

Article

Sentiment Analysis of Chinese E-Commerce Product Reviews Using ERNIE Word Embedding and Attention Mechanism

Weidong Huang, Miao Lin * and Yuan Wang

School of Management, Nanjing University of Posts and Telecommunications, Nanjing 210000, China; huangwd@njupt.edu.cn (W.H.); 1019111912@njupt.edu.cn (Y.W.)

* Correspondence: 1021112611@njupt.edu.cn

Abstract: The development of e-commerce has ushered in a golden age. E-commerce product reviews are remarks initiated by online shopping users to evaluate the quality and service of the products they purchase; these reviews help consumers learn the reality of the product. The sentiment polarity of e-commerce product reviews is the best way to obtain customer feedback on products. Therefore, sentiment analysis of product reviews on e-commerce platforms is greatly significant. However, the challenges of sentiment analysis of Chinese e-commerce product reviews lie in dimension mapping, disambiguation of sentiment words, and polysemy of Chinese words. To solve the above problems, this paper proposes a sentiment analysis model ERNIE-BiLSTM-Att (EBLA). Here, the dynamic word vector generated using the Enhanced Representation through Knowledge Integration (ERNIE) word embedding model is input into the Bidirectional Long Short-term Memory (BiLSTM) to extract text features. Then, the Attention Mechanism (Att) is used to optimize the weight of the hidden layer. Finally, *softmax* is used as the output layer for sentiment classification. The experimental results on the JD.com Chinese e-commerce product review dataset show that the proposed model achieves more than 0.87 in precision, recall, and *F1* values, which is superior to classic deep learning models proposed by other researchers; it has strong practicability in sentiment analysis of Chinese e-commerce product reviews.

Keywords: ERNIE; attention mechanism; BiLSTM; sentiment analysis; deep learning; Chinese e-commerce product reviews



Citation: Huang, W.; Lin, M.; Wang, Y. Sentiment Analysis of Chinese E-Commerce Product Reviews Using ERNIE Word Embedding and Attention Mechanism. *Appl. Sci.* **2022**, *12*, 7182. <https://doi.org/10.3390/app12147182>

Academic Editor: Antonio Sarasa Cabezuelo

Received: 19 June 2022

Accepted: 15 July 2022

Published: 16 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The e-commerce sector is rapidly developing in the age of network information. According to the “China E-commerce Report 2020” released by the Department of E-commerce and Information Technology of the Ministry of Commerce of the People’s Republic of China, the national e-commerce transaction volume in 2020 increased to 37.21 trillion yuan, of which the national online retail sales rose to 11.76 trillion yuan, and China’s online shopping users reached 782 million. The development of E-commerce has ushered in a golden age. Compared with offline shopping, online shopping can be conducted anytime, anywhere, without leaving home, saving time and effort [1]. Although it has become the first choice for many consumers, e-commerce as a business activity based on virtual networks has some defects: being unable to check the real situation of the product, being prone to product purchases that do not meet expectations, or to false business propaganda. The e-commerce product reviews by consumers who have purchased the product provide references for other consumers [2]. They are remarks initiated by online shopping users to evaluate the quality and service of the products they purchase [3], which objectively reflect users’ willingness to purchase products and their satisfaction levels. Due to the vigorous development of e-commerce and the habit of consumers commenting after purchasing products on e-commerce platforms, several e-commerce product reviews have been generated. Consumers can obtain evaluations and opinions on commodities from these

e-commerce platforms. These comments and opinions express the sentiment tendency of consumers who have purchased products, and the commodity-related comments provided by them greatly influence the purchase intentions of other consumers [4]. Consumers are accustomed to consulting product reviews before purchase to reduce perceived uncertainty and online shopping risks due to information asymmetry. The sentiment polarity of e-commerce product reviews is the best way to obtain customer feedback on products. Therefore, positive reviews establish a good brand image for businesses to attract more consumers. Neutral reviews make businesses realize the insufficiency in some aspects of products or services. However, negative reviews greatly affect consumers' purchase intention, product sales, and the normal operation of enterprises. Therefore, sentiment analysis of product reviews on e-commerce platforms is greatly significant.

Sentiment analysis, also known as opinion mining [5], was originally proposed by Yang Bo [6]. It is the analysis of a certain product's review to find consumers' attitudes and opinions about the product. Liu Bing [7] further elaborated the concept of sentiment analysis research: sentiment analysis research or opinion mining refers to the statistical analysis of people's views, emotions, judgments, and attitudes toward commodities, businesses, groups, people, social problems, social events, and their basic attributes. More so, sentiment analysis of e-commerce product reviews is a process of analyzing the content of user reviews on e-commerce platforms and mining the attitudes of these reviews. Recently, sentiment analysis research on e-commerce product reviews has become a hot topic. Panthathi et al. [8] used word2vec to learn word embedding and convolution neural networks to train and classify the sentiment classes of the product reviews. Yan et al. [9] proposed a parallel Convolutional Neural Networks (CNN) and BiLSTM-attention two-channel neural network for sentiment analysis of e-commerce product reviews. However, although the above models have achieved good results, they have not considered dimension mapping, disambiguation of sentiment words, and polysemy of words in input. The sentiment analysis of e-commerce product reviews is considered a multidimensional classification process, which includes three dimensions: document, sentence, and aspect level tasks. These three dimensions correspond to the specific aspects of the whole document, each sentence, and the entity for sentiment polarity classification. The challenges of sentiment analysis of e-commerce product reviews lie in dimension mapping, disambiguation of sentiment words, and polysemy of words. The dimension mapping problem is the problem of mapping the review text to the correct dimension. The sentiment word disambiguation problem refers to the situation where the sentiment word in the review text is connected with two or more dimensions. In the word polysemy problem, for example, the word "Xiaomi" in appliances is a brand name, but as a food, it means the grain.

In order to solve the above problems and to improve the effectiveness of the sentiment analysis model in Chinese e-commerce product review texts, this paper combines the ERNIE word embedding model and Bidirectional Long Short-Term Memory (BiLSTM) and adds an attention mechanism (Att) [10] to the network structure. First, the ERNIE word embedding model is used to obtain the representation of the dynamic word vector of the review text. Then, BiLSTM is used combined with the attention mechanism to construct a sentiment analysis model called ERNIE-BiLSTM-Att (EBLA). The main contributions of this paper are as follows:

1. We target sentiment analysis in the field of Chinese e-commerce product reviews. Based on the ERNIE word embedding model, BiLSTM model, and attention mechanism, we proposed a Chinese e-commerce review sentiment analysis model EBLA, which was tested on the real crawled Jingdong appliance review dataset and achieved good results, the model outperforms many classic deep learning models proposed by other researchers, can provide more accurate reference for consumers and merchants. It has strong practicability in the field of sentiment analysis of Chinese e-commerce reviews.
2. Dynamic word embedding model ERNIE was used to obtain the representation of the dynamic word vector of the review text, which effectively solves the polysemy

problem in Chinese text. In addition, it addresses the challenge of dimension mapping in the sentiment analysis of e-commerce product reviews.

