RESEARCH ON E-COMMERCE CUSTOMER SATISFACTION EVALUATION METHOD BASED ON PSO-LSTM AND TEXT MINING

Qin Yang*

School of Marxism, Jinling Institute of Science and Technology, Nanjing, Jiangsu, 211169, China

yang2161980@sina.com

Reception: 13/11/2022 **Acceptance**: 29/12/2022 **Publication**: 13/01/2023

Suggested citation:

Y., Qin (2023). **Research on E-commerce Customer Satisfaction Evaluation Method Based on PSO-LSTM and Text Mining**. *3C Empresa. Investigación y pensamiento crítico, 12(1)*, 51-66. https://doi.org/10.17993/3cemp.2023.120151.51-66

ABSTRACT

With the increase of social technology, e-commerce platforms have entered a period of rapid development. Improving customer satisfaction and purchase rate is the key to the survival of e-commerce platforms. Text mining and analysis of customer evaluation data will help to grasp the focus of customers and optimize the e-commerce platform. To this end, through text mining technology, the text comment data of five e-commerce platforms such as Amazon, eBay, Alibaba, Jingdong, and Taobao are collected, and the cleaned text is analyzed by particle swarm algorithm (PSO)-long short-term memory (LSTM) model. The data is subject to time scale extraction, and the extraction results are visualized and interpreted. The research shows that the logistics, price, freshness, quality and packaging of e-commerce platform merchants are important factors that affect the evaluation of e-commerce customer satisfaction.

KEYWORDS

text mining; PSO-LSTM; particle swarm algorithm; long short-term memory network

PAPER INDEX

https://doi.org/10.17993/3cemp.2023.120151.51-66

1. INTRODUCTION

Online shopping has become an important form of shopping. Not only are more and more consumers choosing to shop online, but also more and more categories of consumers are shopping online. After several years of development, various e-commerce platforms are in full swing, but fresh food e-commerce, known as the "last blue ocean" in the e-commerce industry, is still in its infancy. The biggest difference between fresh e-commerce and other e-commerce is that the products pay more attention to freshness and are not easy to preserve. Data from Research shows that the e-commerce market is developing rapidly, with an average annual growth rate of more than 50%. However, due to the constraints of many factors, the overall service of e-commerce is still in the immature stage, its business model is still developing, and the service level is also mixed. The degree is not optimistic [1-3]. The epidemic in 2020 has brought great opportunities to e-commerce [4-6]. Therefore, how to seize this opportunity requires in-depth analysis of the factors affecting customer satisfaction on e-commerce platforms, improving e-commerce products and services, increasing customer satisfaction, and enhancing user stickiness, thereby promoting the development of e-commerce.

For the analysis of e-commerce customer satisfaction evaluation methods, scholars at home and abroad mainly use the text mining method based on sentiment analysis, classify the vocabulary contained in the customer evaluation content, and quantify the emotional trend in the text content through weighted assignment. Based on this method, Susan et al. established an evaluation help degree model to describe the enthusiasm of the evaluation, but did not consider the difference between the evaluation of professional buyers and ordinary customers, and lacked accuracy. The problem of simple evaluation content sentiment analysis has been basically solved, but the evaluation of evaluation content is affected by the additional functions of the e-commerce platform, and a variety of restrictive factors must be comprehensively considered. Yang Ligong et al. [7] used Markov Logic Network to combine sentence context and emotional features for sentiment analysis, and realized cross-domain text analysis. Ming Junren [8] applied the association rule method to text mining analysis, and designed a text data analysis method integrating semantics and association mining, which improved the accuracy of text analysis. Cai Xiaozhen et al. [9] selected 4 indicators as the basis for online user comments, and constructed a text mining model to solve the problem of uneven quality of online user comments. Tang Xiaobo et al. [10] combined co-word analysis and polarity transfer method for text analysis, which overcomes the shortcomings of traditional text analysis and has better text analysis results. Liu L[11] and others proposed to combine the feature vector model and the weighting algorithm of product review sentiment analysis for text analysis. Walaa Medhat [12] proposed that the conventional analysis steps for commodity text analysis are commodity review, emotion recognition, feature selection, emotion classification, and emotion polarity judgment. The above studies have improved the accuracy and efficiency of text analysis, and laid the foundation for the use of particle swarm algorithm to mine intelligence value in text.

https://doi.org/10.17993/3cemp.2023.120151.51-66

With the rapid development of society, accurate e-commerce customer satisfaction prediction is becoming more and more important. The accurate prediction of e-commerce customer satisfaction not only plays an irreplaceable role in the long-term stable operation of e-commerce platforms and sellers, but also plays an important role in reducing the cost of e-commerce platforms, improving product quality and market planning. With the emergence of deep learning, many scholars have turned their attention to deep belief networks, convolutional neural networks, and recurrent neural networks. The related research progress is shown in Figure 1.

