The Reputation Evaluation Method of E-Commerce Merchants Based on Improved Bilstm Network

Peng Li

Shandong Polytechnic College, Jining ShanDong 272000, China
*1261184813@protonmail.com

Abstract. To address the issue that the semantic and emotional characteristics of consumer reviews are not considered in traditional e-commerce merchant reputation evaluation , this study combines the bidirectional long short-term memory network, the convolutional neural network and the attention mechanism to design a method based on the improved bidirectional long-term memory network. E-commerce merchant reputation evaluation method using short-term memory network. The outcomes indicate that the accuracy of the design method is 4.7%, 10.6%, and 14.9% higher than the bidirectional long short-term memory network, the long short-term memory network, and the convolutional neural network respectively, and the recall rate is 8.6% and 15.5% higher than the other three algorithms. ,18.2%. It indicates that the design method can accurately identify the semantic and emotional characteristics of consumer reviews, which proves its effectiveness and is of great significance to the healthy development of e-commerce platforms.

Keywords: E-commerce; Merchant reputation assessment; Two-way long short-term memory network

1 INTRODUCTION

As the quick growth of the Internet, online shopping has become an indispensable part of people's daily lives. However, the openness of e-commerce platforms has also brought a series of problems to consumers, among which the credibility of merchants is particularly prominent. Merchant reputation not only affects consumers' shopping experience, but also is related to the development of e-commerce platforms [1-2] . Therefore, how to accurately evaluate the credibility of e-commerce merchants is currently a hot topic for relevant professionals. In the last few years, with the growth of deep learning technology, its application in consumer review analysis on e-commerce platforms has become more and more widespread [3-4] . Among them, Long Short-Term Memory (LSTM) is a deep learning model that can process sequence data. It is often used to model contextual information in natural language processing and can well distinguish the emotional polarity of sentences. However, in emotion classification, sentences usually contain all forward and backward information, and it is difficult to capture all information using LTSM [5-6] . In view of this, the study uses Bidirectional long/short-term memory network (BiLSTM) to evaluate merchant reputation,

V. Vasilev et al. (eds.), *Proceedings of the 2024 5th International Conference on Management Science and Engineering Management (ICMSEM 2024)*, Advances in Economics, Business and Management Research 306,
https://doi.org/10.2991/978-94-6463-570-6_135

and introduces convolutional neural network (CNN) and attention mechanism (AM) to optimize the model. A e-commerce merchant reputation evaluation method based on improved BiLSTM is designed to more accurately evaluate merchant reputation and provide consumers with a more reliable shopping reference.

2 E-COMMERCE MERCHANT REPUTATION EVALUATION METHOD BASED ON IMPROVED LONG-TERM MEMORY NETWORK

2.1 Construction of E-Commerce Merchant Reputation Evaluation Index System

In e-commerce, there are various reasons for transaction problems, but the most important factor is the lack of reliable reputation assessment model [7] . Therefore, the study takes Alibaba Group's Taobao e-commerce platform as an example to construct a reputation evaluation index system based on the characteristics of merchants and products on the platform.The constructed index system is shown in Figure 1.

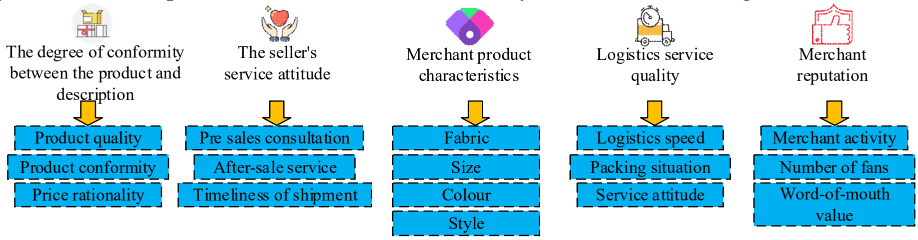


Fig. 1. Accuracy and recall of different methods

The reputation evaluation system includes first-level indicators and second-level indicators. Among them, the first-level indicators include merchant reputation, product characteristics, logistics service quality, seller service attitude, and the degree of compliance between the product and the description. Then, in the Python environment, the obtained data is segmented using jieba word segmentation technology, and the total number of times is calculated. The next step is to use the word frequency-inverse document frequency algorithm to analyze product characteristics. The larger the word frequency-inverse document frequency value, the more important the word is to the document. Finally, the Detail Seller Rating (DSR) method is used to calculate the score of the degree of compliance between the product and the description, the seller's service attitude and the quality of the logistics service . The scoring standard is from 1 to 5, with 5 being very satisfied and 4 being satisfied. 3 points are average, 2 points are dissatisfied, and 1 point is very dissatisfied. The score of each first-level indicator is the average of the total score of its corresponding second-level indicator. The calculation method of merchant reputation is shown in Equation (1).