3. Since BiLSTM can fully capture contextual semantic information from the front and rear directions, it was applied to the sentiment analysis problem to solve the sentiment word disambiguation.

The rest of this paper is as follows: Section 2 introduces related work of this research; Section 3 introduces the composition of the EBLA model; Section 4 describes the experimental process and analysis, and concludes. Section 5 discusses and summarizes these experiments.

2. Related Work

2.1. ERNIE Word Embedding

Many deep learning models in NLP require word embedding results as input features [11]. Word embedding, also known as distributed representation, was proposed by Hinton (1986) [12] as the idea of distributed representation of words: words are represented as dense, low-dimensional, and continuous vectors. The word embedding model is divided into static and dynamic word embedding models.

Common static word embedding models include the Neural Network Language Model (NNLM) [13], Word2Vec [14], Global Vectors for Word Representation (GloVe) [15], etc. These word embedding models have strong semantic representation capabilities. However, NNLM, Word2Vec, and GloVe all belong to the static word embedding representation, and for them, the relationship between words and vectors is one-to-one, so the problem of polysemy cannot be solved.

The above problem can be improved by introducing a more complex neural network for contextual information modeling. Dynamic word embedding technology emerges as a perfect fit. The Enhanced Representation through Knowledge Integration (ERNIE) [16] dynamic word embedding model used in this paper is an improvement of the Bidirectional Encoder Representation from Transformers (BERT) [17] model, and its structure is shown in Figure 1. The main structure of ERNIE includes multi-layer Transformer [10] blocks and knowledge masking. Multi-layer transformer blocks include 12 transformer encoder (Trm) layers, 768 hidden units, and 12 attention heads. A transformer is a Seq2Seq [18] model based on the self-attention mechanism. Its structure is encoder–decoder, which abandons the traditional recursion and convolution. It uses an attention mechanism to capture context information and generate context embedding. ERNIE mainly uses its encoder part.

Compared with the BERT model, ERNIE improves its Mask Language Model (MLM) and proposes an optimized mask strategy, as shown in Figure 2. The masking strategy of BERT is to randomly mask Chinese characters in Chinese sentences during pre-training. However, the Chinese words differ in meaning from those of the characters that make them up [19]. Therefore, this masking policy tends to sever the relationships between Chinese characters, which is not conducive to the learning of knowledge information. The mask strategy adopted by the ERNIE model includes three levels: basic-level, phrase-level, and entity-level. Figure 3 describes different mask levels in a sentence. This improved strategy forcing the model to predict the content of the mask part through global information, it enables the model to learn a more complete semantic representation. During the model training process, the review text is divided into several short text blocks to more effectively map the text blocks to the corresponding dimensions and can solve the dimension mapping problem in sentiment analysis of e-commerce product reviews.

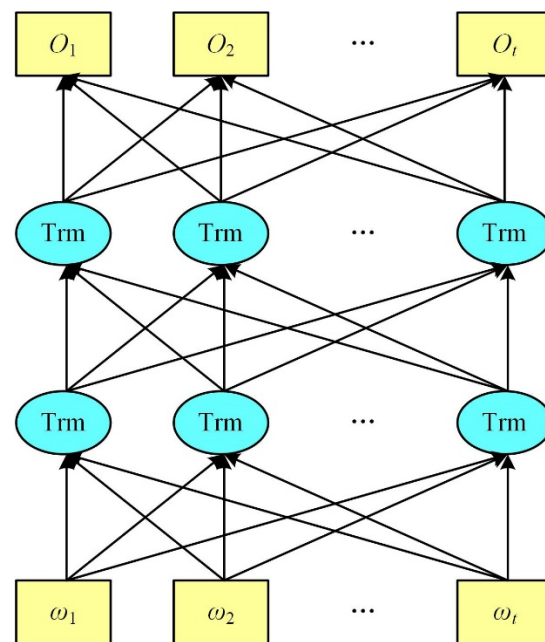


Figure 1. The structure of ERNIE.

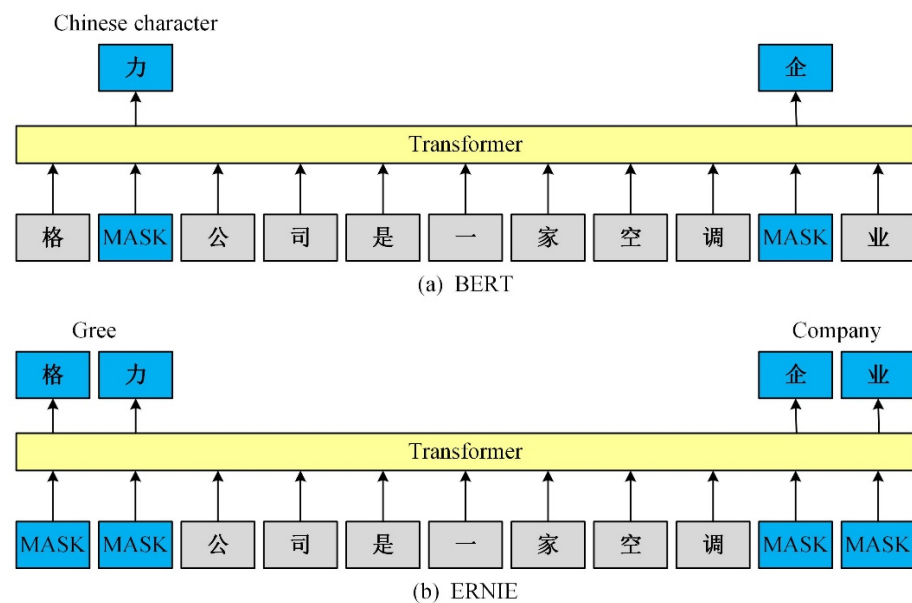


Figure 2. Comparison of BERT and ERNIE masking strategy in Chinese sentence.

Sentence	格	力	公	司	是	一	家	专	业	化	的	空	调	企	业
Basic-Level	MASK	力	公	司	是	一	MASK	专	业	化	的	空	调	企	MASK
Phrase-Level	MASK	MASK	公	司	是	一	MASK	专	业	化	的	空	调	MASK	MASK
Entity-level	MASK	MASK	公	司	是	MASK	MASK	专	业	化	的	空	调	MASK	MASK

Figure 3. Different mask levels of ERNIE in Chinese sentence.

2.2. BiLSTM

Recently, deep learning models have made gratifying achievements in NLP and have been widely used in sentiment analysis. Compared with traditional machine learning

methods, deep learning does not require human intervention features but a massive data corpus as support. The main idea of deep learning is to use deep neural networks to learn complex features extracted from data without extensive external intervention [20]. Typical deep learning models include Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Recurrent Convolutional Neural Network (RCNN), and so on [21]. The text length of e-commerce product reviews is on the rise, especially that of the negative review, and traditional neural network models cannot completely capture the entire context of the text. The traditional RNN is prone to the gradient disappearance problem during the training process, which makes RNN unsuitable for long-term-dependent scenarios. LSTM is a special RNN, whose structural unit includes four parts: memory cell, forget gate, input gate, and output gate. The gate mechanism is introduced to effectively solve the traditional cycle neural network problems with long-term dependencies and better capture longer-distance dependencies. However, LSTM solves the problem of long-term dependence using previous historical information and does not consider the importance of the latter, while BiLSTM solves this problem. BiLSTM is composed of forward and backward LSTMs. The forward and backward LSTMs are responsible for the forward and reverse operations of the input information, respectively. BiLSTM can not only fully capture contextual semantic information, but captures the interaction between words, such as sentiment words, adjectives, degree adverbs, etc., to better capture the bidirectional semantic dependencies of the comment text and solve the problem of sentiment word disambiguation. The structure of BiLSTM is shown in Figure 4.

