Analysis of Product Selection Strategy for Cross-Border E-Commerce with Assistance of Artificial Intelligence

Lin Feng[*]

School of Foreign languages, Ankang University, China
FengLin5678123@yeah.net

Abstract

The expansion of the e-commerce sector beyond international boundaries has opened up new opportunities in international business. The first and most important problem for international e-commerce businesses is choosing the right products to sell. Scientific product selection is important not only for the smooth running of international e-commerce businesses, but also for the growth and development of thriving private labels. With development for Internet and rapid rise of e-commerce platforms, online shopping has become mainstream. While consumers are deeply involved, a large number of user comments appear on the Internet, and these online shopping comments often contain a lot of valuable information. Cross-border e-commerce can get useful suggestions and feedback about products through sentiment analysis of these texts. This can also find product problems, improve where they are not in place, and improve operational efficiency. This work studies the selection strategy of cross-border e-commerce products assisted by artificial intelligence technology. This paper uses artificial intelligence technology to conduct sentiment analysis of product reviews, which can help cross-border e-commerce companies make effective product selection. This work proposes CNN-BiGRU-ATT for sentiment polarity analysis of product reviews. First, the multi-channel convolutional neural network is utilized to extract local feature of different granularities. Then connect BiGRU for context sequence learning to memorize long-distance dependency information. This method not only solves the problem that CNN cannot perform context sequence learning, but also solves issue of gradient disappearance or explosion. Finally, attention is introduced after the BiGRU hidden layer to filter the important features.

Keywords: Product selection, Cross-border E-commerce, AI, CNN-BiGRU-ATT

1 Introduction

After years of development, China has evolved into an economic giant, and its standing on the global stage has also greatly improved. This provides solid support for Chinese enterprises to go global and expand the international market. At the same time, China has actively promoted the upgrading of its industrial structure, the structure of import and export trade has been continuously optimized, and the export competitiveness of products has been significantly improved. As China has successively joined the Asia-Pacific Economic Cooperation, RCEP and other multilateral trade organizations, and actively built the Belt and Road Initiative, it will help China to establish good bilateral and multilateral trade relations. On the other hand, with country's rapid development and promotion of informatization, it also brings new opportunities for the development of digital trade. This is of great significance for reshaping many new trade patterns, extending the industrial chain and expanding the global marketing network. Recently, the state has implemented numerous national and local rules and regulations for the growth of the cross-border e-commerce industry, all of which are helpful to the quick expansion of this sector. The quick growth and increasing significance of e-commerce and cross-border e-commerce are two outcomes of the Internet's rise to mainstream popularity in the business world. However, there are still many obstacles and barriers to the growth of international trade. Several obstacles are tariffs & trade barriers, non-tariff barriers, trade restrictions & embargoes, lack of infrastructure, currency fluctuations & exchange rate risks, trade-related corruption, political instability, and so on. On the one hand, the once rapid rate of expansion in international trade has settled into a new normal of medium-to-low speed growth. However, as global economic development slows, trade tensions are rising and China's exports are subject to outside pressure to undergo a process of reform and upgrade. China's participation in multilateral trade organizations, such as the Asia-Pacific Economic Cooperation (APEC) and the Regional Comprehensive Economic Partnership (RCEP), has had several benefits for its trade relations. Some key advantages are market access, regional integration, tariff reductions, economic diplomacy, supply chain connectivity, and soft power projection. A significant development in international trade is the increasing prevalence of cross-border e-commerce through the use of overseas e-commerce platforms. This can effectively reduce the intermediate links and costs of traditional trade, and help enterprises expand overseas markets, which has become an inevitable choice for foreign trade transformation and upgrading [1-5].

*Corresponding Author: Lin Feng; E-mail: FengLin5678123@yeah.net
DOI: https://doi.org/10.70003/160792642025012601010

At present, domestic enterprises are facing opportunities of international market development and policy support. Behind the vigorous development, China's foreign trade enterprises are also facing pressures such as serious product homogeneity and continuous decline in export profit margins. As the backbone of China's import and export output value, cross-border businesses should take the opportunity to develop into new foreign markets as the economy recovers. Many established firms now have dedicated international trade divisions in an effort to increase market share and sales. Companies involved in international trade have adapted quickly to the rise of cross-border e-commerce by developing both traditional trade and cross-border e-commerce platforms, such as third-party websites. The rules of cross-border e-commerce are customs & import regulations, trade agreements & tariffs, consumer protection & data privacy, payment & financial regulations, intellectual property rights, and logistics & shipping. Companies who are already involved in international e-commerce are working to grow by offering a wider variety of products. Additionally, many organizations are introducing cross-border e-commerce after having transitioned from domestic trade or expanding into international trade. Organizations that are reshaping and growing their cross-border e-commerce will be among those struggling with the issue of product selection. Most international online retailers now focus on selling goods from the so-called hot list. Although the selection of products with reference to the hot-selling list has the advantages of high timeliness and low decision-making risk, there are also problems with cross-border enterprises that refer to hot-selling products for follow-up sales. While popular products are selling well, competition is fierce. Often because the seller's analysis and identification of hot-selling products are not in-depth and scientifically reasonable, they simply follow the sale. This has led to sellers entering the Red Sea market and intensified the homogenization of the market [6-8].

With the advent of the high-tech era, the Internet has also developed efficiently and rapidly, and the information network is affecting people's basic life at an alarming rate. Online platforms have become one of the indispensable consumption channels for users. At the same time, consuming users can express their opinions based on the goods or services they consume. Such a large number of review texts are generated on the Internet, and these review texts express the attitudes of consumers towards products. Platforms that generate comment texts include movie review platforms, medical platforms, e-commerce platforms, and catering platforms. The comment texts generated by these platforms contain the emotional colors and emotional tendencies of users towards products. The feedback of the sentiment of the review text is also a response to the quality of the goods or services. For cross-border e-commerce, users' evaluation of products or services has also become a way for them to know their products or services. This feedback information can let merchants understand the advantages and disadvantages of their products or services, and grasping the problems existing in products and services in real time can help merchants optimize them and improve their own quality. This can further facilitate the product selection strategy of cross-border e-commerce. Therefore, user evaluation information also plays a crucial role in the self-improvement and development of businesses [9-13].

