1. Introduction

This has stimulated offline entities to a certain extent and has formed an integrated online and offline sales strategy. With the increasing scale of online business platforms, recommendation systems have also risen [1] and have become one of the backbone technologies in various fields today. It is also mainly from the perspective of “people-oriented” to make the originally large and complex assembly line more intelligent, which greatly facilitates human beings. Therefore, it is favored by many scholars at home and abroad, and more and more researchers invest in this field [2].

Nowadays, traditional foreign trade is in a state of insufficient growth and continued sluggishness. However, cross-border e-commerce has sprung up and has experienced a “blowout” growth, which has become an important driving force for industrial upgrading and new growth in foreign trade under the new normal of the economy [3].

Banks and other financial institutions jointly established a series of norms and mechanisms based on the collection, processing, credit rating, credit supervision, and rewards and punishments related to credit information, thereby effectively enhancing the trust relationship and creditworthiness between transaction entities. Optimize the credit environment of cross-border e-commerce. The system’s data analysis center converts it into indicators in some way, obtains the credit rating of the transaction subject through system analysis, and provides a basis for system supervision, rewards, and punishments.

This study applies deep learning to the cross-border e-commerce intelligent information recommendation system to overcome the geographical problems of cross-border e-commerce and improve the recommendation effect of cross-border e-commerce.

2. Related Work

The rule-based recommendation algorithm is mainly based on customer-defined rules for information association and information mining and then recommends and analyzes customers through existing rules. The advantage of this algorithm is simple and fast [4].

The recommendation based on content filtering is to make recommendations by comparing the customer profile with existing resources. The key consideration of this recommended technique is the calculation of similarity. Because most customers are unable to accurately analyze and accurately express their needs, therefore, the rule-based
recommendation algorithm has big flaws. The content-based recommendation technology compares the customer’s feature description file with the resource file and searches for similar word features. Typical systems include Person- alWebWatcher [5], WebMate [6], and WebACE [7].

The recommendation algorithms of collaborative filtering technology include user-based collaborative filtering. Among them, user-based collaborative filtering is realized by analyzing the similarity of user characteristics and behaviors. The item-based collaborative filtering method analyzes the similarity relationship between items and then provides users with similar items. The model-based collaborative filtering is based on recommendations which are based on the relationship between existing item features and user features [8].

Hybrid filtration technology: since the content based have their own advantages and disadvantages, in many practical applications, these two different recommendation algorithms are combined for recommendation, which is a hybrid recommendation algorithm [9]. Huang et al. [10] use the Bayesian probability framework to combine content-based features and collaborative filtering features for recommendation calculations and achieve better results. Liang and Qin [11] combine the two recommendation techniques in a weighted form, thereby increasing the diversity of recommendation results. Liu et al. [12] combined model-based collaborative filtering and memory-based model for recommendation and then recommended products effectively. Chen [13] applies the ranking learning algorithm to the hybrid recommendation algorithm for model fusion. Generally speaking, such an algorithm structure combines the advantages of content-based recommendation algorithm, but increases the complexity of the algorithm and the running time cost. Recommendation algorithms play a huge role in all areas of human life and work. However, although there are various differences in its algorithms, it must be admitted that each algorithm has its own advantages and applicable characteristics. There is still no algorithm that is close to perfect because various algorithms have their own advantages and disadvantages to a certain extent. Therefore, how to effectively analyze and integrate these algorithm resources so as to combine their strengths and avoid weaknesses as much as possible and then combine their advantages for recommendation has become one of the optimization directions of recommendation algorithms.