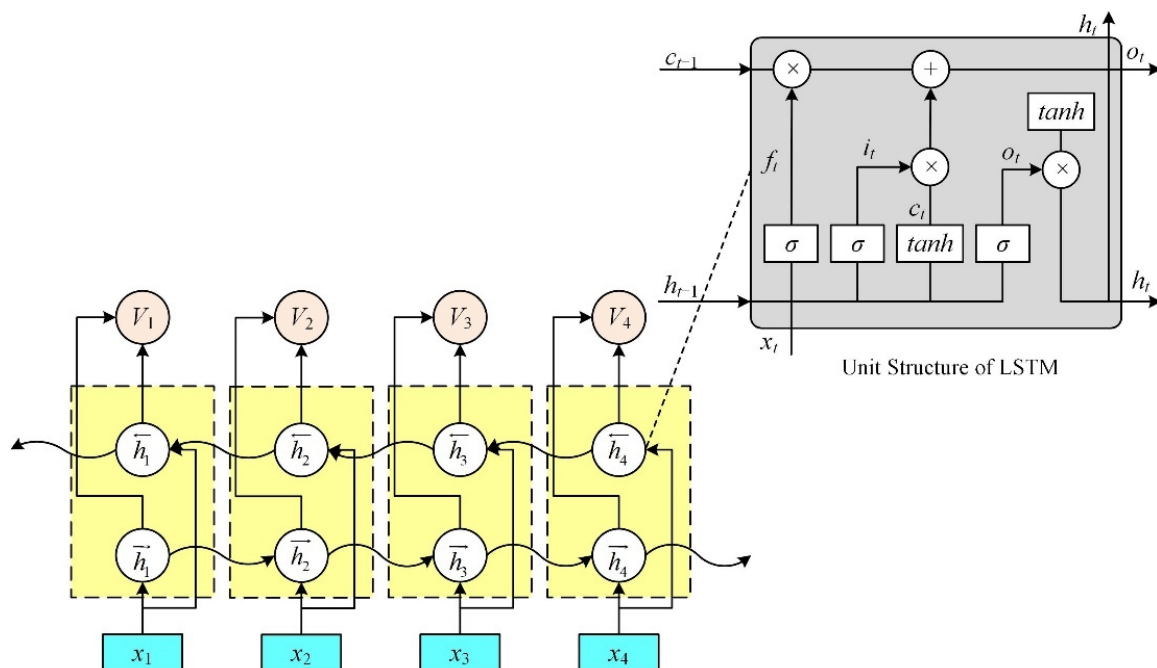


Figure 4. The structure of BiLSTM.

2.3. Attention Mechanism

The attention mechanism has become a powerful technique in deep learning recently, bringing many breakthroughs and has improved the performance of neural networks. Without the attention mechanism, the text translation relies on consulting the entire sentence and compressing all text information into a fixed-length vector. Furthermore, problems such as information loss or inaccurate translation are prone to occur, resulting in inaccurate sentiment analysis. In 2017, Vaswani et al. [10] proposed a Transformer model using the attention structure, which can weigh the information in the sequence, effectively enable the model to pay attention to a part of the input sequence, highlight the characteristics with high correlation using the training task in the training process, enhance its weight

representation, and achieve high-quality feature extraction while realizing the high-speed parallel operation. The introduction of the attention mechanism enables the model to highlight the weight of “sentiment words” in the e-commerce product review, reducing the information redundancy and lack of information in the feature extraction process, and improving the performance of the model.

3. Method and Materials

To improve the information representation ability of the model and accuracy of the sentiment analysis, this paper combined ERNIE, bidirectional LSTM model, and attention mechanism to propose the EBLA e-commerce product review sentiment analysis model. The structure is shown in Figure 5, which mainly includes five parts: input layer, ERNIE word embedding layer, BiLSTM layer, attention mechanism layer, and output layer.