2 Related Work

Reference [16] compared the growth rates of traditional trade and e-commerce, and found that rapid development for e-commerce created huge opportunities for SMEs to find and develop new customers in foreign markets. But this still faces many constraints, including differences in customs, tariff regimes, and tax laws. Literature [14] found that compared with offline trade, the trade cost associated with geographic distance in online transaction decreased significantly. But the cost of language barriers increases, the quality of legal institutions, online payment and logistics systems have an impact on cross-border trade. The impacts of online payment and logistics systems on cross-border e-commerce rates can have significant implications for businesses. Certain points are online payment systems, cost & transparency, fraud prevention & security, logistics & shipping systems, cross-border returns and customer support. Regulators can promote further development by improving the legal and financial system and improving parcel delivery infrastructure. Literature [15] shows that when enterprises compete in global platform e-commerce, they are hindered by problems such as the unsuitability of goods and services for sales, logistics, payment, laws and regulations. In order to facilitate the growth of cross-border e-commerce by small and medium-sized businesses, the government should adopt measures such as giving educational aid, expanding technological infrastructure, and revising laws and regulations. With regards to the current state of international online trade in my country, most academic studies agree that it is undergoing a time of rapid development. It focuses on the current state and future projections of the industry, as well as the issues and solutions surrounding the brand, logistics, payment, and other areas. The potential of international online trade in the Belt and Road region was assessed in reference [16]. Recommendations for fixing the problem are put out, including strengthening e-commerce regulations, expanding opportunities for international trade, and enhancing the information infrastructure upon which businesses rely. From the point of view of logistical cooperation, literature [17] has pointed out that there are issues in cross-border e-commerce in nations and areas along the Belt and Road. That means we need to get on speeding up the development of overseas warehouses for cross-border e-commerce logistics, enhancing the customs clearance environment, and constructing a cross-border e-commerce logistics data platform. In the future, cross-border e-commerce policy is expected to prioritize the regulation of online foreign trade and the coordination of online and offline foreign trade, according to the literature [18]. E-commerce platforms and businesses operating across international borders will increase the use of overseas warehouses, and e-commerce businesses operating across international borders will prioritize the development of their own brands. Commodity quality, cross-border electronic payment, a foreign trade comprehensive service platform, and government supervision are only a few of the issues that have been raised as barriers to the growth of cross-border e-commerce in the published literature [19]. To thrive in the new economic normal, cross-border e-commerce must have a conducive environment for growth, and the supremacy of international rule-making must be sought. Reference [20] predicts five major trends in development of cross-border e-commerce: rapid growth, vigorous development of emerging markets, capitalization and branding, localization, and digitization. This also puts forward corresponding strategies on how to promote the transformation and upgrading of cross-border exports.

Literature [21] believes that category management can not only meet the existing needs of consumers, but also promote consumers to form new consumption preferences. It also explores the impact of category management of alternatives and merchandising on consumer purchasing decisions. It proposes that category management mainly affects the decision-making process of consumers from the perspective of simplifying consumers' purchase judgment and simplifying the process of consumers' collection and processing of commodity information. Reference [22] explores the cooperation model between suppliers and retailers from the perspective of category management. This is divided into different areas through modular technology, and then according to different areas, hybrid strategies, seller-led category management, and category performance-based cooperation methods are proposed. Reference [23] takes agricultural products as an example, and proposes to do a good job in the classification of agricultural products on the basis of continuously improving the logistics and distribution system. This requires building a supply chain driven by consumer demand and driven by category management. Carry out e-commerce through the category synergy model and regional market model, and promote the development of agricultural product e-commerce. Literature [24] believes that only by comprehensively analyzing the advantages and disadvantages of online sales, the characteristics of online consumers, and the marketing situation of commodity attributes can e-commerce obtain appropriate commodity selection. Literature [25] believes that cross-border e-commerce product selection strategies should follow five principles, namely platform orientation, emphasis on intellectual property rights, decision-making based on data, moderate prices, and selection of appropriate channels. It also proposes a method for selecting products by analyzing data, channels and markets, and finally proposes a product selection strategy for price, category, and market segmentation. Literature [26] believes that enterprises should pay attention to target customer groups of this category and the price and quality of existing competitors in the market according to their own product lines and the competitive advantages of each type of product. This also puts forward countermeasures and suggestions for maintaining profitability, selecting products with low after-sales maintenance costs, and creating popular products. Reference [27] proposes a calculation formula for judging the dominant relationship between two commodities based on a commodity selection method based on online evaluation information and consumer expectations, combined with consumers' expectations for attribute evaluation. This provides a new idea for solving the problem of commodity selection based on online evaluation information in reality. Reference [28] discusses four techniques for cross-border e-commerce sellers to select products. First, to discover popular categories by analyzing the data of cross-border e-commerce platforms. Second, fully tap the potential demand and develop the blue ocean market. Third, attach importance to cross-border e-commerce rules and avoid infringement on trademarks, appearances, and invention patents. Fourth, increase profits through the advantages of combination products such as diverting products and increasing MOQ. The potential advantages are inventory management, cost savings, reduced logistics complexity, market demand optimization, enhanced cross-selling opportunities and market differentiation.

3 Research Method

This work proposes CNN-BiGRU-ATT for sentiment polarity analysis of product reviews. First, the multi-channel convolutional neural network is utilized to extract local feature of different granularities. Then connect BiGRU for context sequence learning to memorize long-distance dependency information. This method not only solves the problem that CNN cannot perform context sequence learning, but also solves issue of gradient disappearance or explosion. Finally, attention is introduced after the BiGRU hidden layer to filter the important features. This allows model to focus more on words that are more important for classification, improving classification efficiency.

3.1 CNN Algorithm

Convolutional neural networks, which are deep feedforward networks with the properties of local connection and weight sharing, found their initial use in the field of imaging. The key properties are local connection, receptive field, weight sharing, translation invariance, parameter efficiency, and hierarchical feature extraction. These properties make CNNs particularly well-suited for tasks involving images, videos, and other forms of structured data. Convolutional layers, pooling layers, fully connected layers, and various activation functions make up the common linked model structures of convolutional neural networks. The directional propagation algorithm is commonly used to maximize the model-training process. Backpropagation is an efficient method for computing the gradients of the model parameters with respect to a loss function, enabling the optimization of neural networks through gradient-based methods. Forward pass, loss calculation, backward pass, gradient calculation, and parameter update. It's been estimated that the number of parameters in a convolutional neural network is several times smaller than in a standard feedforward neural network. The number of parameters in a Convolutional Neural Network (CNN) is typically much smaller compared to a standard feedforward neural network, especially when dealing with grid-like input data, such as images. The reasons are parameter sharing, spatial invariance, local connectivity, and input structure exploitation. This is mainly due to the network structure design with a series of operational characteristics such as weight sharing, local connection and pooling. The local connection in a CNN differs from connections in other neural network architectures by exploiting the local receptive fields, parameter sharing, spatial invariance, and hierarchical feature extraction. These properties make CNNs particularly effective in tasks involving grid-like input data, such as images, where local patterns and spatial relationships play a crucial role. Pooling helps achieve spatial invariance by reducing the sensitivity of the network to small spatial translations or distortions in the input. It also reduces the computational requirements by downsampling the feature maps, enabling the network to focus on the most salient features. Weight sharing enables the network to learn spatially invariant features that are effective across the entire input space. It promotes the extraction of translation-invariant features by using the same filters at different locations.