$$M_R(u) = \omega_1 \left[\sum_{i \in G(u)} M_e(i) \times (M_a(i) + 1) \right] + \omega_2 S_D(u) + \omega_3 F_a(u) \tag{1}$$

In formula (1), $M_R(u)$ indicates u the reputation of the merchant, $F_a(u)$ indicates the number of fans $M_e(i)$ of the merchant, represents the word-of-mouth value of store dynamics, $S_D(u)$ represents the total number $M_a(i)$ of store dynamics, represents the activity of store dynamics, G_u represents the set of store dynamics, ω_1, ω_2, ω_3 represents the value of word-of-mouth, the number of dynamics and The weight of the amount of fans. Since users have emotional tendencies when commenting on merchants, the word-of-mouth value of a merchant needs to be calculated based on the emotional value of the comments. The specific calculation method is shown in equation (2).

$$M_e(u) = \omega_4 S_u + \omega_5 (CS_u) C_u \tag{2}$$

In formula (2), S_u represents the total amount of likes for a certain activity in the store, C_u represents the total amount of comments for a certain activity in the store, CS_u represents the emotional value of the comments, ω_4 indicates the weight of the number of likes, and ω_5 indicates the weight of the number of comments.

2.2 Reputation Emotion Quantification Model Based on Improved Long-Term Memory Network

In reputation evaluation, some emotional words may reflect different emotions, whether positive or negative, when used in different sentences [8]. If an emotional dictionary is used to directly calculate the emotional value of a merchant's reputation, there will be a certain degree of error in the final merchant's reputation rating[9-10]. To obtain accurate emotional expression of a certain word in different aspects, this study applies deep learning technology to emotional feature extraction in merchant reputation evaluation. Using the Word2vec model to convert saved product review text data into word vectors, and then combining BiLSTM with Convolutional Neural Networks (CNN), CNN is used to extract the main features of the text data. BiLSTM receives the features output by CNN and learns them. At the same time, an attention mechanism is introduced to calculate the optimal emotional value weight, and an improved LSTM based emotional quantification model is designed. The implementation process is shown in Figure 2.

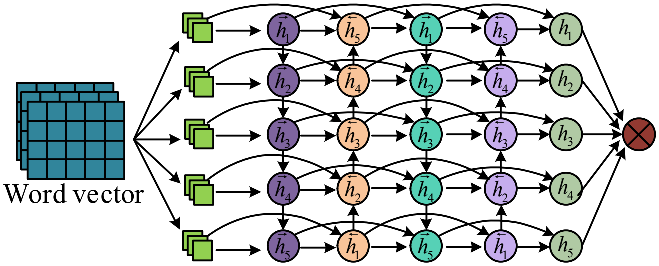


Fig. 2. A sentiment quantification model based on improved LSTM

In the improved LSTM based merchant word-of-mouth sentiment quantification model, first vectorize the input merchant word-of-mouth text data in the Word2vec model, the calculation method is as shown in Equation (3).

$$L = \prod_{t=1}^{T} \prod_{-c \le j, j \ne 0} P(w_{t+j} \mid w_t) \tag{3}$$

In formula (3), T represents the text length, c represents the size of the Word2vec model training window, w_t represents the center word, and w_{t+j} represents w_t the word after the center word j. Due to the different dimensions of merchant reputation characteristics, it is necessary to perform dimensionless processing on the data. The calculation method is as shown in Equation (4).

$$w = (w - w_{\min}) / (w_{\max} - w_{\min}) \tag{4}$$

In formula (4), $w_{\max}$, $w_{\min}$ represent the max and mini values of the merchant's word-of-mouth feature word vector respectively. Input the obtained word vector into a convolutional neural network and extract the features of the word vector using convolutional kernels of different sizes. Then input the extracted features into the BiLSTM layer to learn the feature vectors. In the final data feature vector obtained, some data is useless, so it needs to be distinguished. The study introduces an AM and adds it after the BiLSTM network to identify feature vectors. At this time, the feature vector sequence output by the BiLSTM network $(w_1, w_2, ..., w_n)$ is the input of the attention layer, which is encoded by the encoder, and a semantic coding sequence is introduced to perform nonlinear transformation on it, and then the decoder is used to convert the nonlinearly transformed feature vector Sequence decoding is used to calculate the output of the merchant's word-of-mouth feature vector sequence. The calculation method is shown in Equation (5).

$$t_m = x_j(w_1, w_2, ..., w_n) \tag{5}$$

In formula (5), t_m refers to the output sequence and x_j refers to the semantic encoding. The value of the semantic encoding vector is related to the input latent vector of the attention layer. The latent vector is obtained by using the encoder to map the merchant's word-of-mouth feature data to the current position. All information of the merchant's word-of-mouth feature data is reflected in the latent vector. Different latent vectors represent different surrounding information of concern. The calculation method of the semantic encoding vector As shown in equation (6).