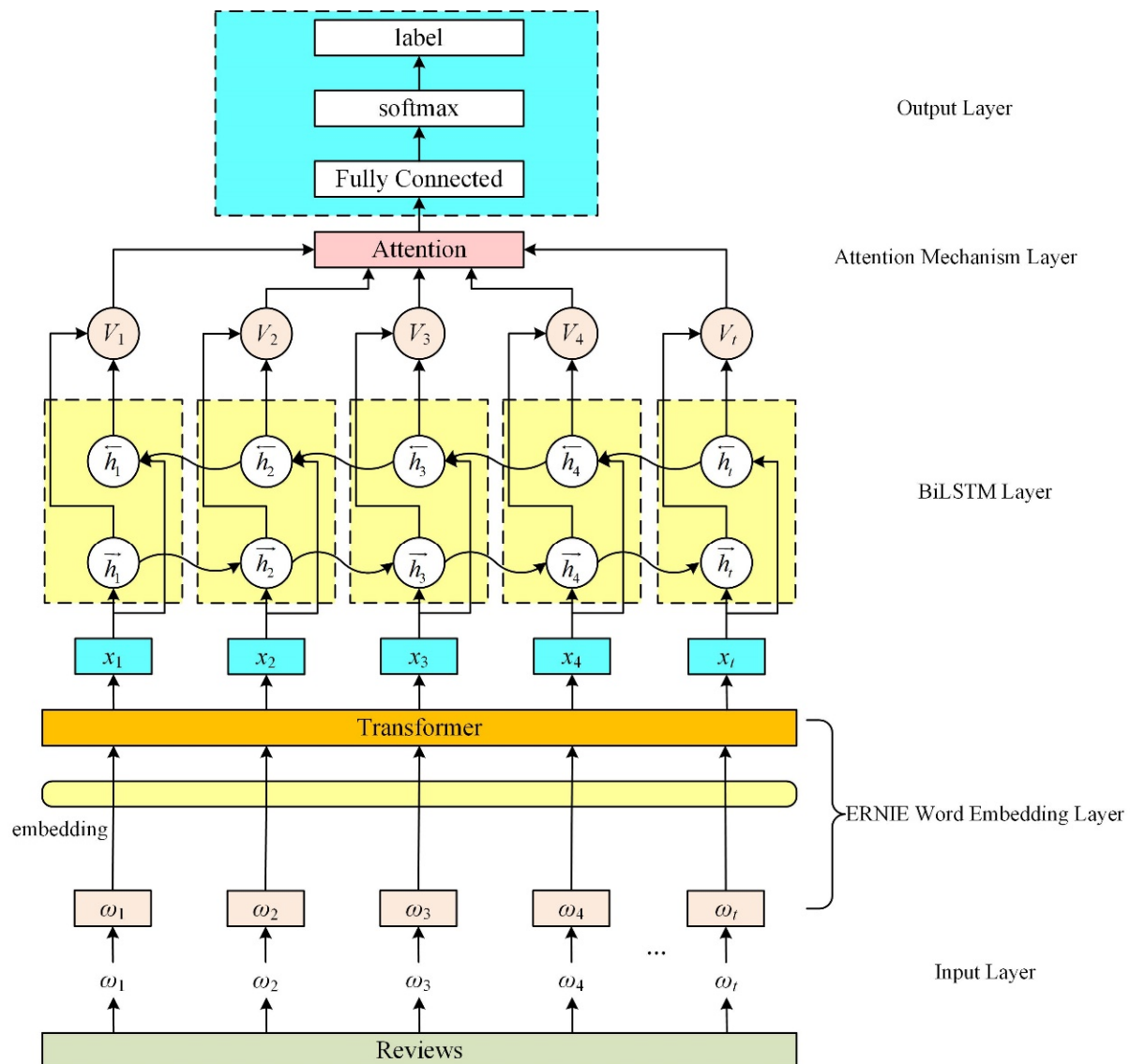


Figure 5. The structure of the EBLA model.

3.1. Input Layer

The e-commerce product reviews crawled from the e-commerce website were pre-processed, including removing stop words, punctuation, removing duplicate content, and

cutting sentences into the same length. Then, the preprocessed data were input into the ERNIE word embedding layer as:

$$w_1 = (\omega_1, \omega_2, \omega_3, \omega_4, \dots, \omega_t) \quad (1)$$

3.2. ERNIE Word Embedding Layer

The ERNIE model essentially adopts multi-layer Transformer blocks, and uses the bidirectional Transformer encoder structure for text vector representation. It converts the preprocessed data into a digital matrix; each column and row of the matrix represent a recognized feature and a specific review, respectively. The Encoder structure in the Transformer unit of the ERNIE model is shown in Figure 6.

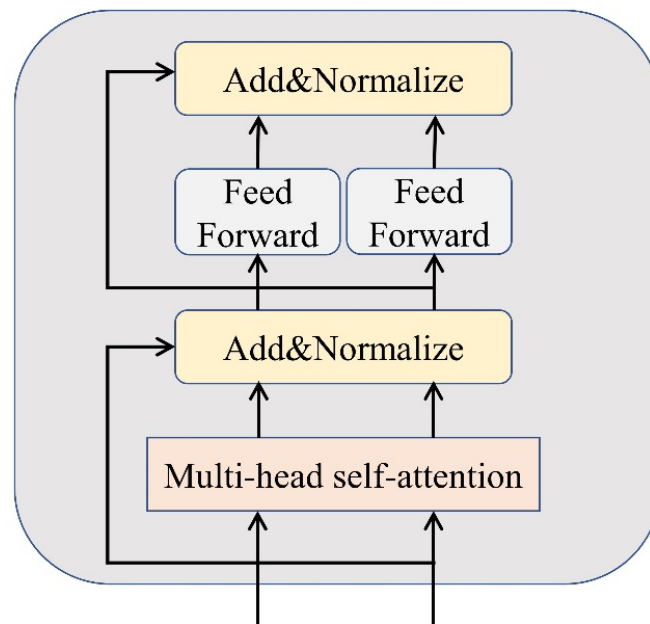


Figure 6. Encoder structure in the Transformer unit of the ERNIE model.

After converting text to word embedding, enter it into the self-attention layer in Transformer to get context information. Then the multi-head self-attention layer mixes word vectors and context information and sends the results to the Add & Normalize layer for residual connection and normalization. Subsequently, the processed text vector is linearly changed through the feedforward neural network. Finally, the residual connection and normalization are performed again, and the dynamic word embedding representation of the reviews is obtained by combining the prior semantic information obtained by training and masking strategies in the heterogeneous corpus.

The calculation process of multi-head self-attention layer in ERNIE is as follows:

First, the input vector matrix is subjected to position encoding to obtain the vector matrix Y . Then, the matrix Y is multiplied by the three weighted matrices, respectively, and the word correlation degree is calculated. Finally, the output of the single self-attention layer is obtained by normalization. The calculation formula is given by Equations (2) and (3):

$$Q = \text{Linear}(Y) = YW^q \quad K = \text{Linear}(Y) = YW^k \quad V = \text{Linear}(Y) = YW^v \quad (2)$$

$$S_h = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (3)$$

where represent W^q , W^k , W^v weighted matrices, d_k is the dimension of the input vector, and the *softmax* function normalizes the output result to obtain the weight sum of the word vector.

Next, the output of the self-attention layer is combined to obtain the output of the multi-head attention mechanism. The calculation formula is given by Equation (4):

$$S_{mh} = \text{Concat}(S_{h1}, S_{h2}, \dots, S_{ht}) \quad (4)$$

Finally, the operation of residual connection-linear change-residual connection is performed and combined with prior semantic information to obtain the dynamic representation of the review:

$$w_2 = (x_1, x_2, x_3, x_4, \dots, x_t) \quad (5)$$

3.3. BiLSTM Layer

When the time is t , the calculation formula for the state update of each LSTM unit is given by Equations (6)–(11):

$$f_t = \sigma(w_f \cdot [h_{t-1}, x_t] + b_f) \quad (6)$$

$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i) \quad (7)$$

$$\tilde{c}_t = \tanh(w_c \cdot [h_{t-1}, x_t] + b_c) \quad (8)$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t \quad (9)$$

$$o_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o) \quad (10)$$

$$h_t = o_t \circ \tanh(c_t) \quad (11)$$

where σ represents the activation Sigmoid function; \circ represents the corresponding multiplication of homologous elements; $[]$ represents the connection of two vectors; x_t represents the LSTM input vector at time t ; h_t and h_{t-1} represent the hidden states at time t and $(t - 1)$, respectively; f_t , i_t , c_t , o_t represent the values of the forget gate, input gate, memory cell, and output gate; w_f , w_i , w_c , w_o represent the weight matrix of the forget gate, input gate, memory cell, and output gate; b_f , b_i , b_c , b_o represent the bias terms of the forget gate, input gate, memory cell and output gate.

Equation (6) is the calculation process for the forget gate, which determines how much of the memory cell state at the previous moment is saved at the current time; Equation (7) is the operation process of the input gate, which determines which input vectors x_t are stored in the unit state c_t at the current moment; Equations (8) and (9) calculate temporary cell states \tilde{c}_t and c_t at the current time, respectively; Equation (10) is the calculation process of the output gate, which controls how much the cell state c_t outputs to the current output value h_t of the LSTM; Equation (11) calculates the hidden layer state h_t at the current time using the calculation results of the output gate and the unit state c_t .