On a small local feature map, a local connection can be thought of as a network in which each neuron is only connected to one other neuron in its own local window in the next layer. In the weight sharing feature in convolutional neural networks, the number of network parameters is not directly related to the number of neurons. Few reasons why network parameters do not scale linearly with the number of neurons are weight sharing, local connections, pooling operations, and network architecture. The number of parameters is determined by the specific design choices and optimization techniques employed in the network architecture, rather than being directly proportional to the number of neurons. The significance of weight sharing is parameter efficiency, generalization, translation invariance, capturing local patterns, and reducing memory requirements. By reducing the number of network parameters, weight sharing in CNNs contributes to more efficient and effective model training and inference. The parameters on the same convolution kernel are the same.

The function of pooling layer is to further filter input information of the previous layer, which reduces the number of parameters and reduces the dimension of the features. Two types of pooling functions are commonly used in models. The first is maximum pooling, which is usually calculated to obtain the maximum value of all results in a local range. The maximum pooling helps in reducing the number of neurons is spatial downsampling, reduction in spatial dimensions, and neuron reduction. By downsampling and reducing the spatial dimensions, maximum pooling helps in extracting the most salient features while discarding less important or redundant information. Average pooling is usually calculated to obtain the average of all results in a local range. The purpose of using average pooling in Convolutional Neural Networks (CNNs) is to downsampled the input data while preserving the general information contained within each pooling region. Unlike maximum pooling, which selects the maximum value within each region, average pooling calculates the average value. The max pooling can effectively reduce neurons in the network. Moreover, ingenious design of structure can maintain the invariance to the local morphological changes of small features, so that it has a wider perception field.

FC layer is generally designed in the last layer of network, its role is to change dimension and learn the rich features of the data. By applying the sigmoid activation function in the FC layer, the network can efficiently transform the combined features into probabilities or scores for different classes. The sigmoid function maps the input values to a bounded range, allowing the network to provide class probabilities that can be interpreted as the likelihood of a given class. However, its parameters are huge, and sometimes in order to avoid its overfitting, a Dropout mechanism is added to alleviate it.

3.2 BiGRU Algorithm

GRU is a superior deep network model than Long Short-Term Memory, or LSTM. The most significant benefit of the recurrent neural network is alleviated with the use of LSTM network. The reasons of how LSTM alleviate the problem are memory cell, gates, forget gate, input gate, update, update and forget, output gate, etc. The main advantages of LSTM network are long-term dependency, vanishing gradient problem, exploding gradient problem, handling variable-length sequences, and handling multiple time scales. Their specialized memory cells and gating mechanisms make them powerful tools for modeling and processing sequential data, leading to improved performance in a wide range of tasks involving temporal information. The LSTM network model is more involved, but it also has drawbacks including slow prediction speeds and lengthy training periods. In order to address these issues with LSTM networks, GRU networks are enhanced by building on the foundation of LSTM networks. The main issues with LSTMs include complexity, difficulty in training, memory cell redundancy, computation and memory requirements. GRUs aim to address the aforementioned issues by offering a more streamlined and computationally efficient design are simplicity, reduced memory cell, fewer gating units, and computational efficiency. While LSTM has been widely used and well-established, GRU offers some advantages that make it a compelling alternative in certain scenarios. Some reasons why GRU is considered a better deep network model than LSTM in certain cases is simplicity & computational efficiency, reduced risk of overfitting, enhanced training dynamics, comparative performance, and model interpretability. The state of an LSTM network cell is determined by whether or not its input information is discarded by the gate structure of the network. By combining the LSTM network's forget gate and input gate into a single update gate, the GRU network improves upon the architecture of the LSTM network's gate structure. Both LSTM and GRU are types of recurrent neural networks (RNNs) designed to address the vanishing gradient problem and capture long-term dependencies in sequential data. The gate structure in LSTM plays a crucial role in achieving these improvements. The significances of LSTM gate are forgetting gate, input gate, update memory cell, and output gate. The cell is fused, the number of gates is reduced from three to two, and additional enhancements are made simultaneously. GRU is demonstrated in Figure 1.

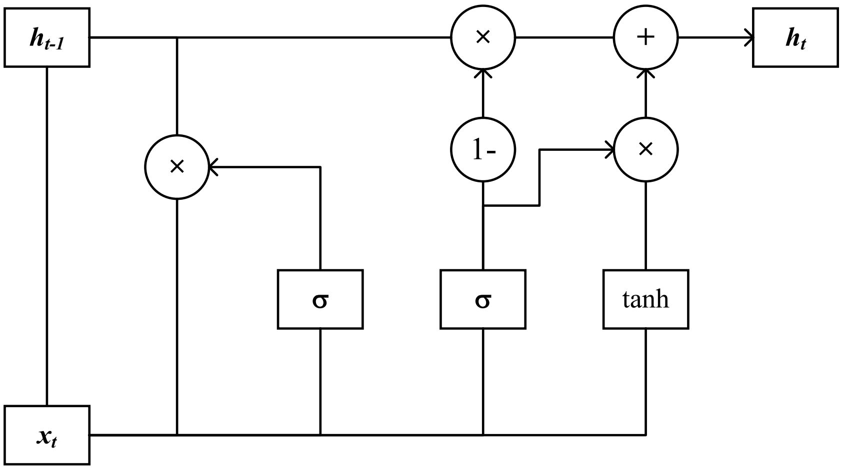

Figure 1. GRU structure

The reset gate regulates the effect of the previous time node's hidden layer's output on the current time node's hidden layer. The update gate in an LSTM (Long Short-Term Memory) network evaluates the impact of the input state on the hidden layer state by controlling how much of the new information should be incorporated into the existing memory. It determines the balance between retaining the previous memory and updating it with the new information. The state of the previous time node's hidden layer is ignored if the reset gate is close to 0, indicating that the output of the previous time node's hidden layer is not relevant to the cell state of the present time node. Therefore, reset gates are designed to discard past hidden layer states that are irrelevant to the future.