He [14] proposed a novel interpretation interface, which focuses on the integration of extracting characteristic emotions and static attributes from product reviews to help users explore and understand the product space more effectively and then learn more from other customers’ experiences and explain the product preferences of target users well. Xu et al. [15] proposed a multitask learning interpretable recommendation algorithm. The algorithm uses tensor decomposition to integrate the user preference modeling in the recommendation task and the modeling of the opinion content in the interpretation task. It can not only predict the target user’s preferences, for items can also provide explanations about the characteristics of specific items. Wei et al. [16] propose a recommendation method that uses tags as features and explain to the user why the recommended movie is related to him based on the features. At the same time, user research experiments have been conducted, and the results show that providing feature-based explanations for the recommendation results can help improve the effectiveness of the recommendation results. Hosseini et al. [17] also verified through user research experiments that feature-based explanations are closely related to user satisfaction and trust in recommendations. Yu [18] proposed a general framework based on graph regularization. The framework first extracts aspects from reviews, then models the user-item-aspect relation as a tripartite graph, which can be used while performing Top-N recommendations, and provides various explanations. Sukrat and Papasratorn [19] developed a phrase-level sentiment analysis toolkit Sentiers1 that can extract user sentiment and product aspects from text reviews on a large scale. This toolkit can not only extract the “aspect-view-emotion” triples from the comments received by the item, for example, extract the “pixel-high-positive” and “cartoon-obvious-negative” three-tuples from the comments on the mobile phone tuple but also detect the emotion of the aspect word based on the context. For example, “pixel” and “high” indicate positive emotions, while “noise” and “high” indicate negative emotions. In addition, since the program constructs a dictionary of “aspect-view-emotion” triples, the toolkit can also detect the triples contained in the review text. Based on the Sentiers toolkit, Fedirko et al. [20] designed an explicit factor model and proposed to use the “aspect-view”the performance of the item in all aspects.

3. Intelligent Information Recommendation Algorithm Based on Deep Learning

This paper proposes an improved topic model to solve the problem of feature extraction in the text of the recommendation system. The model is shown in Figure 1, where shaded circles represent observed variables and unshaded circles represent hidden variables. User comments \( d_{ui} \in D \) on item \( i \). These generated sentences usually focus on a topic, either from the user’s preference or from the characteristics of the item.

When the topic model is used as a text processor, very good performance can be obtained in the recommendation system. The main reason is that the topic model can simulate the process of manually writing comments, generate simulated comments based on user preferences and item characteristics, filter out some irrelevant content, and enhance the representation of users and items.

Convolutional neural network (CNN) is shown in Figure 2. Different from processing image data because text is a characteristic of one-dimensional data, CNN can only use one-dimensional convolution kernel \( h \times d \) for convolution operation, that is, \( d \) is the dimension of a fixed word vector. Multiple convolution kernels with different values of \( h \) can obtain multiple feature maps, where \( h \) refers to convolution with \( h \) words as a window.
wait for the video and don’t rent it

Figure 1: Text processor based on the topic model.

Figure 2: Text classification model based on convolutional neural network.
each time. Specifically, firstly, the convolution kernels with $h$ being 2 and 3 are used for convolution to extract features. Then, the algorithm uses the maximum pooling operation on the obtained features. Finally, the algorithm combines the results of pooling in a splicing way as text features.

When a convolutional neural network is used as a text processor, the model has good performance in text classification and sentiment classification. The main reasons are as follows. The dimension of word vector can be kept within $d$, which avoids the problem of data sparseness and feature dimension disasters and also reduces the amount of model parameters. Convolutional neural networks can extract $h$-ary syntactic features. Convolutional neural networks can obtain more effective feature items in the feature by using the pooling layer and filter out certain noise to improve the accurate expression of features. The second layer is the convolutional layer, which contains $m$ convolution kernels $K \in \mathbb{R}^{d \times d}$. The feature formula extracted by the $j$th neuron from the text is as follows:

$$z_j = \text{ReLU}(w_{1:n} * K_j + b_j),$$

where $*$ is the convolution operation, ReLU is the nonlinear activation function, and $b_j$ is the bias. The text feature extracted by the $j$th element using the sliding window of $i$ is $z_1, z_2, \ldots, z_{n-i+1}$. The third layer is the maximum pooling layer, which can reduce dimensionality and capture the most important features. The formula is

$$o_j = \max(z_1, z_2, \ldots, z_{n-i+1}).$$

Finally, the features extracted by $m$ neurons are pieced together to obtain text features $O = (o_1, o_2, \ldots, o_m)$. The model is shown in Figure 3.

In the recommendation system, the text processor based on the convolutional neural network is adopted by many researchers. The experimental results prove that the text features it extracts can better express users and items, thereby improving the recommendation performance.