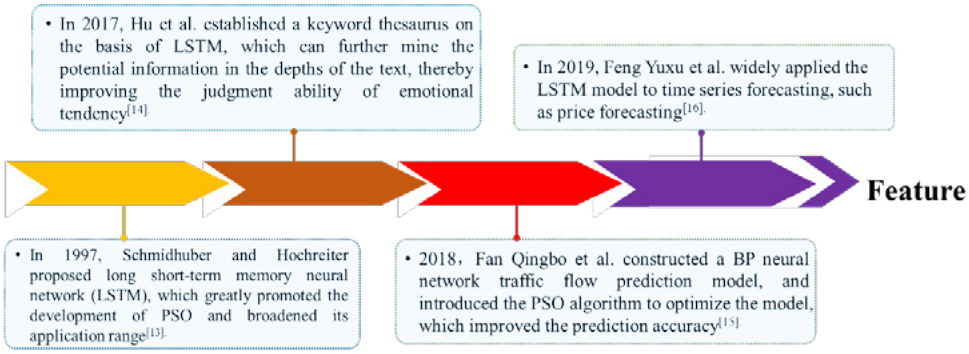

Figure 1 Research progress of LSTM model

As a special form of regular recurrent neural network (RNN), long short-term memory network (LSTM) was first proposed by Hochreiter et al. in 1997, and it is widely used in text mining data prediction. In order to improve the prediction accuracy of e-commerce customer satisfaction, particle swarm optimization (PSO) optimization of long short-term memory (LSTM) neural network hyperparameters for e-commerce customer satisfaction prediction model (PSO-LSTM) is widely used. Aiming at the problem that LSTM hyperparameters are difficult to select, the PSO algorithm can effectively find the global optimal solution to optimize the hyperparameters of the LSTM model, and continuously train to find suitable hyperparameters and verify them. Therefore, this paper combines the PSO-LSTM algorithm with text mining technology to study and analyze the factors that affect the evaluation of e-commerce customer satisfaction.

2. RESEARCH ON TEXT MINING AND ANALYSIS METHODS

Online reviews of e-commerce customers are messy but valuable unstructured or semi-structured data. The information contained in online reviews has an important impact on the strategic decision-making of merchants and the purchasing decisions of consumers [13-14]. At present, the academic community has recognized the importance of online reviews, but the research on online review willingness has not received much attention. It is necessary to link online reviews with Internet word-of-mouth, and study consumers' review willingness from the perspective of Internet word-of-mouth dissemination willingness. This paper directly studies the influencing factors of customer satisfaction evaluation, and uses the method of text mining in the

https://doi.org/10.17993/3cemp.2023.120151.51-66

research to explore which factors will affect consumers' evaluation behavior and willingness [15].

Text mining analysis has a strong foundation under the support of network data. Sem Eval, an international evaluation expert, defines evaluation objects as expressions that can be used to express the characteristics of evaluation entities in specified texts. When evaluating an entity as a commodity, the evaluation object at this time may be the product features, functions, and parts corresponding to the entity. For product reviews on e-commerce platforms, the evaluation objects include express delivery, services, etc. in addition to the inherent attributes of the product. For example: "The phone looks tall and the system runs very fast, but the power is not durable, so I'm a little disappointed." For this comment, "appearance" and "battery" are the evaluation objects, "tall", "Running very fast" and "Not durable" are the comments corresponding to the three evaluation objects. Therefore, the extraction of evaluation objects and comments plays an important role in the analysis of commodity evaluation texts. By analyzing the above comments, it can be found that the comments corresponding to the first two evaluation objects are positive, while the comments corresponding to the third evaluation object are negative. In the actual analysis process, whether the reviews are classified as positive, negative or neutral will affect the accuracy of text mining analysis of product reviews [16-17]. Therefore, the text mining process will extract the evaluation objects and their corresponding comments in each product review, and analyze them, which increases the accuracy of text mining analysis to a certain extent, and can be better applied to e-commerce customer satisfaction. Under evaluation. In addition, it helps e-commerce companies understand their competitive environment and position, discover the breakthrough points of the industry, and further adjust their own development strategies, so that consumers' purchasing experience in merchants can be optimized, and it is also helpful to improve the overall industry. performance level. Taking computer furniture, seafood and aquatic products, fruits and vegetables, and daily necessities as examples, some data of specific customer full score evaluation are shown in Table 1.