$$r_t = \sigma\left(w_r * \left[h_{t-1}, x_t\right]\right) \quad \textbf{(1)}$$

The update gate is responsible for determining whether the current time node's input state will be disregarded. The update gate of an LSTM network is able to evaluate the impact of the input state of the current time node on the hidden layer state, just as the input gate does. The state of the hidden layer of the previous node will be written into the current node's cell if the update gate is near 1 all the time. Therefore, update gates are designed to solve the problem of gradient decay in recurrent neural networks while better capturing widely spaced dependencies in timing information. When the gradients vanish, the network's ability to learn long-term dependencies is compromised. The problem is particularly pronounced in RNNs because the gradients need to flow through multiple time steps. If the gradients diminish exponentially as they propagate backward, the influence of earlier time steps on the current time step diminishes significantly. Gradient explosion can occur when the gradients are multiplied repeatedly during backpropagation, causing their magnitudes to increase exponentially.

$$z_t = \sigma\left(w_z * \left[h_{t-1}, x_t\right]\right) \quad \textbf{(2)}$$

The GRU network is a refined version of the LSTM network, hence there are parallels between the BiLSTM and BiGRU models. By incorporating BiGRUs into context sequence learning, models can overcome the limitations of CNNs in capturing long-range dependencies, leveraging bidirectional information flow, enhancing representation learning, capturing contextual embeddings, and handling variable-length sequences. These improvements make BiGRUs a powerful tool for tasks that involve understanding and processing sequential data. The BiGRU network's central concept is similar to that of proposing two GRU networks, one forward and one backward, for each training sequence. The BiGRU model leverages the forward and backward GRU networks to enhance understanding are capturing past and future context, concatenating hidden states, enhanced contextual understanding, and contextual bidirectional interactions. Moreover, the two networks share a common pathway to the final layer of the data structure. For each node in the input sequence's timeline, this structure conveys to the output layer comprehensive historical and prospective context. Six distinct weights are required for the BiGRU network, and these weights are recycled between time steps. This should also know that the forward and backward concealed layers do not communicate with one another in any way. Thus, at each node in time, the BiGRU network is guaranteed to be acyclic.

3.3 Attention Algorithm

The essence of attention is to only focus on the part you want to focus on, and weaken the part that does not help the result. This greatly reduces the amount of calculation and effectively strengthens the target information, thereby ignoring unimportant information. In this paper, the attention mechanism is introduced to obtain emotional features of the evaluation data, focusing on the key part of the information, thus enhancing the accuracy of emotional prediction. Attention is introduced in the neural network machine translation via an encoder-decoder framework. The encoder-decoder architecture combined with attention allows the model to focus on different parts of the source sentence while generating the target translation. The working principles are encoder, attention mechanism, decoder, attention calculation, and contextual alignment. After introducing attention, the translation process will focus on different words, and the model will give the most concentrated attention to the words currently being translated. This dynamic pattern of continuously adjusting the input focus as the task progresses is the flow of the attention.

Encoding is to map the text into feature vectors, and decoding is to convert the feature vectors into a specific sequence of sentences. Encoder-Decoder is considered a widely used framework. The Encode-Decoder framework often used in natural language processing is demonstrated in Figure 2.

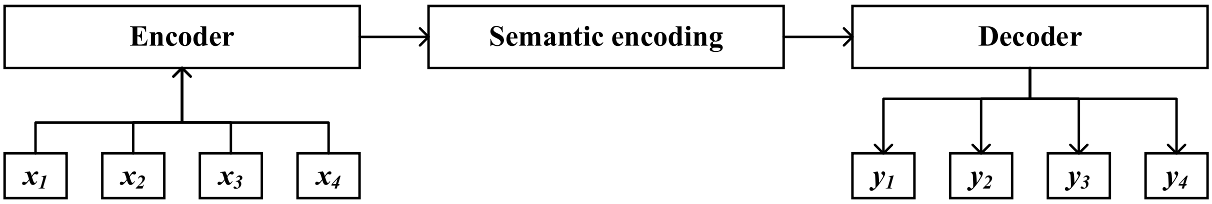

Figure 2. Schematic diagram of Encoder-Decoder framework

First, the sequence is input to the encoding module for encoding, and the encoded semantic representation is obtained.

$$C = F(X) \quad \textbf{(3)}$$

Then, the decoding module decodes C. Sequence output depends on C and previous output

$$y_i = G\left(C, y_1, ..., y_{i-1}\right) \quad \textbf{(4)}$$

The encoder in the Encoder-Decoder model evenly compresses the input into a continuous space representation with a fixed number of dimensions, which is manifestly unrealistic given the variability in the size and amount of input information. Due to the attention process, not all input vectors contribute equally to the final result.

To represent the input in an attention-based Encoder-Decoder model, the encoder can use a series of vectors, each of which must convey a predetermined amount of information about the input's focus. Decoding is a multi-step process, and the decoder network can choose use encoded inputs at each stage. By computing the context vector at each time step, this decoding removes the need for the encoder to flatten inputs of varying lengths into a single fixed-dimensional vector, thereby revealing a representation of the input sequence that is itself varying in size.

The core mechanism of the attention mechanism is demonstrated in Figure 3.

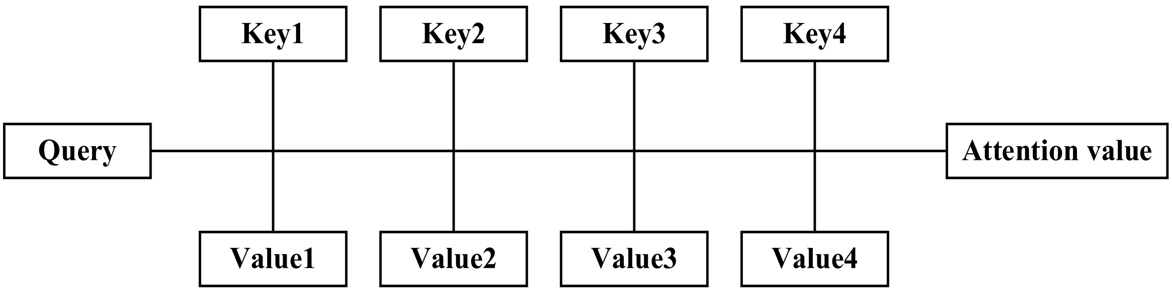

Figure 3. Structure diagram of attention mechanism

3.4 CNN-BiGRU-ATT Architecture

From the principles of CNN and RNN, the convolutional neural network obtains the local feature information of the text through convolution and pooling operations, but cannot learn the context-dependent information of the text. The CNN-BiGRU-ATT network utilizes the BiGRU layer effectively to extract data features. By leveraging the strengths of both CNNs and RNNs, the network captures sequential dependencies, improves contextual understanding, performs dynamic feature extraction, and incorporates attention mechanisms. The recurrent neural network is used for sequence learning, memorizing text long-distance dependent information, but there may be problems such as gradient explosion. The model resolves the issues are long-term dependencies, bidirectional GRU, gradient flow & information, attention mechanism, and convolutional layers. It enables more effective sentiment polarity analysis by considering both past and future contextual information while mitigating the issues associated with gradient disappearance and explosion. Therefore, CNN-BiGRU-ATT model is constructed using their respective characteristics, as demonstrated in Figure 4.