The dynamic word vector obtained from ERNIE was input into the BiLSTM layer. Then, we obtained the forward and backward hidden layer vector sums using the forward and backward modules, respectively. Next, both vector sums are spliced to obtain the hidden layer vector. As shown in Equation (12), V represents the set of hidden layer vectors at all times, as shown in Equation (13):

$$v_t = \left[\begin{matrix} \vec{h}_t, \overleftarrow{h}_t \end{matrix} \right] \quad (12)$$

$$V = [v_1, v_2, v_3, \dots, v_t] \quad (13)$$

3.4. Attention Mechanism Layer

The attention mechanism was applied to the output of the BiLSTM layer to highlight features highly relevant to the training task.

The attention mechanism aims to obtain the corresponding weight value by calculating a series of key-value pairs. The calculation of the attention mechanism is mainly divided into three steps:

First, given the query (vector-matrix *query*), key (vector-matrix *key*) and the input (vector-matrix *value*) of the task, the similarity calculation is performed using the vector-matrix *query*, and vector-matrix *key*, where *Sim* represents the attention scoring function of the dot product model, and s_i represents the attention score, as shown in Equation (14):

$$s_i = \text{Sim}(\text{query}, \text{key}) = \text{query} \otimes \text{key}^T \quad (14)$$

The second step is to normalize the obtained weight coefficients to obtain a probability distribution with a weight coefficient sum of one and highlight the weights of important elements, where a_i represents the attention distribution, as shown in Equation (15):

$$a_i = \text{SoftMax}(\text{Sim}(q, k)) = \frac{\exp(s_i)}{\sum_{j=1}^N \exp(s_j)} \quad (15)$$

The third step is to multiply the weight coefficient by the corresponding *value* to obtain the attention-weighted vector. The attention calculation formula is shown in Equation (16):

$$\begin{aligned} \text{Attention}((K, V), q) &= \sum_{i=1}^N \alpha_i \otimes \text{value} \\ &= \sum_{i=1}^N \frac{\exp(s(k_i, q))}{\sum_{j=1}^N \exp(s(k_j, q))} \otimes \text{value} \end{aligned} \quad (16)$$

3.5. Output Layer

The weighted output vector obtained using the attention mechanism layer is input into the full connection layer for feature fusion, and the feature sequence R is obtained. Finally, the output results of the full connection layer are normalized using the *softmax* function, and the feature sequence of the full connection layer output is transformed into the probability distribution of sentiment polarity for sentiment classification. The calculation process is given in Equation (17)

$$\begin{aligned} y &= \text{softmax}(w \cdot R + b) \\ w &: \text{Weight matrix} \\ b &: \text{Bias} \end{aligned} \quad (17)$$

4. Experimental Design and Result Analysis

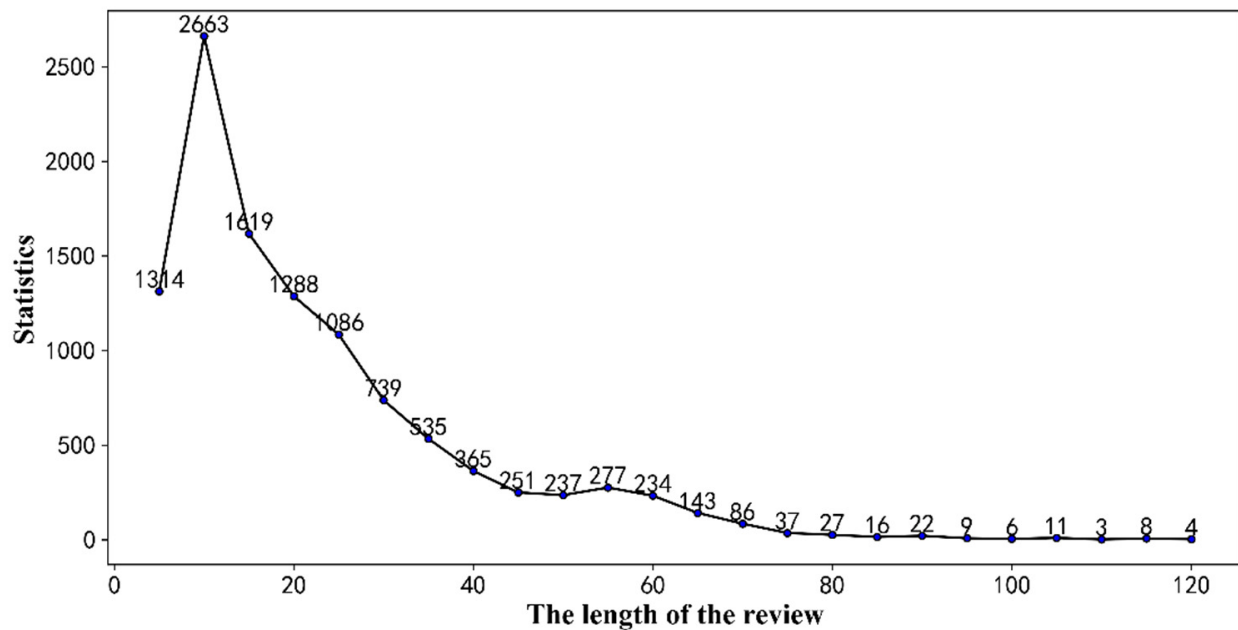
4.1. Dataset

The research data used in this paper are reviews of electrical products obtained from the Jingdong e-commerce website using a web crawler. We sent a request to the API of the comment content through Python's Requests module and stored the comment-related data obtained from the e-commerce platform in a dictionary list structure. Next, we added it continuously through paging access, and finally, saved reviews and their corresponding scores in csv and xlsx formats using the pandas module. Subsequently, we manually label the dataset, with score 1 and 2 labeled as negative reviews, score 3 as neutral reviews, and score 4 and 5 as positive reviews.

The dataset contains 11,000 reviews, including 6000 positive reviews, 2500 neutral reviews, and 2500 negative reviews. After preprocessing the dataset, the training, test, and validation sets were divided in the ratio of 6:2:2. The comment data contain three emotion labels, among which 0 represents positive emotion, 1 represents neutral emotion, and 2 represents negative emotion. Table 1 shows the obtained text data information and Figure 7 shows the statistical length of the review data.

Table 1. Text data information.

	Positive	Neutral	Negative	Total
Training set	3600	1500	1500	6600
Test set	1200	500	500	2200
Validation set	1200	500	500	2200
Total	6000	2500	2500	11,000

**Figure 7.** Length statistics of reviews.

4.2. Analysis of Word Frequency and Word Cloud

Word frequency counts and analyzes the occurrences of words in the text, and intuitively reflect the product attributes and supporting services that consumers focus on in e-commerce product reviews. In this paper, the Chinese word segmentation library Jieba in Python was used to remove stop words and word segmentation in the review texts and count the top fifteen words using word frequency. These data was used to make a word frequency diagram, as shown in Figure 8.

The word cloud map enables consumers to obtain relevant information about a product for the first time and meet their aesthetic needs for the preference of word appearance attributes [22]. Here, we generated a word cloud map using the WordCloud library in Python. The larger the font of the word in the word cloud map, the higher its frequency. The word cloud is shown in Figure 9.