Table 1 Some product review data

E-commerce customer	Comment	product category
M***P	I received the phone, and the pink color is too beautiful. I really like this phone. Can last all day on a single charge	phone
Little***j	The logistics is very fast, and the lining protection of the packaging is very characteristic. The dual graphics cards work together, and the video and audio effects are good~~~	computer

https://doi.org/10.17993/3cemp.2023.120151.51-66

Call***girl	The vegetables are very fresh and the price is affordable. There are also garlic moss and eggplants that are currently scarce in the market. It is really good. Thank you for the efforts made by the ** platform for us!	vegetable and fruit
OH***!	The platform is fast, I place an order in the morning and arrive in the afternoon. After the logistics arrives, the frozen state is kept well, and the fish is very fresh.	seafood
See***door	My son loves reading books very much, he bought a lot of books on the e-commerce platform, the quality is very good, very beautiful	book
MY***baby	All solid wood furniture, no glue, no paint, with a touch of pine fragrance, I like it very much	furniture
YOU***left	The fabric and form are really, really good! Bought it for my husband, the e-commerce customer service is very greasy and loving! Buy with confidence!	clothing

The process of text mining analysis method includes text feature extraction, text data cleaning, high-frequency text selection and processing. The feature representation of text refers to a method that can use words, words, phrases or sentences as feature items to represent the entire text, and complete unstructured text processing by processing these feature items. Text features can generally be accurately expressed by words, words and phrases, while sentences and paragraphs can be further divided into words, words and phrases. The text representation method is currently uncommon. Commonly used text feature representation types are as follows:

(1) Words. Individual numbers, letters, spaces, special symbols, and Chinese characters that constitute the structural units of phrases, phrases, and concepts cannot reflect the characteristic meaning and emotional expression of the text, and need to be effectively combined. Therefore, word-based text feature representation requires Extraction and optimization of text words [18-19].

(2) Phrases. To a certain extent, it can be used as a combination of words at the most basic semantic level. Phrase-based text feature representation In a certain limited field, the phrase feature space may contain tens of thousands or even millions of phrases, which requires special dimensionality reduction processing.

https://doi.org/10.17993/3cemp.2023.120151.51-66

(3) Phrases. Single words or compound phrases extracted from the original corpus directly by entity extraction methods are generally composed of specific words, and most of them exist in typical text dictionaries.

(4) Concept. Based on rules or hybrid classification methods, through preprocessing procedures, specific unit combinations are formed by manual counting, identifying individual words, compound phrases, entire sentences and even larger syntactic units.

This text mining analysis collects text comment data from five e-commerce platforms such as Amazon, eBay, Alibaba, JD.com, and Taobao. The product types include furniture, home appliances, digital products, fresh fruits, seafood and meat, and daily necessities[20-22]. The collected text data is processed by data cleaning, word segmentation, etc., and text mining analysis is performed on the data to construct high-frequency words in comments to study the customer satisfaction evaluation of e-commerce platforms.

1) Use web crawler technology to capture customer comment texts of merchants on various e-commerce platforms, filter and clean the data, remove word segmentation and other preprocessing.

2) Perform text mining analysis on the preprocessed data. This stage mainly includes the following steps: ① Select an appropriate method to determine the number of topics (because the number of topics is too large or too small will affect the interpretation of text analysis results); ② Use text mining analysis to extract topics, and thus generate document-topic Distribution matrix and topic-vocabulary distribution matrix; ③Simple sorting and interpretation of the extracted topics; ④Visually display the results of text analysis.

3) Through the interpretation of the text analysis results, it is finally verified how the merchants of the e-commerce platform affect the satisfaction evaluation of consumers, and whether there are other factors in the influence process that mediate or moderate the influence process [23-24].

In order to study the e-commerce customer satisfaction evaluation criteria more accurately, the text topics identified by text mining analysis are in common: logistics, price, service, quality, and packaging. Based on the five types of e-commerce platforms such as Amazon, eBay, Alibaba, Jingdong and Taobao, the text analysis method is used to calculate the five text themes in the five categories of furniture, home appliances, digital products, fresh fruits, seafood and meat, and daily necessities. The weight occupied, and the radar chart is drawn according to the weight, as shown in Figure 2.

https://doi.org/10.17993/3cemp.2023.120151.51-66

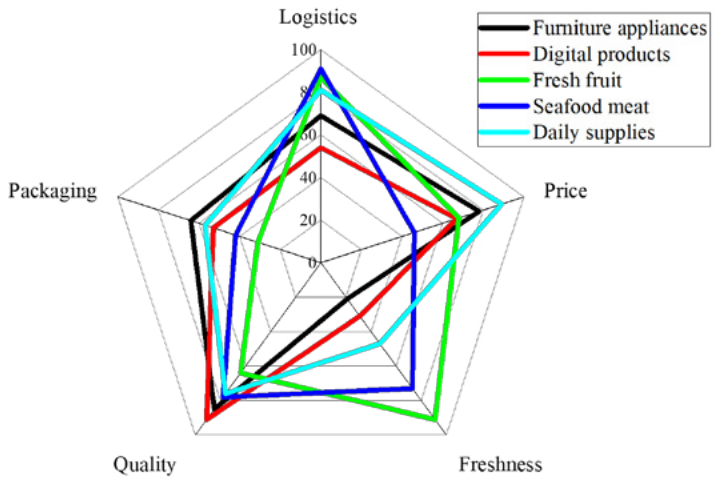

Figure 2 Radar chart of various themes of various e-commerce platforms

It can be seen from Figure 2 that customers value different characteristics of different types of products on the e-commerce platform. For example, for commodities such as fresh fruits, seafood and meat, consumers place the most importance on freshness and quality, and for other types of products, consumers place the most importance on quality. On the whole, consumers pay more attention to the quality, price and logistics of the products of the e-commerce platform. According to the research results, for the major e-commerce platforms, the following suggestions are put forward to improve customer satisfaction:

quality. The above e-commerce platforms all have stable brand suppliers and controllable product sources. There are certain advantages in terms of quality. E-commerce platforms can also achieve quality traceability through technologies such as blockchain and the Internet of Things. After the information is visualized, consumers have more trust in the quality of the products, and the platform will pay more attention to the quality problems in the production process. Price [25]. Some products of the above-mentioned e-commerce platforms are directly sourced, and some products have established their own product bases, so they can provide relatively favorable prices. E-commerce platforms can also compare multiple parties when selecting suppliers. Under the premise of ensuring quality, they can cooperate with suppliers for a long time. At the same time, they can optimize logistics, warehousing and other solutions to reduce costs, so as to give customers the best price.

3. E-COMMERCE CUSTOMER SATISFACTION EVALUATION METHOD BASED ON PSO-LSTM AND TEXT MINING

In the previous section, we conducted text mining on e-commerce customer satisfaction evaluation texts to explore how e-commerce platforms affect consumer satisfaction evaluation behavior, and to find important factors that affect e-commerce

https://doi.org/10.17993/3cemp.2023.120151.51-66

customers' satisfaction evaluation behavior. The factor is single, only high-frequency words are counted, and the time changes of customer evaluations are not considered, so the accuracy is low. The PSO-based LSTM algorithm transforms its structure on the basis of the traditional artificial neural network algorithm, so as to achieve the purpose of enabling the network to remember past time information, so that the network can not only realize the connection from bottom to top (input-output), but also It can also realize information transmission and recording from left to right (time t-time t+1). Therefore, the combination of PSO-LSTM calculation and text mining analysis can more accurately study the satisfaction evaluation of e-commerce customers. The algorithm flow of PSO-LSTM will be discussed in depth below.

1 The forward propagation process

(1) Input-hidden layer

The PSO-LSTM algorithm is a two-dimensional feature algorithm. After the input layer-hidden layer-output layer operation, the local features of the input data are extracted. In this paper, ft is used to represent the channel input of the PSO-LSTM algorithm model, yt is the output weight of the PSO-LSTM algorithm model, Wf is the bias vector of the input layer, and Vht is the receptive field of the output layer on the input data. The algorithm process is shown in formulas (1)-(2):

$$f_t = \sigma\left(W_f h_{t-1} + U_f x_t + b_f\right). \quad (1)$$

(2) Hidden layer - output layer

$$y_t = \sigma\left(V h_t + b_y\right). \quad (2)$$

2. Backpropagation process

(1) Define the error function of the sequence

At time t, the actual output of the LSTM is $y_t = \left\{y_1(t), y_2(t), \ldots, y_p(t)\right\}^T$, while the expected output is $d_t = \left\{d_1(t), d_2(t), \ldots, d_p(t)\right\}^T$, Then the loss function of the entire time series is

$$L = \sum_{t=1}^{T} L_t. \quad (3)$$

Therefore,

$$L_t = \frac{1}{2}\left(y_t - d_t\right)^2. \quad (4)$$

(2) Define the error term at time t

That is, the partial derivative of the loss function to the output value. First, define the error term of the output layer at time t as $\delta y_t = \frac{\partial L_t}{\partial y_t}$, Second, define the error term in the hidden layer at time t.

https://doi.org/10.17993/3cemp.2023.120151.51-66

Error in output gate:

$$\delta h_t = \frac{\partial L_t}{\partial h_t} = \frac{\partial L_t}{\partial y_t}\frac{\partial y_t}{\partial h_t} = \delta y_t \cdot y_t \quad (1 - y_t) \quad \cdot$$

$$\delta o_t = \frac{\partial L_t}{\partial o_t} = \frac{\partial L_t}{\partial h_t}\frac{\partial h_t}{\partial o_t} = \delta h_t \cdot \tanh(C_t). \tag{5}$$

Because the information of all moments is stored in the memory cell state, the error term δCt needs to be accumulated at each moment, and the error returned to the memory cell is divided into two parts, the first part is returned from the output ht of the hidden layer The second part is the error returned from the memory cell state at the next moment Ct+1:

$$\delta C_t = \frac{\partial L_t}{\partial C_t} = \frac{\partial L_t}{\partial h_t}\frac{\partial h_t}{\partial C_t} = \delta h_t \cdot \left[1 - \tanh(C_t)^2\right] \quad \cdot \tag{6}$$