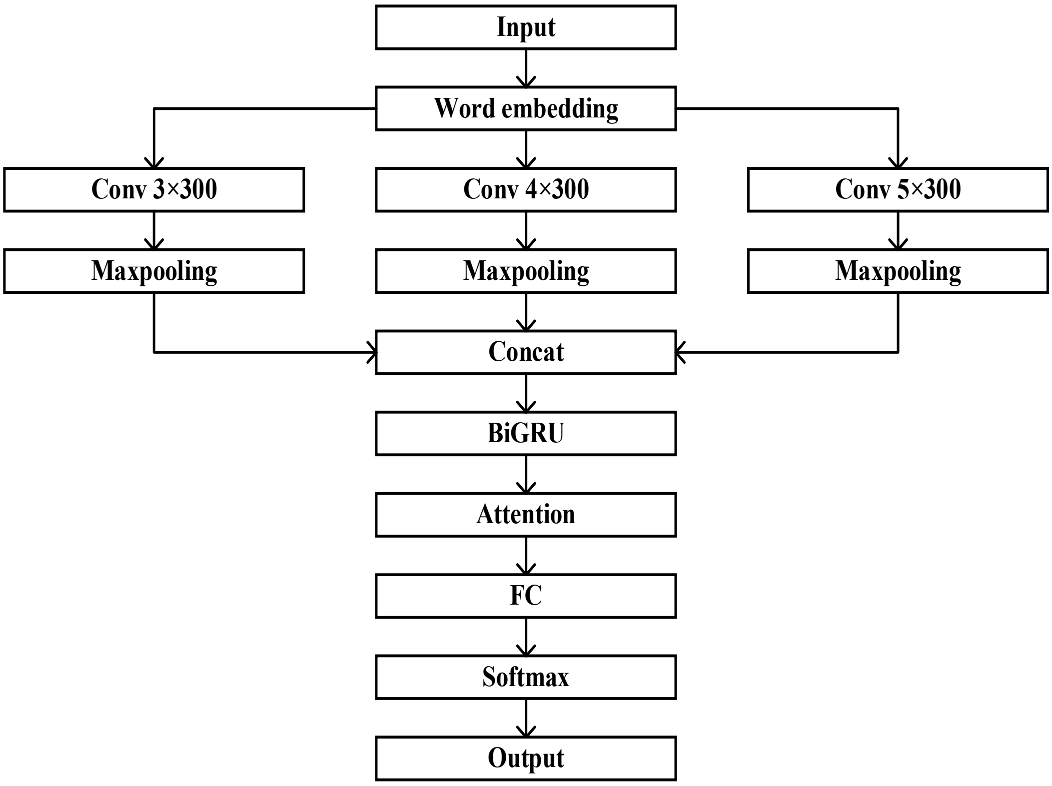

Figure 4. CNN-BiGRU-ATT pipeline

Input layer inputs words obtained after word segmentation of each comment text. The word embedding layer plays a crucial role in representing words as dense, continuous vectors, capturing their semantic relationships. These word vectors help the neural network understand and learn from the input text more effectively, as compared to using one-hot encoded representations of words. The function of word embedding layer is to convert input text into a sequence of word vectors, and the word vector dimension is 300. Convert all sample data into word embedding matrix, which is often done in batch mode in the actual model training process. That is, samples of batch size are taken as input each time to speed up training

The convolution layer extracts the local feature information through convolution operation. The dimension of word vector is set to 300. In order to extract different n-gram features, three kinds of convolution kernels are set, with the sizes of 3×300, 4×300, and 5×300, and each kind of convolution kernel is 128.

Pooling is to down sample the local features of sentences obtained after convolution to obtain more salient features. Here, the maximum pooling is selected, that is, the maximum value in the vector after the convolution operation is selected as the local feature. By choosing the maximum value in the vector after the convolution operation, max pooling promotes spatial invariance, robustness to noise, translation invariance, and dimensionality reduction. It allows the network to focus on the most significant and discriminative features, making CNNs more efficient and effective in tasks such as image classification, object detection, and feature extraction.

The primary function of the BiGRU layer is to refine the feature vectors generated by the CNN network by additional feature extraction. The BiGRU layer serves as a sequential modeling component that refines and enriches the feature vectors generated by the preceding CNN layers. By incorporating bidirectional information flow and capturing sequential dependencies, the BiGRU layer enhances the network's contextual understanding and feature representation, enabling it to make more informed and accurate predictions in tasks involving sequential data. Maximum pooling contributes to feature extraction in CNNs by downsampling the input, achieving translation invariance, enhancing robustness to local variations, reducing dimensionality, promoting feature invariance, and facilitating hierarchical feature extraction. These contributions help CNNs capture and represent important features in the input data, making them effective in various computer vision tasks such as image classification, object detection, and semantic segmentation. The BiGRU network architecture takes the sequence of inputs from both the front and back to feed into the GRU model. Both forward and reverse GRU have their hidden layer sizes set to 64. The final output of the BiGRU layer is obtained by weighting and splicing the outputs of the two hidden layers.

The attention mechanism layer adopts the forward attention to extract important word information. The purpose of utilizing the attention mechanism layer in neural networks is to enhance the model's ability to focus on relevant parts of the input or context while generating an output. The attention mechanism helps address the limitations of fixed-length representations by allowing the model to selectively attend to different parts of the input dynamically. Attention assigns different weights to the semantic encoding of text vectors. This can achieve purpose of distinguishing the importance of information, thereby improving classification accuracy.

$$v = \tanh(wh + b) \quad \textbf{(5)}$$

$$p = soft\max(v_t, v_w) \quad \textbf{(6)}$$

$$a = \sum ph \quad \textbf{(7)}$$

After the effective features of the text are obtained through the attention mechanism, a dropout layer is added to prevent the model from overfitting. Finally, add the FC layer and use the sigmoid function activation to classify the text features.