The research focus in natural language processing is how to extract rich features from sequence data. For example, in tasks such as translation systems, question answering systems, dialogue systems, and sentiment analysis, recurrent neural networks (RNN) are often used to encode or decode text. Among them, the most common recurrent neural networks for solving sequence problems are the long and short-term memory recurrent neural network. This section will take LSTM as an example to illustrate the process of recurrent neural network processing text.

LSTM proposes to encode sentences according to the order of words. In addition to maintaining the sequence information of the text, it can also solve the problem of long-distance dependence. In addition, when LSTM considers the current word information, it also considers the historical code output by the last recurrent neural network unit. Specifically, LSTM has three gate mechanisms: input gate, forget gate, and output gate. The model is shown in Figure 4.

When coding word $x_t$, the unit calculation process is shown in the following formula:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f),$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i),$$

$$ct = \tanh(W_c x_t + U_c h_{t-1} + b_c),$$

$$c_t = f_t \odot c_{t-1} + i_t \odot ct,$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o),$$

$$h_t = o_t \odot \tanh(c_t).$$

Recurrent neural networks are commonly used to deal with sequence problems. However, due to the limitation of memory, the long-distance dependence of the cyclic neural network is still lost. Because the attention mechanism is not restricted by distance, it has become a popular technique for dealing with sequence problems. The text processor based on self-attention first converts the input matrix $X$ into query matrix $Q$, key matrix $K$, and value matrix $V$, as shown in Figure 5. The calculation formula is as follows:

$$Q = XW^Q,$$

$$K = XW^K,$$

$$V = XW^V,$$

where $W_q \in \mathbb{R}^{d \times d}, W_k \in \mathbb{R}^{d \times d}, W_v \in \mathbb{R}^{d \times d}, Q \in \mathbb{R}^{b \times d}, K \in \mathbb{R}^{b \times d}$, and $V \in \mathbb{R}^{b \times d}$. Next, the algorithm calculates the dot product of query and key, divides the result by $\sqrt{d_k}$, and finally applies the softmax function to obtain the weight of value. The formula is as follows:

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

Because the target vector matrix Attention($Q$, $K$, $V$) calculated by this self-attention mechanism is obtained by point operation, that is, there is interaction between two words, therefore, using it as the coding matrix of the sentence itself, it can capture the dependency between any two words.

The self-attention text processor can extract the dependency relationship between words at any distance, and the execution efficiency is higher than that of the text processors of the convolutional neural network and the recurrent neural network.

As we all know, only part of the words or part of the comments in the comment-based recommendation is useful. There have been many ways using the attention mechanism in conjunction with text processors of convolutional neural networks or cyclic neural networks to capture the usefulness of the word level or sentence level, but there is also a small amount of work that fully utilizes the attention mechanism to achieve this goal. The MPCN model is shown in Figure 6. MPCN uses the co-attention mechanism to deeply interact with user and item comments and captures the word-level and comment-level attention distribution. Specifically, the model first takes the encoded comment $x \in \mathbb{R}^{bd}$ as input.
and uses the co-attention network to first extract the comment-level co-attention. Among them, $l$ is the length of the comment and $d$ is the size of the word embedding. The formula is as follows:

$$ s_{ij} = F(x_i)^T M F(x_j), $$  \hspace{1cm} (6)

where $i$ and $j$ represent the $i$th and $j$th comments, respectively, and $F(.)$ is the feedforward neural network, $M \in \mathbb{R}^{d \times d}$, and $S = s_{ij} \in \mathbb{R}^{l \times l}$. Next, the algorithm calculates comments $x_i$ and $x_j$ and considers the representations $\overline{x_i}$ and $\overline{x_j}$ of the comment-level co-attention; the formula is as follows:

$$ \overline{x_i} = (\text{gumbel}(\max_{\text{col}}(S))^T x_i, $$
$$ \overline{x_j} = (\text{gumbel}(\max_{\text{row}}(S))^T x_j, $$  \hspace{1cm} (7)

