 2. The overall structure of the proposed BA-BiLSTM model.

is represented as a real-valued vector $w \in \mathbb{R}^{d_{emb} \times 1}$ characterized by its low dimensionality and continuity, where d_{emb} is the dimension of the word vector.

Then, a weighted average target word representation is performed. Since the target word may contain multiple words, this paper designs a weighted average method to convert the target word consisting of multiple words into a word vector v_t representation.

$$v_t = \frac{1}{m} \sum_{i=\tau+1}^{\tau+m} w_i \tag{9}$$

where w_i denotes the word vector for each word in the target word.

To optimize the utilization of target word data, the model combines the target word vector v_t with each vector representing a contextual word, feeding these into the model as inputs. At this point, the hidden state output by each context word through the BiLSTM unit will contain the target word information, so that the interdependence between the context word and the target word can be modelled in the next step of calculating the attention weights.

After the BiLSTM layer, the attention mechanism is introduced, which enables the model to focus on those context words that contribute significantly to sentiment classification by calculating the importance weights of the context words to the target words.

With the above improved methods, the model is able to represent the semantic information of the target word more accurately and effectively capture the key context words related to the semantic orientation of the focal term, thus improving the accuracy of sentiment classification.

4.3. **BiLSTM-based feature extraction.** In text sentiment analysis, accurate feature extraction is crucial to the performance of the model. In this paper, BiLSTM is used to capture the dependencies between target words and context words. The bidirectional structure of BiLSTM can effectively utilise contextual information to provide richer feature representation for sentiment classification [27].

Some existing methods, such as unidirectional LSTM, can only capture contextual information in a single direction, which may lose some of the important contextual relationships when dealing with complex sentence structures [28]. In addition, simple target word representations (e.g., averaging word vectors of multiple words) are prone to lose information or introduce noise. So, BiLSTM is adopted in this paper for feature extraction, and the specific process is as follows:

The input sentence $S = \{c_1, \ldots, c_n\}$ is converted to a word vector representation, where each word c_i is represented as a vector $w_i \in \mathbb{R}^{d_{emb} \times 1}$, where d_{emb} is the the dimension of the word vector. The BiLSTM layer uses two independent LSTM networks, one interpreting the input sequence sequentially from start to finish (forward LSTM) and the counterpart interpreting from finish to start (backward LSTM). For each time step t, the hidden state of the forward LSTM is h_t and the hidden state of the backward LSTM is h_t. The final hidden state is denoted as follow:

$$h_t = [h_t^{\rightarrow}; h_t^{\leftarrow}] \tag{10}$$

This bi-directional processing captures complete information about the context before and after the target word. Then, target word context modelling is performed. The target word is divided into three parts: the target word itself, the context word in front of the target word, and the context word behind the target word. The contexts before and after the target word are modelled by forward LSTM and backward LSTM, respectively.

$$h_t^{\rightarrow} = \text{LSTM}_{\text{forward}}(w_t, h_{t-1}^{\rightarrow}) \tag{11}$$

$$h_t^{\leftarrow} = \text{LSTM}_{\text{backward}}(w_t, h_{t+1}^{\leftarrow}) \tag{12}$$

The final hidden state h_t contains all the context information before and after the target word. The hidden state h_t of a target word is concatenated with the hidden states of the contexts before and after it to form the final feature representation [29]. For a target word T, its feature representation is

$$H_T = [h_{\tau+1}; h_{\tau+2}; \ldots; h_{\tau+m}] \tag{13}$$

where $h_{\tau+i}$ denotes the hidden state of the i-th word in the target word.

The BiLSTM model not only effectively utilises the contextual information of the target word, but also reduces the noise introduced by the fact that the target word consists of multiple words. The feature representation obtained after bi-directional modelling of the input sentence by BiLSTM captures the complex interrelationships linking the focal word to its surrounding text, providing high-quality features for subsequent sentiment classification.