4 Experiment

4.1 Experimental Details

This work collects the corresponding review data based on the crawler technology to form the training set and test set required by the CNN-BiGRU-ATT network. To train and test a CNN-BiGRU-ATT network, one approach to collecting review data could involve obtaining a dataset from publicly available sources or through collaborations with organizations that curate and provide access to labeled review data. A general outline of how crawler technology could potentially be used as identifying relevant data sources, defining the data collection strategy, data extraction & processing, splitting into training & test sets. The process of using crawler technology to form training and test sets involves careful consideration of legal and ethical aspects, data quality control, and adherence to best practices for web scraping. This work is via deep learning framework, and experimental environment used is demonstrated in Table 1. In addition, this work requires the setting of many parameters for the experiment, and the specific settings are demonstrated in Table 2.

Table 1. Experimental environment data

Experimental environment	Configuration data
Language	Python3.6
Tool	Pycharm
Frame	Tensorflow1.17.0
Processor	NVIDIA GeForce 940MX
Operating system	Windows10

Table 2. Model hyperparameter setting

Hyperparameter	Set value
Word vector dimension	300
BiGRU hidden layer	64
Optimizer	Adam
Learning rate	0.001
Dropout	0.2

4.2 Comparison with Different Methods

To verify the reliability of the CNN-BiGRU-ATT network for product review analysis, this work compares it with other review analysis methods. Some review analysis methods are bag-of-words (BoW), support vector machines, recurrent neural networks, convolutional neural networks, attention-based models, transformer-based models. The experimental data are demonstrated in Table 3.

Table 3. Comparison with different methods

Method	Acc	F1
CNN	90.7	88.3
LSTM	91.8	89.7
GRU	93.2	90.6
CNN-BiGRU-ATT	95.9	93.1

Compared with other methods, CNN-BiGRU-ATT can achieve the highest accuracy and F1 score. This validates the reliability of CNN-BiGRU-ATT for product review analysis.

4.3 Evaluation on Single and Multi-Scale Feature

The CNN module in the CNN-BiGRU-ATT network uses multi-scale convolutional features. To verify that this measure can improve performance, this work compares the network performance corresponding to multi-scale features and single-scale features. One significant advantage of using multi-scale convolutional features compared to single-scale features in convolutional neural networks (CNNs) is the ability to capture information at different levels of detail or granularity. Some advantages are hierarchical representation, robustness to scale variations, local and global context, Richer feature representation, and improved translation invariance. The experimental data is demonstrated in Figure 5.

After using multi-scale features, the CNN-BiGRU-ATT network can achieve improvements in both accuracy and F1 score. This verifies the superiority of CNN-BiGRU-ATT using multi-scale features.

4.4 Evaluation On BiLSTM and BiGRU

CNN-BiGRU-ATT network uses BiGRU to extract date feature. To verify that this measure can improve performance, this work compares the network performance corresponding to BiGRU and BiLSTM. The primary differences between BiLSTM and BiGRU models lie in their underlying RNN architectures and the complexity of memory management. BiLSTM tends to have a more complex architecture with separate memory cells and three gates, allowing them to capture long-term dependencies effectively. The experimental data is demonstrated in Figure 6.