where $\max_{\text{col}}$ and $\max_{\text{row}}$ are the maximum pooling operations performed on the columns and rows of the matrix $S$, respectively. The Gumbel function can obtain the one-hot encoded vector, that is, the useful comments are selected.
Then, the word-level co-attention is processed similarly. The algorithm takes $\mathbf{x}_i$ and $\mathbf{x}_j$ as input and outputs comments $\mathbf{x}_i$ and $\mathbf{x}_j$ that consider the word-level co-attention. The formula is as follows:

$$ w_{ij} = F(\mathbf{x}_i)^T M_w F(\mathbf{x}_j), $$

$$ \mathbf{x}_i' = \text{softmax}(\text{avg}_{\text{col}}(w)) \mathbf{x}_i, $$

$$ \mathbf{x}_j' = \text{softmax}(\text{avg}_{\text{row}}(w)) \mathbf{x}_j, $$

where $\text{avg}_{\text{col}}$ and $\text{avg}_{\text{row}}$ are the average pooling operations performed on the columns and rows of the matrix $w$, respectively.

MPCN has greatly improved its performance on 24 publicly accessible datasets.

In order to deal with the translation problem, an end-to-end sequence-to-sequence learning method is proposed. The model of this method is shown in Figure 7, which consists of two recurrent neural networks (RNN). One acts as an encoder to process the input sequence $x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow x_4$, and the other acts as a decoder (sentence generator) to generate a sequence $y_1 \rightarrow y_2 \rightarrow y_3$.

The specific steps are as follows:

(1) The algorithm uses the word embedding model to encode the input sequence and then enters the encoded words into the RNN one by one. The formula is as follows:
The application of the seq2seq model has been widely used in other fields by translation systems, such as question answering systems, dialogue systems, and recommendation systems. Although LSTM and GRU are proposed to improve the dependence of words in sentences, the long-distance dependence is still not ideal. In order to solve this problem, the seq2seq joint attention mechanism is used to achieve better translation. The model is shown in Figure 8. The specific steps are as follows:

1. The algorithm uses the word embedding model to encode the input sequence and then uses RNN to process the encoded words one by one. The formula is as follows:
   \[ h_t = \text{RNN}_{\text{encoder}}(x_t, h_{t-1}). \]
   \[ (9) \]
   where \( x_t \) is the input of the current unit, \( h_{t-1} \) is the output of the previous hidden unit, and \( h_t \) is the output of the current unit.

2. The algorithm uses the output of the encoder as the first hidden state \( h_0 \) of the decoder and uses the decoder to generate words one by one:
   \[ h_t' = \text{RNN}_{\text{decoder}}(\tilde{y}_{t-1}, h_{t-1}'). \]
   \[ (10) \]

3. The algorithm uses the classifier to predict each word. The number of classifications is the size of the vocabulary. The formula is as follows:
   \[ \tilde{s}_t = \text{softmax}(W\tilde{y}_t + b). \]
   \[ (11) \]

4. The algorithm uses the attention weight to weight the hidden layer of the encoder as a context vector. The formula is as follows:
   \[ c_t = \sum_{j=1}^{T} \alpha_{ij} h_j. \]
   \[ (15) \]

5. The algorithm concatenates the context vector and the hidden state of the decoder as the output of the predicted word and then inputs the result to the classifier to predict each word:
   \[ P(y_t | y < t, x) = \text{softmax}(W_s \tanh(W_c[c_t; s_t])). \]
   \[ (16) \]

seq2seq combined with attention has better effect. The RNN unit of the seq2seq encoder and decoder can use a long short-term memory network (LSTM) or a gated recurrent unit (GRU), and the decoder is commonly used in recommendation systems as a generator of interpretable sentences. This study uses hierarchical GRU as a translator to generate natural language sentences, and the generated sentences can explain the recommended results well.

The formula for multihead attention is as follows:
   \[ \text{head}_i = \text{Attention}(QW^q_i, KW^K_i, VW^v_i). \]
   \[ \text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \text{head}_2, \text{head}_n)W^0, \]
   \[ (17) \]
   where Concat (.) is the splicing operation. In a multihead decoder, \( K \) and \( V \) come from the output of the encoder and \( Q \) comes from the output of the upper layer of the decoder.