First calculate the error term of the memory cell state in the hidden layer at the previous moment:

$$\delta C_{t-1} = \frac{\partial L_t}{\partial C_{t-1}} = \frac{\partial L_t}{\partial C_t}\frac{\partial C_t}{\partial C_{t-1}} = \delta C_t \cdot f_t, \tag{7}$$

From this formula, the error returned from the memory cell state at time t+1 Ct+1:

$$\delta C_t = \delta C_{t+1} \cdot f_{t+1}, \tag{8}$$

Combining (6) and (8), the error term of the memory cell state is obtained as:

$$\delta C_t = \delta C_{t+1} \cdot f_{t+1} + \delta h_t \quad \left[1 - \tanh(C_t)^2\right] \quad \cdot \tag{9}$$

Error in input gate:

$$\delta i_t = \frac{\partial L_t}{\partial i_t} = \frac{\partial L_t}{\partial C_t}\frac{\partial C_t}{\partial i_t} = \delta C_t \cdot C_t',$$

$$\delta C_t' = \frac{\partial L_t}{\partial C_t'} = \frac{\partial L_t}{\partial C_t}\frac{\partial C_t}{\partial C_t'} = \delta C_t \cdot i_t. \tag{10}$$

Error in output gate:

$$\delta f_t = \frac{\partial L_t}{\partial f_t} = \frac{\partial L_t}{\partial C_t}\frac{\partial C_t}{\partial f_t} = \delta C_t \cdot C_{t-1}. \tag{11}$$

(3) Calculate the error gradient of the weight coefficient matrix

The first step is to calculate the error gradient of the weight matrix V from the hidden layer to the output layer, and the two sides of the loss function L take the partial derivative of V to get:

$$\frac{\partial L}{\partial V} = \sum_{t=1}^{T}\frac{\partial L_t}{\partial V} = \sum_{t=1}^{T}\frac{\partial L_t}{\partial y_t}\frac{\partial y_t}{\partial V} = \sum_{t=1}^{T}\delta y_t \cdot y_t \quad (1 - y_t) h_t^T. \tag{12}$$

https://doi.org/10.17993/3cemp.2023.120151.51-66

The next step is to calculate the error coefficient in the hidden layer:

$$\frac{\partial L}{\partial W_f} = \sum_{t=1}^{T} \frac{\partial L_t}{\partial W_f} = \sum_{t=1}^{T} \frac{\partial L_t}{\partial f_t} \frac{\partial f_t}{\partial W_f} = \sum_{t=1}^{T} \delta f_t \square f_t \quad \square - f_t) h_{t-1}^T. \tag{13}$$

The error gradient from the input layer to the output layer weight matrix can also be obtained in the same way:

$$\frac{\partial L}{\partial U_f} = \sum_{t=1}^{T} \frac{\partial L_t}{\partial U_f} = \sum_{t=1}^{T} \frac{\partial L_t}{\partial f_t} \frac{\partial f_t}{\partial U_f} = \sum_{t=1}^{T} \delta f_t \square f_t \quad \square - f_t) x_t^T. \tag{14}$$

(4) Update of the weight matrix

$$V_{new} = V + \eta \frac{\partial L}{\partial V}. \tag{15}$$

In this paper, the PSO-LSTM algorithm uses an activation function in the neurons and the final prediction layer, which can combine the linear input nonlinearly, so that the PSO-LSTM algorithm has nonlinear factors, and the prediction is more accurate. The commonly used activation functions are Sigmoid function, Relu function, Tanh function. The activation functions of the LSTM network at the neurons are the Relu function and the Tanh function, which are often used because of their simple calculation and fast iteration speed [26-27]. The Sigmoid function is used in the output layer for binary classification probability calculation.

The mathematical formula for the sigmoid function is：

$$sigmoid(x) = \frac{1}{1 + e^{-x}} \tag{16}$$

The mathematical formula for the Tanh function is:

$$\tanh(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}} \tag{17}$$

The mathematical formula for the Relu function is：

$$f(x) = \max(0, x) \tag{18}$$

According to the traditional method of training deep models, the LSTM model in this paper uses the stochastic gradient descent algorithm to train parameters. Choose a cross-entropy loss function according to the classification target definition：

$$E^N = -\frac{1}{N} \sum_{n=1}^{N} \sum_{k=1}^{c} t_k^n \log y_k^n \tag{19}$$

Among them, N represents the number of samples, and c represents the number of categories. tkn represents the true category of the nth sample, and ykn represents the prediction result of the nth sample. The purpose of training the model is to minimize the loss function. Define the output of layer Z as：

https://doi.org/10.17993/3cemp.2023.120151.51-66

$$x^l = f\left(u^l\right)$$
$$u^l = \omega^l x^{l-1} + b^l \quad (20)$$

Among them, f represents the activation function, xl−1 represents the output of the L-1 layer, and for the L layer, it is also its input, l represents the weight of the L layer, and b represents the L layer bias. In the process of error back propagation algorithm, this model uses gradient descent method to update the weight of the calculation network. The update calculation formula of gradient descent method is as follows:

$$\omega_{\text{new}}^l = \omega_{\text{old}}^l - \eta \frac{\partial E}{\partial \omega_{\text{old}}^l}$$
$$b_{\text{new}}^l = b_{\text{old}}^l - \eta \frac{\partial E}{\partial b_{\text{old}}^l} \quad (21)$$