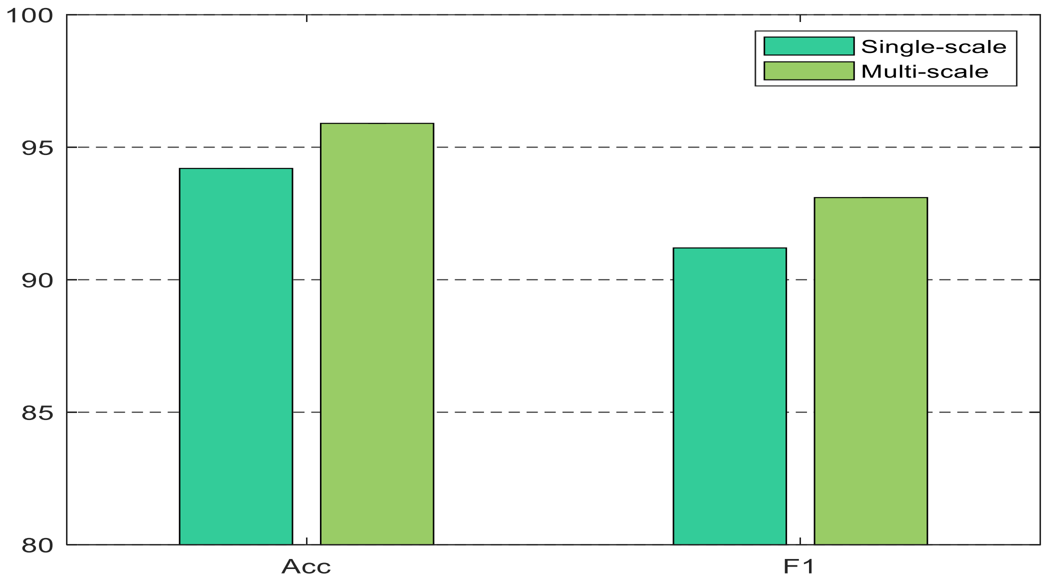

Figure 5. Evaluation on single and multi-scale feature

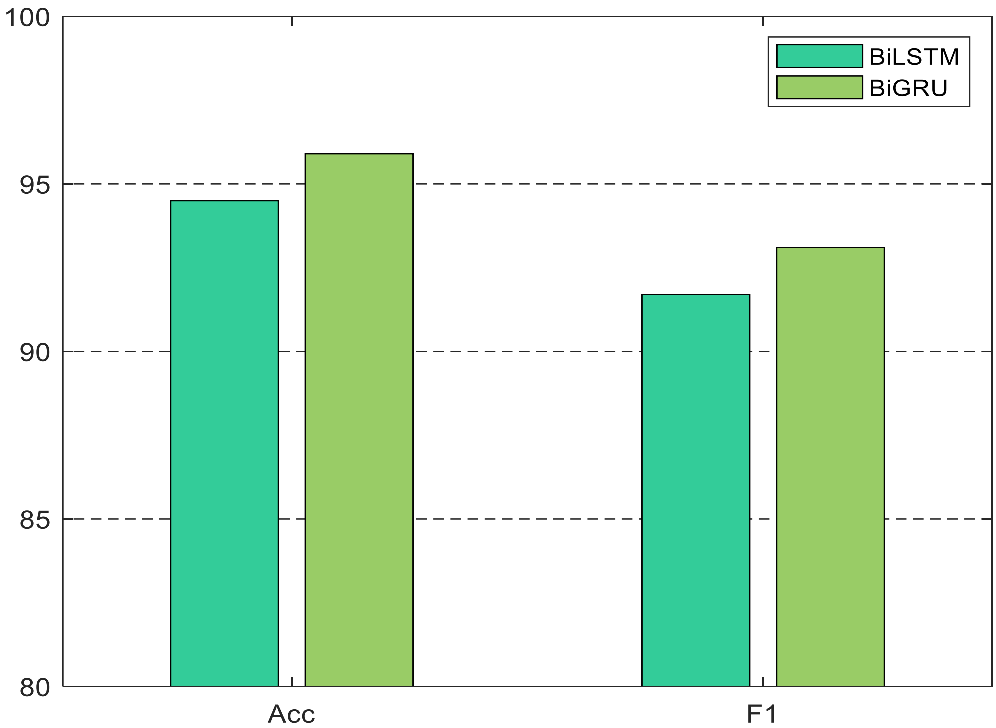

Figure 6. Evaluation on BiLSTM and BiGRU

In addition, the hidden layer of BiGRU can be manually set. To verify impact of different hidden layers on performance, this work compares the correct rate and F1 score corresponding to different BiGRU hidden layers. The experimental data are demonstrated in Table 4.

Table 4. Evaluation on BiGRU hidden layer

Layer	16	32	64	96	128
Acc	93.7	94.3	95.9	95.2	94.6
F1	91.3	91.8	93.1	92.7	92.3

With the superposition of BiGRU hidden layers, the model performance first increases and then decreases. When the number of hidden layers is set to 64, CNN-BiGRU-ATT can achieve the highest performance.

4.5 Evaluation on With and Without Attention

CNN-BiGRU-ATT network uses attention to enhance date feature. To verify that this measure can improve performance, this work compares the network performance corresponding to using attention and no attention. The experimental data is demonstrated in Figure 7.

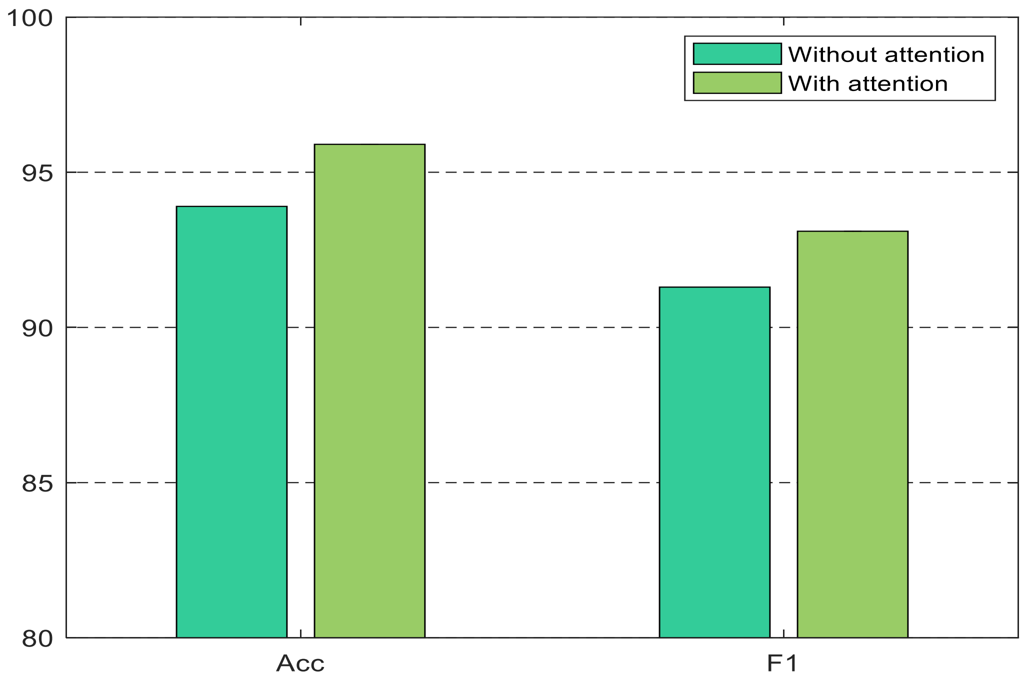

Figure 7. Evaluation on With and Without attention

After using attention, the CNN-BiGRU-ATT network can achieve improvements in both accuracy and F1 score. This verifies the superiority of CNN-BiGRU-ATT using attention.

5 Conclusion

At present, most cross-border e-commerce companies only blindly follow method of selling popular products when conducting market layout. Obviously, this method has greater risks to the business operation. However, there is currently no good way for cross-border e-commerce companies to avoid this risk. This is mainly because these companies lack a set of strategies to drive product selection with data mining. With artificial intelligence, more and more Internet users have become the producers of Internet information. Various e-commerce product review texts on the Internet have potential huge value for cross-border e-commerce. Mining this information and performing sentiment analysis on it is of great research significance. This work proposes CNN-BiGRU-ATT for sentiment polarity analysis of product reviews. First, the multi-channel convolutional neural network is utilized to extract local feature of different granularities. Then connect BiGRU for context sequence learning to memorize long-distance dependency information. This method not only solves the problem that CNN cannot perform context sequence learning, but also solves issue of gradient disappearance or explosion. Finally, attention is introduced after the BiGRU hidden layer to filter the important features. This allows the model to focus more on words that are more important for text classification, improving classification efficiency. In this work, many experiments are conducted to verify feasibility for this method.

Acknowledgements

Research and Development Program of Ankang Science and Technology:

Research on the Demand and Training Path Mechanism of Cross-border E-commerce Talents in Ankang under the Background of “Belt and Road” Initiative

Project Number: AK2023-RK09

References

[1] X. Liu, Z. Dou, W. Yang, Research on influencing factors of cross border E-commerce supply chain resilience based on integrated fuzzy, DEMATEL-ISM, *IEEE Access*, Vol. 9, pp. 36140-36153, February, 2021.

[2] H. Taherdoost, M. Madanchian, Empirical modeling of customer satisfaction for E-services in cross-border E-commerce, *Electronics*, Vol. 10, No. 13, Article No. 1547, June, 2021.