The final output of the decoder can predict the probability of each word through a linear transformation function. The formula is as follows:
   \[ P_{E_{\text{pos},2i}} = \frac{\sin \text{pos}}{10000^{\frac{2i}{d_{\text{mod}}}}}, \]
   \[ P_{E_{\text{pos},2i+1}} = \frac{\cos \text{pos}}{10000^{\frac{2i}{d_{\text{mod}}}}}, \]
   \[ (18) \]
   where pos is the position of the word, \( i \) is the dimension of the word, and \( d_{\text{mod}} \) is the dimension of the position code.

The encoder and decoder models based on the self-attention mechanism further improve the effect of machine translation. The translation model is similar to seq2seq, in which the decoder can also be used as a sentence generator.
Figure 9: Cross-border e-commerce intelligent information recommendation system based on deep learning.
in the recommendation system to generate natural language sentences as the interpretation of the recommendation results.

4. Cross-Border e-Commerce Intelligent Information Recommendation System Based on Deep Learning

The cross-border e-commerce intelligent information recommendation system based on deep learning is shown in Figure 9. The model is divided into 3 layers. The first layer is the embedding layer, which initially expresses users and points of interest through the TransR model. The second layer is the relational reasoning layer, which forms a multilayer relational reasoning through random walks on the knowledge graph, calculates the information attenuation after each walk, and uses a deep neural network to perform a new vector mapping. Each mapping of a vector is a performance of relational information. Each mapping is concatenated to become the final vector representation of the user or point of interest, and the user set and the point of interest set are divided. The third layer is the output layer, which uses the vector representation of the new user’s points of interest and recommends top-N points of interest to the user based on the collaborative filtering algorithm.

The new model constructed in this study links the status quo of various businesses in reality and changes the starting point of the previous recommendation system. We use the long tail theory to mine the mass commodities in the niche,
recommend these products to users as suggestions, and combine the graph model to propose a niche commodity recommendation algorithm based on the graph search strategy. The specific framework is shown in Figure 10.

On the basis of the above analysis, the model proposed in this paper is verified. First of all, this article conducts simulation research on the effect of intelligent information processing and builds the system model of this article through simulation software to simulate the intelligent information processing of cross-border e-commerce. The results obtained are shown in Table 1.

From the above research, it can be seen that the method proposed in this paper can have a good effect. On this basis, the recommendation effect test is carried out, and the results shown in Table 2 are obtained.

Through the above experiments, it can be seen that the cross-border e-commerce intelligent information recommendation system based on deep learning proposed in this paper has a good recommendation effect and meets the recommendation needs of cross-border e-commerce.

5. Conclusion

With the continuous expansion of the scale of online platforms, online trading has become one of the indispensable activities on people’s daily lives in the new era. This has also stimulated offline entities to a certain extent, forming an integrated online and offline sales strategy. With the increasing scale of online business platforms, recommendation systems have also risen, and they have become one of the backbone technologies in various fields today. It is also mainly from the perspective of “people-oriented” to make the originally large and complex assembly line more intelligent, which greatly facilitates human beings. This paper applies deep learning to the cross-border e-commerce intelligent information recommendation system to overcome the geographical problems of cross-border e-commerce and improve the recommendation effect of cross-border e-commerce. The experiment shows that the cross-border e-commerce intelligent information recommendation system based on deep learning proposed in this paper has a good recommendation effect and meets the recommendation needs of cross-border e-commerce.

Data Availability

The labeled dataset used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares no conflicts of interest.

Acknowledgments

This study was sponsored by research about the impact of network interaction on consumers’ purchase decision behavior Humanities and Social Sciences Planning Project of Henan Provincial Department of Education (2020–22JH-283) and Key Scientific Research Project of High Education in Henan Province (19A880024): “Research on Innovation and Entrepreneurship System Based on Cloud Computing Theory”.