Among them, represents the learning rate in the gradient descent calculation, and the update formula of the weight of each layer can be obtained according to the chain derivation rule. In the process of classifying information texts, there are many situations in the classification results. The following four possibilities exist for the classification results:

(1) True class True positives (TP): The number of positive samples that were successfully identified;

(2) False positives (FP): The number of negative samples that are misidentified;

(3) False negatives (FN): The number of positive samples that are incorrectly identified;

(4) True negatives (TN): The number of negative samples that are correctly identified;

Based on the method of text mining in Chapter 3, and the high-frequency vocabulary logistics, price, freshness, quality and packaging it excavated are the text positive classification data of PSO-LSTM. The sample data are customer evaluations of furniture appliances, digital products, fresh fruits, seafood and daily necessities of the five major e-commerce platforms, from March 2017 to March 2021. The essence of the experiment in this paper is to classify information texts, so only the text and category labels in the collected forum data are extracted as experimental data. The purpose of selecting category labels is to distinguish normal information and spam information with 0 and 1 respectively, the class labels in the category are verified manually, so the labels of the data are true and reliable. The total number of collected data is 83,961 texts, including 14,678 spam messages and 69,283 normal messages. Formulas (1)-(15) are the calculation equations for filtering text data by PSO-LSTM, and the activation function is used in the neurons and the final prediction layer, which can combine the linear input nonlinearly, so that the PSO-LSTM algorithm has non-linear combination. Linear factor, the prediction is more accurate, and the high-frequency words of text mining are classified and calculated according to formulas (22)-(25). Through the previous theoretical research basis and data processing

https://doi.org/10.17993/3cemp.2023.120151.51-66

process, the time division of the weights of high-frequency vocabulary logistics, price, freshness, quality and packaging is carried out, as shown in Figure 3

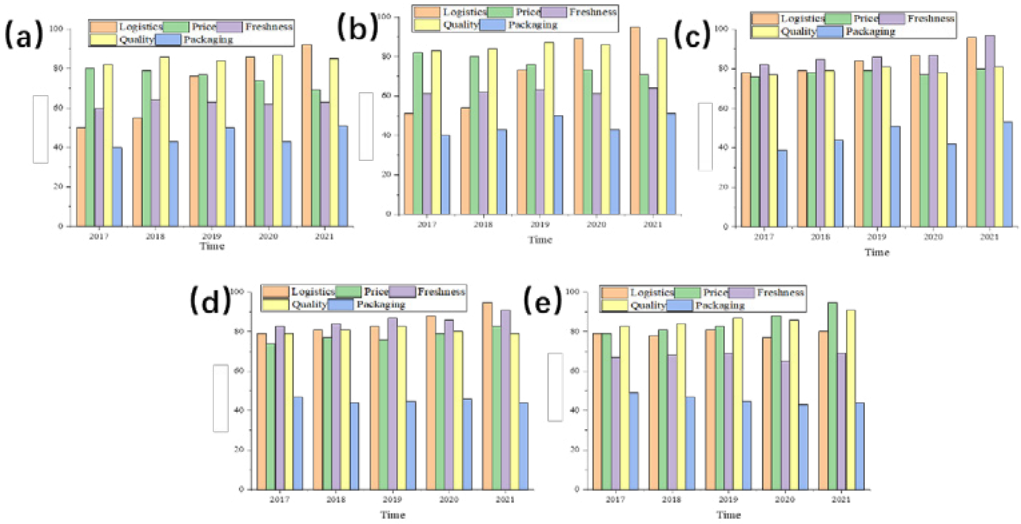

Figure 3 The relationship between the characteristics of various e-commerce platforms (a) furniture and home appliances (b) digital products (c) fresh fruit (d) seafood and meat (e) daily necessities over time