4.4. **Internal Attention Mechanisms.** It can be seen from the above analysis that the traditional attention mechanism approach usually uses averaging of the word vectors of individual words in the target word when dealing with target words containing multiple words. However, this approach suffers from semantic dilution, information loss and noise

introduction. Simple averaging leads to dilution of important information and makes it difficult for the model to accurately capture the core meaning of the target word. Different words have different importance in the target word, and simple averaging ignores these differences, leading to information loss. Certain words in the target word may introduce noise, affecting the model's judgement of sentiment polarity.

In order to better capture the semantic information of target words and accurately identify their sentiment tendencies in different contexts, this paper designs an internal attention mechanism. The mechanism avoids the semantic dilution problem caused by simple averaging by calculating the internal attention weight of the target word, thus improving the accuracy of sentiment analysis. Assuming that the target word consists of m words, denoted as $\{w_{\tau+1}, w_{\tau+2}, \ldots, w_{\tau+m}\}$, each word w_i is embedded as a vector $v_i \in \mathbb{R}^{d_{emb} \times 1}$.

For each word w_i, calculate its activation value e_i by the activation function:

$$e_i = \tanh(W_{att} \cdot v_i + b_{att}) \tag{14}$$

where W_{att} and b_{att} are trainable parameters.

The activation values are normalised using the Softmax function to obtain the attentional weights for each word α_i:

$$\alpha_i = \frac{\exp(e_i)}{\sum_{j=\tau+1}^{\tau+m} \exp(e_j)} \tag{15}$$

The word vectors of the target word are weighted and summed according to the attention weights to obtain the representation vector of the target word v_t:

$$v_t = \sum_{i=\tau+1}^{\tau+m} \alpha_i v_i \tag{16}$$

Through the above process, the internal attention mechanism is able to effectively capture the key parts of the target word, thus generating a more accurate representation of the target word.

4.5. **Bilateral interactive attention mechanisms.** Traditional unidirectional attention mechanisms tend to consider only one direction of dependency when capturing the relationship between the target word and the context word, which makes it difficult for the model to make full use of the context information. This paper proposes a bilateral interactive attention mechanism, which enhances the ability to capture sentiment information by simultaneously considering the interaction among the focal term and its associated textual elements on the left and right sides.

The bilateral interaction attention mechanism consists of two main steps: the interaction amongst the target and the left context, and the interaction amongst the target and the right context. By calculating the attention weights of these two parts separately, the model is able to understand the semantic information of the target more comprehensively. The target word is set to be represented as r_t and the hidden state of the left context is $\{h_1, h_2, \ldots, h_L\}$. Firstly, the similarity score activation function $g(h_i, r_t)$ between the left context and the target is calculated by the

$$g(h_i, r_t) = \tanh(W_l \cdot h_i + U_l \cdot r_t + b_l) \tag{17}$$

where W_l and U_l are trainable weight matrices and b_l is a bias vector.

The similarity scores are normalised using the Softmax function to obtain the attention weights α_i:

$$\alpha_i = \frac{\exp(g(h_i, r_t))}{\sum_{j=1}^{L} \exp(g(h_j, r_t))} \tag{18}$$

Weighted summation of the hidden states of the left context words using the attention weights yields the left context representation c_l:

$$c_l = \sum_{i=1}^{L} \alpha_i h_i \tag{19}$$

Set the hidden state of the right-hand context word to $\{h_{L+1}, h_{L+2}, \ldots, h_n\}$. Calculate the similarity score between the right-hand context word and the target word

$$g(h_i, r_t) = \tanh(W_r \cdot h_i + U_r \cdot r_t + b_r) \tag{20}$$

where W_r and U_r are trainable weight matrices and b_r is a bias vector.