From Figures 7 and 8, it can be seen that the words “installation”, “delivery”, “logistics”, and “after-sales” appear frequently in the commentary text. This indicates that consumers attached great importance to the after-sales service of products in the sales of electrical products, which greatly affects consumer satisfaction. For example, although the quality and performance of a certain product can be excellent, its slow installation or logistics can still earn negative reviews. Furthermore, the words “appearance”, “refrigeration”, “effect”, and “quality” show that consumers paid attention to the performance and quality of their products. The appearance of the word “customer service” shows that the level of customer serviceability in the e-commerce platform is also an important factor affecting consumer satisfaction. Customer service is a fundamental role in product sales, customer maintenance, and product after-sales-related work. The efficiency of the service determines the customer’s experience and invariably determines the customer’s satisfaction.

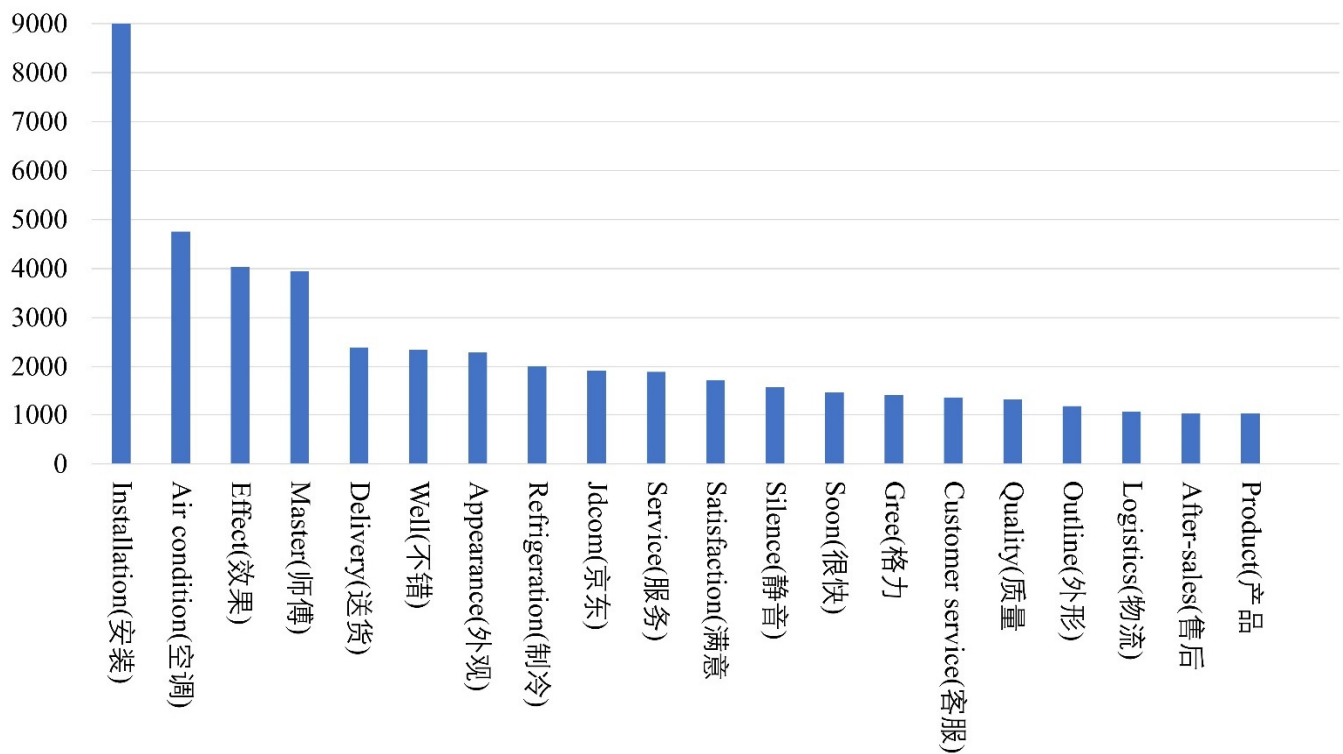


Figure 8. Word frequency diagram.



Figure 9. Word cloud.

4.3. Experimental Parameters

The operating environment of this experiment was Window10, Intel (R) Xeon (R) Gold 5218 CPU @ 2.30 GHz processor with a graphics card of GeForce RTX 3090. The Python programming language was used, and the training platform is Pytorch 1.11.0 deep learning framework. Furthermore, we set a stop training trigger on the model to stop the experiment when the experimental data exceeded one thousand batches and the effect of the model did not improve.

(1) The public parameter settings of the experiment are shown in Table 2:

Table 2. Public parameter settings.

Parameters	Value
require_improvement	1000
sequence_length	64
epochs	30
learning_rate	4×10^{-5}
batch_size	128
dropout	0.1

(2) The special parameter settings of each model are shown in Table 3:

Table 3. Special parameter settings.

Model	Parameters
TextCNN	filter_sizes = (2, 3, 4), num_filters = 256
RCNN	hidden_size = 256, num_layers = 1, pooling = Max
BiLSTM	hidden_size = 128, num_layers = 2
BiLSTM-Att	hidden_size = 128, num_layers = 2
BERT-CNN	hidden_size = 768, filter_sizes = (2, 3, 4), num_filters = 256
BERT-BiLSTM-Att	hidden_size = 768, filter_sizes = (2, 3, 4), num_filters = 256, rnn_hidden = 768, num_layers = 2
ERNIE-CNN	hidden_size = 768, filter_sizes = (2, 3, 4), num_filters = 256
ERNIE-RCNN	hidden_size = 768, filter_sizes = (2, 3, 4), num_filters = 256, rnn_hidden = 256, num_layers = 2, pooling=Max
ERNIE-BiLSTM	hidden_size = 768, filter_sizes = (2, 3, 4), num_filters = 256, rnn_hidden = 768, num_layers = 2
ERNIE-BiLSTM-Att	hidden_size = 768, filter_sizes = (2, 3, 4), num_filters = 256, rnn_hidden = 768, num_layers = 2

4.4. Performance Indicators

The performance indicators of precision (P), recall (R), and $F1$ value are used to evaluate the sentiment analysis model; these performance indicators range from 0 to 1, and the closer to 1, the better the performance of the model. Precision (P) represents the ratio of true correct samples in all the data predicted as correct samples; recall (R) represents the ratio of predicted correct samples in actual correct samples. $F1$ value trades off precision versus recall [23], which can comprehensively reflect the performance of the model. However, the $F1$ value is only applicable to the scenario with binary classification task, Macro $F1$, Micro $F1$, and Weighted $F1$ are often used as performance indicators for multi-classification tasks. In this experiment, the $F1$ value is calculated by weighted method, the Weighted $F1$

value solves the problem that Macro F1 does not consider the imbalance of samples, and the calculation formula is shown in Equation (18):

$$F1_i = \frac{2 \times P_i \times R_i}{P_i + R_i} \quad P_i = \frac{TP}{TP + FP} \quad R_i = \frac{TP}{TP + FN} \quad (18)$$

$$P_w = \sum_{i=1}^N c_i \times P_i \quad R_w = \sum_{i=1}^N c_i \times R_i \quad F1_w = \sum_{i=1}^N c_i \times F1_i$$

Here:

TP is True positive;

FP is False positive;

TN is True negative;

FN is False negative;

P_i is the precision (P) of the i -th classification;

R_i is the recall (R) of the i -th classification;

$F1_i$ is the F1 value of the i -th classification;

c_i is the proportion of the i -th classification sample to the total sample;

N is the total number of classifications. In this experiment, $N = 3$.