As can be seen from Figure 3(a), with the increase of time, the highest change rate of high-frequency words in e-commerce customers' satisfaction with furniture and home appliances is logistics speed, and its weight is 52 in 2017, and gradually increases to 91 in 2021. This also means that when other factors remain unchanged, the faster the logistics of furniture and home appliances on the e-commerce platform, the higher the customer satisfaction rating; as shown in Figure 3(b), as time goes by, e-commerce customers are more interested in The highest rate of change in the high-frequency words of satisfaction of digital products is also the logistics speed, and its weights from 2017 to 2021 are 51, 54, 73, 89 and 95 respectively; as can be seen from Figure 3(c)-(d), With the increase of time, the most frequent change rate of e-commerce customers' satisfaction with fresh fruits, seafood and meat is their freshness, followed by logistics. The weight of the freshness of the subject word was 82 in 2017 and gradually increased to 97 in 2021. This also means that when other factors remain unchanged, the faster the freshness of fresh fruits and seafood on the e-commerce platform, the higher the customer satisfaction rating; as shown in Figure 3(d), with the increase of time, E-commerce customers' satisfaction with daily necessities has the highest change rate of high-frequency words in price, followed by quality.

In summary, in 2017, the development of e-commerce platforms was relatively weak, the number of e-commerce shoppers was relatively small, and customer satisfaction at that time was more inclined to price, which also meant that the higher the product prices of the merchants on the e-commerce platform. When it is low, the customer satisfaction rating is higher. With the development of e-commerce platforms, customers pay more and more attention to the speed of logistics and the freshness of products. Through data processing through the PSO-LSTM algorithm, it can be seen that in 2018, the proportion of logistics and freshness has increased significantly,

https://doi.org/10.17993/3cemp.2023.120151.51-66

which also means that e-commerce The faster the product logistics speed of the merchants on the platform and the higher the freshness, the higher the customer satisfaction rating. With the advent of the epidemic, the e-commerce platform has entered an ice age, but through data analysis, it can be seen that from 2019 to 2021, product orders on the e-commerce platform will still grow steadily.

4. CONCLUSION

This paper mainly studies the analysis process of e-commerce customer satisfaction evaluation based on PSO-LSTM and text mining, and conducts analysis and research according to the obtained results. The text data of the text all come from five categories of products such as furniture, home appliances, digital products, fresh fruits, seafood and meat, and daily necessities on the five e-commerce platforms of Amazon, eBay, Alibaba, Jingdong, and Taobao. When using text mining technology to analyze the evaluation of e-commerce customer satisfaction, five subject words that have the highest impact on customer satisfaction evaluation are obtained: logistics, price, freshness, quality and packaging.

Since the text mining technology does not consider the time factor, the accuracy of the measured factors is low. In order to more accurately study the factors affecting the evaluation of e-commerce customer satisfaction on the time scale, the text analysis technology and the PSO-LSTM model were combined to further analyze the five types of high-frequency words that affect the evaluation of e-commerce customer satisfaction, and to analyze Its proportion is analyzed. Analyzed the proportion of factors affecting the evaluation of e-commerce customer satisfaction in each time period. For furniture and home appliances, with the increase of time, the highest change rate of high-frequency words in e-commerce customers' satisfaction with furniture and home appliances is logistics speed, the maximum weight is 91 in 2021, and the minimum weight is 52 in 2017; E-commerce customers who buy fresh fruit or seafood are more concerned about the freshness of the goods, and their weight will gradually increase from 82 in 2017 to 97 in 2021. Through the data processing of PSO-LSTM algorithm, it can be seen that due to the impact of the epidemic, the logistics speed and quality have shown a trend of substantial improvement. On the whole, consumers pay more attention to the quality, price and logistics of the products of the e-commerce platform. The evaluation method of e-commerce customer satisfaction based on PSO-LSTM and text mining designed in this paper can effectively solve the imbalance of information datasets, and is exp

REFERENCES

(1) Le S , Wu Y , Guo Y , et al. **Game Theoretic Approach for a service function chain routing in NFV with coupled constraints[J]**. *Circuits and Systems II: Express Briefs, IEEE Transactions on*, 2021, PP(99):1-1.

(2) Mitsuru, Toyoda, Yuhu. **Mayer-Type Optimal Control of Probabilistic Boolean Control Network With Uncertain Selection Probabilities.[J]**. I*EEE transactions on cybernetics*, 2019.

https://doi.org/10.17993/3cemp.2023.120151.51-66

(3) Wu Y, Guo Y, Toyoda M. **Policy Iteration Approach to the Infinite Horizon Average Optimal Control of Probabilistic Boolean Networks[J]**. I*EEE Transactions on Neural Networks and Learning Systems*, PP(99):1-15.

(4) Zhang Y, Qian T, Tang W. **Buildings-to-distribution-network integration considering power transformer loading capability and distribution network reconfiguration[J]**. *Energy*, 2022, 244.

(5) Li, H., Deng, J., Feng, P., Pu, C., Arachchige, D., and Cheng, Q., (2021) **Short-Term Nacelle Orientation Forecasting Using Bilinear Transformation and ICEEMDAN Framework**. *Front. Energy Res*. 9, 780928.