The similarity scores are normalised using the Softmax function to obtain the right-hand side attention weights β_i:

$$\beta_i = \frac{\exp(g(h_i, r_t))}{\sum_{j=L+1}^{n} \exp(g(h_j, r_t))} \tag{21}$$

Weighted summation of the hidden states of the right-hand context words using the attention weights yields the right-hand context representation c_r:

$$c_r = \sum_{i=L+1}^{n} \beta_i h_i \tag{22}$$

The left contextual representation c_l, the target word representation r_t and the right contextual representation c_r are concatenated to obtain the final representation c:

$$c = [c_l; r_t; c_r] \tag{23}$$

5. Experimental results and analyses.

5.1. **Experimental dataset.** The dataset used to conduct the simulation experiments in this paper is the cross-border e-commerce sentiment dataset collected by Blitzer et al. from the University of Pennsylvania [30], which specifically includes four domains: Books (B), DVDs (D), Electronics (E), and Kitchen (K). Among them, each domain contains 1000 positive and negative reviews each, in addition to a large amount of unlabelled data. In order to verify the effectiveness of cross-domain learning, ten sentiment analysis tasks across different domains have been established with the help of this dataset: B→D, B→E, B→K, D→B, D→E, D→K, E→B, E→D, E→K, and K→B. Where letters before the symbols are the source domains. letters following the symbols are the target domains. Separately, the positive and negative comments of each domain are divided into training and validation sets by five-fold cross-validation. The statistical information of the sentiment dataset is shown in Table 2.

Table 2. The statistical information of the sentiment dataset

Realm	Training set	Test set	Unlabelled data sets	Percentage of negative comments/%
Books	1500	500	6000	50
DVD	1500	500	34741	50
Electronics	1500	400	13153	50
Kitchen	1500	500	16785	50

During pre-processing, the comment text was cleaned, including operations such as modifying abbreviations and irregular punctuation to ensure a more standardised and tidy text.

5.2. **Experimental parameter settings.** In order to ensure the validity and stability of the model, the hyperparameters are optimally set, and the specific parameters are shown in Table 3.

Table 3. Experimental parameter settings

Parameter description	Parameter value
Dimension of a word vector	300
Batch size	64
Hidden layer size	300
Dropout ratio	0.5
Optimisation function	Adam
Learning rate	0.001
Number of selected generations trained	20

The experiments are conducted on a PC configured with Intel i7-8750H@2.20GHz CPU, 16GB, Windows 10 operating system. To evaluate the model performance, this paper mainly adopts the Accuracy as the evaluation index. The development environment is Python 3.5 and TensorFlow 1.10.0. In addition, in order to prevent model overfitting, this paper adopts Dropout strategy and introduces L2 regularisation mechanism. In backpropagation, the AdaGrad optimization method is used to adjust the model parameters to minimize the value of the loss function. The loss function in this paper is the cross-entropy.

$$\text{Loss} = -\sum_{i=1}^{C} y_i \log \hat{y}_i + \lambda \|\theta\|^2 \tag{24}$$

where y_i is the actual category, $\hat{y}_i$ is the predicted category, λ is the regular term coefficient (to prevent overfitting), C is the number of categories (2 in this experiment), and θ denotes all the trainable parameters of the model (including all the weight matrices and bias vectors).

5.3. **Attention visualisation analysis.** The feature extraction work in this paper is visualised through a simple example. Figure 3 shows a visual image of the attention mechanism of the model proposed in this paper.