4.5. Results and Discussion

To evaluate the effectiveness of the EBLA model for sentiment analysis of e-commerce product reviews, we conducted two control experiments. The first group compares the sentiment analysis performance of models using different word embeddings, and the second group compares the sentiment analysis performance of different deep learning models.

(1) Comparative Experiment for Different Word Embedding Models

To verify that the ERNIE word embedding model has better semantic representation ability, we compared its performance with that of Word2Vec and BERT word embedding models. ERNIE used the ERNIE1.0 Chinese pre-training model published by Baidu with the word vector dimension of 768. BERT used the BERT-base network structure launched by Google with the word vector dimension of 768. Word2Vec used the Chinese word vector “Chinese Word Vector” constructed by Li [24] with the word vector dimension of 300. In this experiment, we divided the word embedding model into two combined deep learning models: TextCNN and BiLSTM-Att. The experimental results are shown in Table 4.

Table 4. Comparative results using different word embedding models %.

Model	R	P	F1
Word2Vec-CNN	83.21	83.59	83.33
BERT-CNN	85.59	85.50	85.50
ERNIE-CNN	86.37	86.77	86.36
Word2Vec-BiLSTM-Att	81.51	82.05	81.58
BERT-BiLSTM-Att	84.80	85.14	84.77
EBLA	87.87	87.55	87.48

The experimental results show that compared with the traditional static word embedding model Word2Vec, dynamic word embedding models, such as BERT and ERNIE, can dynamically adjust the different semantic representations of a word using the text context information, thereby effectively solving the polysemy problem of the static word vector.

(2) Comparative Experiment of Different Sentiment Analysis Model

To verify the effectiveness of this model in sentiment analysis of e-commerce product reviews, the sentiment classification performance of different models was compared in the same experimental environment. ERNIE adopted the ERNIE1.0 Chinese pre-training model released by Baidu, and other traditional deep learning models adopted Word2Vec constructs word vectors. Table 5 shows the experimental results of the experiment.

Table 5. Performance comparison of different sentiment analysis models %.

Model	R	P	F1
BiLSTM	74.72	75.73	74.44
TextCNN	83.21	83.59	83.33
RCNN	81.29	82.09	81.35
BiLSTM-Att	81.51	82.05	81.58
ERNIE	86.37	86.77	86.51
ERNIE-BiLSTM	86.56	86.64	86.52
ERNIE-CNN	86.37	86.77	86.36
ERNIE-RCNN	86.81	85.77	85.01
EBLA	87.87	87.55	87.48

BiLSTM: Bidirectional long short-term memory network, the forward hidden layer vector, and the backward hidden layer vector sum were obtained using the forward LSTM and backward LSTM modules, respectively, then both were spliced to obtain the hidden layer vector. Finally, the output is connected to the *softmax* function for sentiment classification.

TextCNN: This is a convolutional neural network for text classification, including convolutional and maximum pooling layers, and finally, the output is connected to a *softmax* function for sentiment classification.

RCNN: Here, the convolution layer in the general convolutional neural network is replaced with a BiLSTM network, and finally, the maximum pooling layer is taken and the output is connected to the *softmax* function for sentiment classification.

BiLSTM-Att: The attention mechanism is introduced using the BiLSTM network to obtain features highly relevant to the training task, and finally, the output is connected to an external *softmax* function for sentiment classification.

ERNIE: The ERNIE pre-training model adds [CLS] symbols and uses the corresponding output as the semantic representation of the text. Finally, the output is connected to a *softmax* function for sentiment classification.

ERNIE-BiLSTM: After the input text is vectorized and represented by the word embedding model ERNIE, it is input to the BiLSTM model to extract features, and finally, the output is externally connected to the *softmax* function for sentiment classification.

ERNIE-RCNN: After the input text is vectorized and represented by the ERNIE word embedding model, it is input to the RCNN model to extract features, and finally, the output is externally connected to the *softmax* function for sentiment classification.

ERNIE-CNN: After the input text is vectorized and represented by the ERNIE word embedding model, it is input to the TextCNN model to extract features, and finally, the output is externally connected to the *softmax* function for sentiment classification.

EBLA: After the input text is vectorized and represented by the ERNIE word embedding model, it is input to the BiLSTM-Att model to extract features, and finally, the output is externally connected to the *softmax* function for sentiment classification.

From the experimental results, all the metrics of TextCNN are optimal for the neural network using traditional static word embedding. This is attributed to the TextCNN model having a simple structure, which only contains a convolution layer and a maximum pooling layer. Therefore, it has a good classification effect on the small corpus. The introduction of the maximum pooling layer of TextCNN reduces the risk of overfitting, reduces parameters, speeds up the calculation speed, and obtains better experimental results. The worst effect is the BiLSTM model, which can be explained by analyzing the principle of its model structure. Its structure contains two LSTM layers with several model parameters, and the corresponding amount of data required is relatively large. The amount of data collected in this paper is small, so the effect is poor. However, after adding the attention mechanism (BiLSTM-Att), the classification effect significantly improved, indicating that the attention mechanism enables the model to learn more important features, thereby improving the performance of the neural network. Additionally, compared with the neural network using the traditional static word embedding, the neural network built with the ERNIE

word embedding model greatly improved its classification effect, with an average increase of 6.7% in precision, 5.82% in recall, and 6.17% in *F1* value. Furthermore, it shows that the ERNIE word embedding model has a strong semantic representation ability, and achieved good results in dealing with the dimension mapping and polysemy problems in the Chinese e-commerce product review. The combination of ERNIE and BiLSTM-Att effectively improved the performance of model sentiment classification.

The EBLA model used in this paper achieved the best results in the sentiment analysis experiment of the Chinese e-commerce product reviews. Compared with the BiLSTM, TextCNN, RCNN, BiLSTM-Att, ERNIE, ERNIE-BiLSTM, ERNIE-CNN, and ERNIE-RCNN models, the *F1* value of the EBLA model increased by 13.04%, 4.15%, 6.13%, 5.9%, 0.97%, 0.96%, 1.12%, and 2.47%, respectively. The EBLA model, too, achieves more than 0.87 in precision, recall, and *F1* values, and it can be seen that this model has an excellent performance on the sentiment analysis task of the Chinese e-commerce product reviews.

5. Conclusions

The rapid development of computer and network technology has seen a rapid increase in the number of Internet users and has become a platform for e-commerce to flourish. There are many products on the e-commerce platform, and the quality is uneven, which is prone to the phenomenon of inconsistency between product promotion and reality. To a certain extent, the massive comments on e-commerce platforms reflect the real information about the products, and the sentiment polarity in these comment texts contains the real feedback from consumers on the products. Therefore, the sentiment analysis of these e-commerce product reviews is of great significance.