[3] X. Lin, J. Wu, A. K. Bashir, J. Li, W. Yang, M. J. Piran, Blockchain-based incentive energy-knowledge trading in IoT: Joint power transfer and AI design, *IEEE Internet of Things Journal*, Vol. 9, No. 16, pp. 14685-14698, August, 2022.

[4] N. Chen, Y. Yang, The impact of customer experience on consumer purchase intention in cross-border E-commerce-Taking network structural embeddedness as mediator variable, *Journal of Retailing and Consumer Services*, Vol. 59, Article No. 102344, March, 2021.

[5] Y. He, R. Wu, Y. J. Choi, International logistics and cross-border E-commerce trade: who matters whom, *Sustainability*, Vol. 13, No. 4, Article No. 1745, February, 2021.

[6] X. Wang, J. Xie, Z. P. Fan, B2C cross-border E-commerce logistics mode selection considering product returns, *International Journal of Production Research*, Vol. 59, No. 13, pp. 3841-3860, 2021.

[7] G. Raja, A. Ganapathisubramaniyan, S. Anbalagan, S. B. M. Baskaran, K. Raja, A. K. Bashir, Intelligent reward-based data offloading in next-generation vehicular networks, *IEEE Internet of Things Journal*, Vol. 7, No. 5, pp. 3747-3758, May, 2020.

[8] Y. Wang, F. Jia, T. Schoenherr, Y. Gong, L. Chen, Cross-border e-commerce firms as supply chain integrators: The management of three flows, *Industrial Marketing Management*, Vol. 89, pp. 72-88, August, 2020.

[9] Z. Guo, Y. Shen, A. K. Bashir, K. Yu, J. C. W. Lin, Graph embedding-based intelligent industrial decision for complex sewage treatment processes, *International Journal of Intelligent Systems*, pp. 1-19, June, 2021.

[10] S. Ma, Q. Liang, Industry competition, life cycle and export performance of China's cross-border e-commerce enterprises, *International Journal of Technology Management*, Vol. 87, No. 2-4, pp. 171-204, 2021.

[11] S. Chen, Q. He, H. Xiao. A study on cross-border e-commerce partner selection in B2B mode, *Electronic commerce research*, Vol. 22, No. 2, pp. 1-21, June, 2020.

[12] J. Mou, W. Zhu, M. Benyoucef, Impact of product description and involvement on purchase intention in cross-border e-commerce, *Industrial Management & Data Systems*, Vol. 120, No. 3, pp. 567-586, March, 2020.

[13] G. H. Gessner, C. R. Snodgrass, Designing e-commerce cross-border distribution networks for small and medium-size enterprises incorporating Canadian and U.S. trade incentive programs, *Research in Transportation Business & Management*, Vol. 16, pp. 84-94, September, 2015.

[14] E. Gomez-Herrera, B. Martens, G. Turlea, *The drivers and impediments for online cross-border trade in goods in the EU*, Institute for Prospective Technological Studies Digital Economy Working Paper No. 2012/01, pp. 1-38, September, 2012.

[15] M. Savrul, A. Incekara, S. Sener, The potential of e-commerce for SMEs in a globalizing business environment, *Procedia-Social and Behavioral Sciences*, Vol. 150, pp. 35-45, September, 2014.

[16] N. Zhu, L. Wu, Research on the Present Situation and Trend of Chinese Cross-border E-Commerce Under "One Belt One Road", *Reformation & Strategy,* Vol. 31, No. 12, pp. 134-137, December, 2015.

[17] X. Liu, B. Zhang, The Study of Collaborative Development of Cross-border E-commerce Logistics in Countries along the Belt and Road, *China Business and Market*, Vol. 30, No. 5, pp. 115-120, May, 2016.

[18] J. Wen, Z. Yan, Development of Cross-border E-commerce in China: New Features, New Problems and New Trends, *International Economy and Cooperation*, Vol. 11, pp. 29-35, 2017.

[19] Y. Zhang, Analysis on the development of cross-border e-commerce under the background of the new normal economy, *Journal of Commercial Economics*, No. 7, pp. 69-72, 2018.

[20] W. Xue, D. Li, Y. Pei, The Development and Current of Cross-border E-commerce, *WHICEB 2016 proceedings*, Vol. 53, pp. 131-138, July, 2016.

[21] B. Dewsnap, C. Hart, Category management: a new approach for fashion marketing, *European Journal of Marketing*, Vol. 38, No. 7, pp. 809-834, July, 2004.

[22] S. Han, Y. Ye, X. Fu, Z. Chen, Category role aided market segmentation approach to convenience store chain category management, *Decision Support Systems*, Vol. 57, pp. 296-308, January, 2014.

[23] C. Yu, C. Qi, Research on the Complementarity and Comparative Advantages of Agricultural Product Trade between China and CEE Countries—Taking Poland, Romania, Czech Republic, Lithuania and Bulgaria as Examples, *Journal of Service Science and Management*, Vol. 8, No. 2, pp. 201-208, April, 2015.

[24] S. Schwarzl, M. Grabowska, Online marketing strategies: the future is here, *Journal of International Studies*, Vol. 8, No. 2, pp. 187-196, 2015.

[25] W. Su, Y. Wang, L. Qian, S. Zeng, T. Balezentis, D. Streimikiene, Creating a sustainable policy framework for cross-border e-commerce in China, *Sustainability*, Vol. 11, No. 4, Article No. 943, February, 2019.

[26] H. Zhao, Research on the development of China's cross-border E-commerce, *Management & Engineering*, Vol. 28, pp. 79-84, September, 2017. https://www.proquest.com/openview/5b289a300a242fcf46c14880dae94d1c/1?pq-origsite=gscholar&cbl=2028702

[27] M. Y. Li, X. J. Zhao, L. Zhang, X, Ye, B. Li, Method for product selection considering consumer's expectations and online reviews, *Kybernetes*, Vol. 50, No. 9, pp. 2488-2520, August, 2021.

[28] J. Mou, G. Ren, C. Qin, K. Kurcz, Understanding the topics of export cross-border e-commerce consumers feedback: an LDA approach, *Electronic Commerce Research*, Vol. 19, No. 4, pp. 749-777, December, 2019.

Biography

Lin Feng received her BA degree in English Language from Northwest University, and got MA degree in Business Administration from Xi'an University of Technology. She is the author of more than 10 journal papers and has written two book chapters. Her current research interests include ELT, business English, and Cross-border E-commerce research. She has certificated with FTBE from LCCIIQ.