As can be seen from the weight visualisation diagram above, the proposed bilateral interaction attention mechanism serves to capture the words that are more strongly related to the emotional polarity of the target word, assign it a larger weight, and pay more attention to the semantic information of the key contextual words as a way to improve the accuracy of the judgement of the emotional orientation of the target. The internal attention mechanism and bilateral interaction attention mechanism are good at constructing the dependency relationship among the target and the context to complete the task of focusing on the key point and capturing the information of the important context, which has a strong and obvious effect on the improvement of the effect of the model in this chapter.
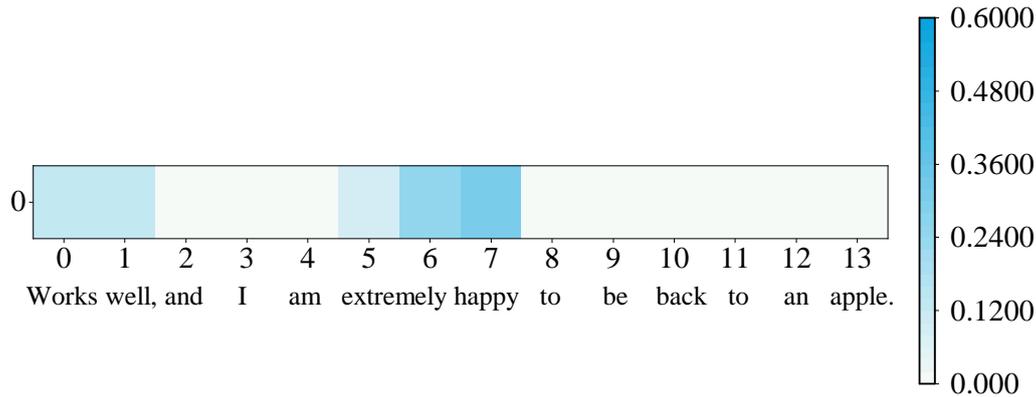
Figure 3. Visual analysis of the proposed model. The bilateral interaction attention mechanism serves to capture the words that are more strongly related to the emotional polarity of the target word, assign it a larger weight, and pay more attention to the semantic information of the key contextual words.

5.4. **Results analysis.** To confirm the viability of the proposed technique, it is compared with Bag-Bayesian [11], SVM [12], CNN [13], A-BiLSTM [14], Transformer [15], AEEAM [17], TEAN [19]. The results of comparison of classification accuracy of various methods mentioned above are shown in Table 4.

Table 4. Comparison results of classification accuracy of different methods

Field	Bag-Bayesian	SVM	CNN	A-BiLSTM	Transformer	AEEAM	TEAN	BA-BiLSTM
D→B	73.01	73.81	77.61	81.01	75.61	76.71	83.41	84.01
E→B	68.01	69.11	76.11	73.51	70.61	72.51	78.11	78.81
K→B	69.41	69.91	74.61	73.51	71.81	70.61	79.51	80.11
B→D	77.41	79.41	83.21	79.71	74.31	80.31	81.91	85.31
E→D	70.91	71.01	77.61	75.21	71.61	73.71	77.31	79.71
K→D	72.71	72.81	77.61	76.31	72.01	74.41	80.71	81.91
B→E	71.91	72.51	72.61	74.81	68.61	75.51	81.81	84.01
D→E	71.31	72.11	77.61	77.71	75.11	75.61	81.11	83.41
K→E	81.91	82.81	85.21	83.51	82.71	83.01	83.11	87.11
B→K	75.11	77.81	78.61	78.71	79.41	76.01	83.11	85.31
D→K	73.81	74.61	81.61	78.41	78.21	78.01	84.21	86.01
E→K	81.61	83.51	87.71	85.51	85.11	85.61	90.61	91.11
Average	73.93	74.95	79.18	78.16	75.43	76.84	82.08	83.90

The findings from the experiments indicate that the BA-BiLSTM model, achieves the top precision across all sentiment classification tasks, averaging an accuracy of 83.90%. The BA-BiLSTM model has enhanced its mean accuracy by 1.83% compared to the TEAN method, which has better classification results.

6. **Conclusion.** In this paper, a cross-border e-commerce product review sentiment analysis model based on interactive attention mechanism and BiLSTM is proposed. By introducing the bilateral interactive attention mechanism, the model is able to more accurately capture the complex dependencies between target words and their contexts, thus improving the accuracy and robustness of sentiment classification. Meanwhile, the improved target word representation effectively solves the problems of information dilution and noise introduction when dealing with multi-word target words in traditional methods.