(6) Li, H., Deng, J., Yuan, S., Feng, P., and Arachchige, D., (2021) **Monitoring and Identifying Wind Turbine Generator Bearing Faults using Deep Belief Network and EWMA Control Charts**. *Front. Energy Res*. 9, 799039.

(7) Yang Ligong, Tang Shiping, Zhu Jian. **Sentence Love Based on Markov Logic Network Sense Analysis Method [J]**. *Journal of Beijing University of Technology*, 2013, 33(6): 600-604.

(8) Ming Junren. **Research on Text Sentiment Analysis Algorithm Integrating Semantic Association Mining [J]**. *Library and Information Work*, 2012, 56(15): 99-103.

(9) Cai Xiaozhen, Xu Jian, Wu Sizhu. **Research on User Comment Filtering Model for Sentiment Analysis [J]**. *Modern Library and Information Technology*, 2014, 30(4): 58-64.

(10) Tang Xiaobo, Xiao Lu. **Research on Comment Mining Model Based on Sentiment Analysis [J]**. *Intelligence Theory and Practice*, 2013, 36(7): 100-105.

(11) Liu L, Song W, Wang H, et al. **A novel feature-based method for sentiment analysis of Chinese product reviews[J]**. *China Communications (English version)*, 2014, 11(3): 154-164.

(12) Medhat W, Hassan A, Korashy H. **Sentiment analysis algorithms and applications: A survey[J]**. *Ain Shams Engineering Journal*, 2014, 5(4): 1093-1113.

(13) Hochreiter S, Schmidhuber J. **Long Short-Term Memory[J]**. *Neural Computation*, 1997, 9(8): 1735-1780.

(14) Zhou Ying , Liu Yue , Cai Jun . **Sentiment Analysis of Weibo Based on Attention Mechanism [J]**. *Intelligence Theory and Practice*, 2018,41(03):89-94

(15) Fan Qingbo, Jiang Fucai, Ma Quandang, et al. **BP Neural Network-Markov Ship Traffic Prediction Model Based on PSO [J]**. *Journal of Shanghai Maritime University*, 2018, 39(2): 22-27, 54.

(16) Feng Yuxu, Li Yumei. **Research on CSI 300 Index Prediction Model Based on LSTM Neural Network [J]**. *Practice and Understanding of Mathematics*, 2019, 49(7): 308-315.

(17) Zhao Jingpeng. **Research on hierarchical Web service classification method based on deep learning [D]**. *North China Electric Power University (Beijing)*, 2021.

(18) Luo Zhengjun, Ke Mingsong, Zhou Dequn. **Research on Text Sentiment Analysis Model Based on Improved LSTM [J]**. *Computer Technology and Development*, 2020, 30(12): 40-44.

https://doi.org/10.17993/3cemp.2023.120151.51-66

(19) Li Qichao. **Analysis of web data mining technology based on cloud computing environment [J]**. *Information Technology and Informatization*, 2020(09):54-56.

(20) Kong Jia, Deng Sanhong, Zhang Yue, Zhang Yiwei. **Thematic mining of science and technology news in the Web environment: Taking the field of international polar expeditions as an example [J]**. *Science and Technology Management Research*, 2020, 40(17): 173-179.

(21) Liu Yong. **Design and implementation of substation equipment defect management system based on Web technology [D]**. *University of Electronic Science and Technology of China*, 2020. https://doi.org/10.27005/d.cnki.gdzku.2020.004586.

(22) Yang Mengda. **Temperature prediction based on improved PSO-LSTM neural network [J]**. *Modern Information Technology*, 2020, 4(04): 110-112.

(23) Xia Maosen, Jiang Lingling. **Analysis of "Maker" Hot Words Based on Web Text Mining [J]**. *Journal of Hubei University of Science and Technology*, 2019, 39(05): 56-62.

(24) He Yun. **Research on intelligent traffic flow prediction based on PSO-LSTM algorithm [J]**. *Qinghai Transportation Science and Technology*, 2019(04):76-79.

(25) Zhou Zhou. **Analysis and Research on the Curriculum Setting of Secondary Vocational E-commerce Majors Based on Web Text Mining [D]**. *Guangdong Normal University of Technology*, 2019.

(26) Frayssinet, M., Esenarro, D., Juárez, F. F., y Díaz, M. (2021). **Methodology based on the NIST cybersecurity framework as a proposal for cybersecurity management in government organizations**. *3C TIC. Cuadernos de desarrollo aplicados a las TIC*, 10(2), 123-141. https://doi.org/10.17993/3ctic.2021.102.123-141

(27) Che Xiangbei,Li Man,Zhang Xu,Alassafi Madini O. & Zhang Hongbin.(2021). **Communication architecture of power monitoring system based on incidence matrix model**. *Applied Mathematics and Nonlinear Sciences (1)*. https://doi.org/10.2478/AMNS.2021.1.00098.

https://doi.org/10.17993/3cemp.2023.120151.51-66