References

[1] J. Anitha and M. Kalaiarasu, “Optimized machine learning based collaborative filtering (OMLCF) recommendation system in e-commerce,” Journal of Ambient Intelligence and Humanized Computing, vol. 12, no. 6, pp. 6387–6398, 2021.
[2] H. Chen, “Personalized recommendation system of e-commerce based on big data analysis,” Journal of Interdisciplinary Mathematics, vol. 21, no. 5, pp. 1243–1247, 2018.
[3] F. Yang, “A hybrid recommendation algorithm-based intelligent business recommendation system,” Journal of Discrete Mathematical Sciences and Cryptography, vol. 21, no. 6, pp. 1317–1322, 2018.
[4] L. Zhou, “Product advertising recommendation in e-commerce based on deep learning and distributed expression,” Electronic Commerce Research, vol. 20, no. 2, pp. 321–342, 2020.
[5] H. Pan and Z. Zhang, “Research on context-awareness mobile tourism e-commerce personalized recommendation model,”
Journal of Signal Processing Systems, vol. 93, no. 2, pp. 147–154, 2021.

[6] V. Subramaniyaswamy, R. Logesh, M. ChandraShekhar, A. Challa, and V. Vijayakumar, “A personalised movie recommendation system based on collaborative filtering,” International Journal of High Performance Computing and Networking, vol. 10, no. 1-2, pp. 54–63, 2017.

[7] J. Shen, T. Zhou, and L. Chen, “Collaborative filtering-based recommendation system for big data,” International Journal of Computational Science and Engineering, vol. 21, no. 2, pp. 219–225, 2020.

[8] D. Chang, H. Y. Gui, R. Fan, Z. Z. Fan, and J. Tian, “Application of improved collaborative filtering in the recommendation of e-commerce commodities,” International Journal of Computers, Communications & Control, vol. 14, no. 4, pp. 489–502, 2019.

[9] Z. Wang, M. Wan, X. Cui et al., “Personalized recommendation algorithm based on product reviews,” Journal of Electronic Commerce in Organizations, vol. 16, no. 3, pp. 22–38, 2018.

[10] Y. Huang, Y. Chai, Y. Liu, and J. Shen, “Architecture of next-generation e-commerce platform,” Tsinghua Science and Technology, vol. 24, no. 1, pp. 18–29, 2018.

[11] L. Liang and X. Qin, “Research on consumers online shopping decision-making and recommendation of commodity based on social media network,” Cluster Computing, vol. 22, no. 3, pp. 6529–6539, 2019.

[12] D. Liu, C. Huo, and H. Yan, “Research of commodity recommendation workflow based on LSH algorithm,” Multimedia Tools and Applications, vol. 78, no. 4, pp. 4327–4345, 2019.

[13] J. Chen, “The application of commodity recommendation in cross-border E-commerce: current situation and prospect,” Frontiers in Economics and Management, vol. 2, no. 1, pp. 266–274, 2021.

[14] G. He, “Enterprise E-commerce marketing system based on big data methods of maintaining social relations in the process of E-commerce environmental commodity,” Journal of Organizational and End User Computing, vol. 33, no. 6, pp. 1–16, 2021.

[15] J. Xu, Z. Hu, and J. Zou, “Personalized product recommendation method for analyzing user behavior using DeepFM,” Journal of Information Processing Systems, vol. 17, no. 2, pp. 369–384, 2021.

[16] C. Wei, J. Niu, and Y. Guo, “DLGNN: a double-layer graph neural network model incorporating shopping sequence information for commodity recommendation,” Sensors and Materials, vol. 32, no. 12, pp. 4379–4392, 2020.

[17] F. Hosseini, H. Sadighi, S. A. Mortazavi, and H. Farhadian, “An E-commerce SWOT analysis for export of agricultural commodities in Iran,” Journal of Agricultural Science and Technology A, vol. 21, no. 7, pp. 1641–1656, 2019.

[18] T. Yu, “A case study of B2C cross-border e-commerce challenges in China from Customs to consumers,” World Customs Journal, vol. 12, no. 2, pp. 121–132, 2018.

[19] S. Sukrat and B. Papasratorn, “An architectural framework for developing a recommendation system to enhance vendors’ capability in C2C social commerce,” Social Network Analysis and Mining, vol. 8, no. 1, pp. 1–13, 2018.

[20] O. Fedirko, T. Zatonatska, T. Wolowiec, and S. Skowron, “Data science and marketing in E-commerce amid COVID-19 pandemic,” European Research Studies Journal, vol. 24, Special 2, pp. 3–16, 2021.