Experimental results show that the proposed BA-BiLSTM model exhibits high accuracy in multiple cross-domain sentiment classification tasks, with an average classification accuracy of 83.90%, which is significantly improved compared to existing methods. In addition, the interpretability of the model is enhanced by visualising the attention weights, which enables users to understand the decision-making process of the model more intuitively. This study not only provides an effective method for cross-border e-commerce sentiment analysis, but also offers new ideas to solve the problems of cross-domain sentiment analysis, complex target word processing and data imbalance. However, there are still some challenges in the application of the model in this paper, such as further improving the computational efficiency of the model and validation on more diverse cross-domain datasets. Future research can continue to optimise the model structure and explore more efficient feature extraction and representation methods to further improve the performance and application scope of sentiment analysis.

REFERENCES

[1] E. Gomez-Herrera, B. Martens, and G. Turlea, "The drivers and impediments for cross-border e-commerce in the EU," *Information Economics and Policy*, vol. 28, pp. 83-96, 2014.

[2] A. Liu, M. Osewe, Y. Shi, X. Zhen, and Y. Wu, "Cross-border e-commerce development and challenges in China: A systematic literature review," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 17, no. 1, pp. 69-88, 2021.

[3] B. B. Hazarika, and R. Mousavi, "Review of cross-border e-commerce and directions for future research," *Journal of Global Information Management (JGIM)*, vol. 30, no. 2, pp. 1-18, 2021.

[4] T.-Y. Wu, H. Li, S. Kumari, and C.-M. Chen, "A Spectral Convolutional Neural Network Model Based on Adaptive Fick's Law for Hyperspectral Image Classification," *Computers, Materials & Continua*, vol. 79, no. 1, pp. 19-46, 2024.

[5] F. Zhang, T.-Y. Wu, J.-S. Pan, G. Ding, and Z. Li, "Human motion recognition based on SVM in VR art media interaction environment," *Human-centric Computing and Information Sciences*, vol. 9, 40, 2019.

[6] V. M. Joshi, R. B. Ghongade, A. M. Joshi, and R. V. Kulkarni, "Deep BiLSTM neural network model for emotion detection using cross-dataset approach," *Biomedical Signal Processing and Control*, vol. 73, pp. 103407, 2022.

[7] G. Xu, Y. Meng, X. Qiu, Z. Yu, and X. Wu, "Sentiment analysis of comment texts based on BiLSTM," *IEEE Access*, vol. 7, pp. 51522-51532, 2019.

[8] P. Ye, and L. Liu, "A Study of the Factors That Influence the Willingness of Users to Leave Comments on E-Commerce Platforms," *Information Resources Management Journal (IRMJ)*, vol. 34, no. 1, pp. 1-18, 2021.

[9] S. Li, C. Jiang, and L. Li, "Research on Chinese well-known e-commerce enterprises' innovation ability based on real comment," *International Journal of Computing Science and Mathematics*, vol. 11, no. 1, pp. 54-62, 2020.

[10] J. Ke, Y. Wang, M. Fan, X. Chen, W. Zhang, and J. Gou, "Discovering e-commerce user groups from online comments: An emotional correlation analysis-based clustering method," *Computers and Electrical Engineering*, vol. 113, pp. 109035, 2024.

[11] A. Almestekawy, and M. Abdulsalam, "Sentiment analysis of product reviews using bag of words and bag of concepts," *International Journal of Electronics and Information Engineering*, vol. 11, no. 2, pp. 49-60, 2019.

[12] C. Torres-Valencia, M. Álvarez-López, and Á. Orozco-Gutiérrez, "SVM-based feature selection methods for emotion recognition from multimodal data," *Journal on Multimodal User Interfaces*, vol. 11, pp. 9-23, 2017.