Deep learning is currently the most popular technique in sentiment analysis, but the classification performance of many existing traditional models can be further improved. In this paper, we designed the EBLA sentiment analysis model to perform sentiment analysis on the Chinese e-commerce product reviews. The model obtains a more accurate dynamic text semantic representation using the ERNIE word embedding model, extracts contextual features using a bidirectional LSTM, combines the attention mechanism to highlight the features highly relevant to the training task, and performs sentiment classification. The problem of dimension mapping, sentiment word disambiguation, and word polysemy in the sentiment analysis of Chinese e-commerce product reviews is well handled. The experimental results on the [JD.com](https://jd.com) (accessed on 2 March 2022) e-commerce product review dataset show that the model achieved a better classification effect in the sentiment analysis task of Chinese e-commerce product reviews.

Based on our research, we found two directions for our future work. One direction is that the ERNIE pre-training language model has high hardware requirements and needs a massive data corpus as support. How to train a more lightweight ENIRE model is one of our future works. Another direction is that for some customers and business, it is not enough to know only the overall sentiment polarity of reviews; they would like to realize what aspects of the product or service are negative, neutral, or positive from other consumer reviews. Therefore, our future research can be extended to aspect-level sentiment analysis of e-commerce product reviews, and more detailed mining of consumers' specific opinions on various attributes of products, so that other consumers can further choose their favorite products according to the product attributes they value more, but also enables merchants to improve a specific aspect of the product.

GitHub: <https://github.com/Linshuis/Sentiment-analysis-model-ERNIE-BiLSTM-Att>.

Author Contributions: Conceptualization, W.H. and M.L.; methodology, W.H.; software, M.L. and Y.W.; validation, M.L.; formal analysis, W.H. and M.L. data curation, M.L.; writing—original draft preparation, M.L.; writing—review and editing, W.H.; supervision, Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Major Projects of National Social Science Fund of China (No. 16ZDA054) and the Jiangsu Provincial Department of education, the key research base of philosophy and social sciences.

Data Availability Statement: The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this article:

Att	Attention Mechanism
BERT	Bidirectional Encoder Representation from Transformer
BiLSTM	Bidirectional Long Short-term Memory
CNN	Convolutional Neural Networks
ELMo	Embeddings from Language Models
ERNIR	Enhanced Representation through Knowledge Integration
GloVe	Global Vectors for Word Representation
LSTM	Long Short-term Memory
MLM	Mask Language Model
NNLM	Neural Network Language Model
NSP	Next Sentence Prediction
RNN	Recurrent Neural Networks
RCNN	Recurrent Convolutional Neural Network

References

1. Yang, L.; Li, Y.; Wang, J. Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. *IEEE Access* **2020**, *8*, 23522–23530. [\[CrossRef\]](#)
2. Hu, M.; Liu, B. Mining and Summarizing Customer Reviews. In Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, WA, USA, 22–25 August 2004; pp. 168–177.
3. Guo, B.; Li, S.; Wang, H.; Zhang, X.; Gong, W.; Yu, Z.; Sun, Y. Examining Product Reviews with Sentiment Analysis and Opinion Mining. *Data Anal. Knowl. Discov.* **2017**, *12*, 1–9.
4. Ward, J.C.; Ostrom, A.L. The internet as information minefield: An analysis of the source and content of brand information yielded by net searches. *J. Bus. Res.* **2003**, *56*, 907–914. [\[CrossRef\]](#)
5. Pang, B.; Lee, L. Opinion Mining and Sentiment Analysis. *Found. Trends Inf. Retr.* **2008**, *2*, 1–135. [\[CrossRef\]](#)
6. Pang, B.; Lee, L.; Vaithyanathan, S. Thumbs up? Sentiment classification using machine learning techniques. *arXiv* **2002**, arXiv:cs/0205070.
7. Liu, B. *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*; Cambridge University Press: Cambridge, UK, 2020; pp. 1–2.
8. Panthathi, J.; Bhaskar, J.; Ranga, T.K.; Challa, M.R. Sentiment Analysis of Product Reviews Using Deep Learning. In Proceedings of the 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Bangalore, India, 19–22 September 2018; pp. 2408–2414.
9. Yan, L.; Zhu, X.; Chen, X. Emotional classification algorithm of comment text based on two-channel fusion and BiLSTM-attention. *J. Univ. Shanghai Sci. Technol.* **2021**, *43*, 597–605.
10. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is all you need. *Adv. Neural. Inf. Process Syst.* **2017**, *30*, 5998–6008.
11. Collobert, R.; Weston, J.; Bottou, L.; Karlen, M.; Kavukcuoglu, K.; Kuksa, P. Natural language processing (almost) from scratch. *J. Mach. Learn. Res.* **2011**, *12*, 2493–2537.
12. Hinton, G.E. Learning distributed representations of concepts. In Proceedings of the Eighth Annual Conference of the Cognitive Science Society, Amherst, MA, USA, 15–17 August 1986; p. 12.
13. Bengio, Y.; Ducharme, R.; Vincent, P. A neural probabilistic language model. *Adv. Neural. Inf. Process Syst.* **2000**, *13*, 1137–1155.
14. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient estimation of word representations in vector space. *arXiv* **2013**, arXiv:1301.3781.
15. Pennington, J.; Socher, R.; Manning, C.D. Glove: Global Vectors for Word Representation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014; pp. 1532–1543.
16. Sun, Y.; Wang, S.; Li, Y.; Feng, S.; Chen, X.; Zhang, H.; Wu, H. Ernie: Enhanced representation through knowledge integration. *arXiv* **2019**, arXiv:1904.09223.
17. Devlin, J.; Chang, M.; Lee, K.; Toutanova, K. BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: NAACL-HLT 2019, Minneapolis, MN, USA, 2–7 June 2019; Volume 1, pp. 4171–4186.

18. Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to sequence learning with neural networks. *Adv. Neural. Inf. Process Syst.* **2014**, *27*.
19. Hu, Y.; Tong, T.; Zhang, X.; Peng, J. Self-attention-based BGRU and CNN for Sentiment Analysis. *Comput. Sci.* **2022**, *49*, 252–258.
20. Araque, O.; Corcuera-Platas, I.; Sánchez-Rada, J.F.; Iglesias, C.A. Enhancing deep learning sentiment analysis with ensemble techniques in social applications. *Expert. Syst. Appl.* **2017**, *77*, 236–246. [[CrossRef](#)]
21. Yadav, A.; Vishwakarma, D.K. Sentiment analysis using deep learning architectures: A review. *Rev. Artif. Intell. Rev.* **2020**, *53*, 4335–4385. [[CrossRef](#)]
22. Zhou, Y.; Zhang, X.; Yu, X. User Preference Analysis Based on Product Review Mining. *Inf. Sci.* **2022**, *40*, 58–65.
23. Manning, C.D.; Raghavan, P.; Schütze, H. *Introduction to Information Retrieval*; Cambridge University Press: Cambridge, UK, 2008; pp. 155–156.
24. Li, S.; Zhao, Z.; Hu, R.; Li, W.; Liu, T.; Du, X. Analogical reasoning on chinese morphological and semantic relations. *arXiv* **2018**, arXiv:1805.06504.