[13] H. Li, Y. Lu, H. Zhu, and Y. Ma, "A Novel AB-CNN Model for Multi-Classification Sentiment Analysis of e-Commerce Comments," *Electronics*, vol. 12, no. 8, 1880, 2023.

[14] S. Wang, Y. Zhu, W. Gao, M. Cao, and M. Li, "Emotion-semantic-enhanced bidirectional LSTM with multi-head attention mechanism for microblog sentiment analysis," *Information*, vol. 11, no. 5, 280, 2020.

[15] X. Wang, and Y. Tong, "Application of an emotional classification model in e-commerce text based on an improved transformer model," *Plos One*, vol. 16, no. 3, e0247984, 2021.

[16] X. Yan, H. Xue, S. Jiang, and Z. Liu, "Multimodal sentiment analysis using multi-tensor fusion network with cross-modal modeling," *Applied Artificial Intelligence*, vol. 36, no. 1, 2000688, 2022.

[17] A. Lieto, G. L. Pozzato, S. Zoia, V. Patti, and R. Damiano, "A commonsense reasoning framework for explanatory emotion attribution, generation and re-classification," *Knowledge-Based Systems*, vol. 227, 107166, 2021.

[18] F. Ma, Y. Li, S. Ni, S.-L. Huang, and L. Zhang, "Data augmentation for audio-visual emotion recognition with an efficient multimodal conditional GAN," *Applied Sciences*, vol. 12, no. 1, 527, 2022.

[19] C. Luna-Jiménez, D. Griol, Z. Callejas, R. Kleinlein, J. M. Montero, and F. Fernández-Martínez, "Multimodal emotion recognition on RAVDESS dataset using transfer learning," *Sensors*, vol. 21, no. 22, 7665, 2021.

[20] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artificial Intelligence Review*, vol. 55, no. 7, pp. 5731-5780, 2022.

[21] M. Birjali, M. Kasri, and A. Beni-Hssane, "A comprehensive survey on sentiment analysis: Approaches, challenges and trends," *Knowledge-Based Systems*, vol. 226, 107134, 2021.

[22] P. Mehta, and S. Pandya, "A review on sentiment analysis methodologies, practices and applications," *International Journal of Scientific and Technology Research*, vol. 9, no. 2, pp. 601-609, 2020.

[23] P. Nandwani, and R. Verma, "A review on sentiment analysis and emotion detection from text," *Social Network Analysis and Mining*, vol. 11, no. 1, 81, 2021.

[24] J. Hartmann, M. Heitmann, C. Siebert, and C. Schamp, "More than a feeling: Accuracy and application of sentiment analysis," *International Journal of Research in Marketing*, vol. 40, no. 1, pp. 75-87, 2023.

[25] A. Yadav, and D. K. Vishwakarma, "Sentiment analysis using deep learning architectures: a review," *Artificial Intelligence Review*, vol. 53, no. 6, pp. 4335-4385, 2020.

[26] K. Chakraborty, S. Bhattacharyya, and R. Bag, "A survey of sentiment analysis from social media data," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 2, pp. 450-464, 2020.

[27] F. Sabahi, "Introducing convex BiLSTM-based controller applied to 3-PSP spatial parallel robot manipulator," *Transactions of the Institute of Measurement and Control*, vol. 45, no. 16, pp. 3126-3137, 2023.

[28] C. Sekhar, and R. Dahiya, "Robust framework based on hybrid deep learning approach for short term load forecasting of building electricity demand," *Energy*, vol. 268, 126660, 2023.

[29] Z. Hameed, and B. Garcia-Zapirain, "Sentiment classification using a single-layered BiLSTM model," *IEEE Access*, vol. 8, pp. 73992-74001, 2020.

[30] A. Mitra, and S. Mohanty, "Sentiment analysis using machine learning approaches," *Journal of Ubiquitous Computing and Communication Technologies (UCCT)*, vol. 2, no. 3, pp. 145-152, 2020.