$$x_j = \sum_{j=1}^{T} \mu_{t,j} \times h_j \tag{6}$$

In Formula (6), $\mu_{t,j}$ represents the th t contribution rate h_j at time and j represents the th j input latent vector of the attention layer. The next step will be to classify the merchant reputation feature data through the Softmax classifier in the fully connected layer, and the activation function used is the Sigmoid function. The calculation method for the emotion classification results of consumers' reviews of merchants is shown in Equation (7).

$$V_i = Softmax(S_i \times h + b_z) \tag{7}$$

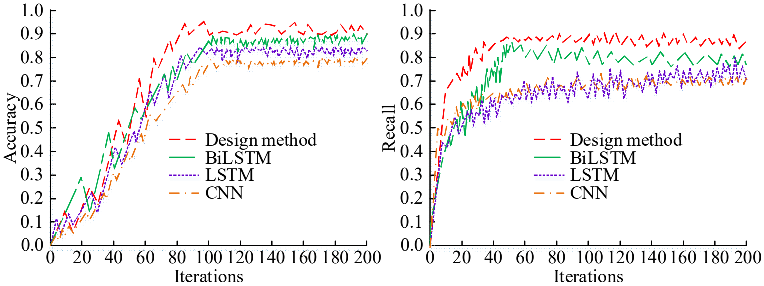
In Formula (7), V_i represents the emotion classification label, b_z represents the bias, and h represents the weight. Finally, the merchant reputation value is calculated through the obtained features. The specific calculation method is shown in Equation (8).

$$C_m = \omega_1 I_d + \omega_2 S_a + \omega_3 S_q + \omega_4 M_c + \omega_5 M_R \tag{8}$$

In formula (8), C_m means the reputation value, I_d represents the degree of compliance of the product with the description, S_a represents the seller's service attitude, S_q represents the logistics service quality, M_c represents the product characteristics, and M_R represents the merchant's reputation. To reduce model complexity, the study first learns its connections through normal network training, then modifies small weight connections, and finally retrains the network to learn the final weights of the remaining connections, improving computational efficiency. At the same time, the research uses the Tableau visualization analysis platform to display the decision-making process and feature weights of the model, so as to have a more intuitive understanding of the working process of the model.

3 ANALYSIS ON THE EFFECTIVENESS OF E-COMMERCE MERCHANT REPUTATION EVALUATION METHOD BASED ON IMPROVED LONG SHORT-TERM MEMORY NETWORK

In order to verify the effect of the designed credit evaluation method of e-commerce merchants based on improved long and short term memory network, the research was carried out in the operating system with Intel Core i7-8565U CPU, 16GB running memory, Intel UHD Graphics 620 independent graphics card and Windows 10. Simulation experiments were performed using Python 3.6.5. The learning rate of the network model is set to 0.0001, the batch size is 256, the convolution kernel size is 55, and the maximum number of iterations is 200. The data set adopted is the Taobao comments obtained by crawler technology and processed by a series of processes. First, two indicators of recall rate and accuracy are introduced to train the design model on the obtained data set, and the two index values are calculated to verify the performance of the model. The results are compared with CNN, LSTM and BiLSTM, as shown in Figure 3.



(a) Accuracy of different methods (b) Recall rates for different methods

Fig. 3. Accuracy and recall of different methods

Figure 3 (a) shows the variation of accuracy with iteration times for the design method, BiLSTM, LSTM, and CNN methods, while Figure 3 (b) shows the variation of recall rates with iteration times for the four methods. As can be seen from Figure 3(a), compared with BiLSTM, LSTM, and CNN, the accuracy of the design method has increased by 4.7%, 10.6%, and 14.9% respectively. From Figure 3(b), the recall rates of the four methods are 89.1%, 80.5%, 73.6%, and 70.9% respectively. The recall rate of the design method is significantly higher than other methods. The above outcomes denote that the precision rate and recall rate of the design method are high, which proves that it has a better recognition effect on merchant reputation characteristics and proves its effectiveness. To further validate the performance of the design method, two different e-commerce book datasets, Fashion MINIST and Online Auctions, were introduced to train the model. F1, accuracy, and runtime were calculated and compared with the results of CNN, LSTM, and BiLSTM. The results are shown in Table 1.

Table 1. F1, accuracy, and runtime of each model on different datasets

Model	F1	Precision	Run time (s)
Design method	0.947	0.984	133
BiLSTM	0.935	0.917	135
LSTM	0.871	0.872	186
CNN	0.846	0.9833	197

From Table 1, it can be seen that the F1 value, accuracy, and running time of the design method are 0.947, 0.984, and 133 seconds, respectively, which are significantly higher than the three indicator values of the other three methods, indicating the recognition accuracy and computational efficiency of the design method. At the same time, it proves that its generalization ability is good and its robustness is high. To verify the effect of the design method in practical applications, the study randomly selected 10 merchants of the same type from Taobao, making the 10 merchants respectively merchant 1 to merchant 10, and used the design method to calculate the reputation score of the merchants. The calculation results are as follows: As shown in Figure 4.

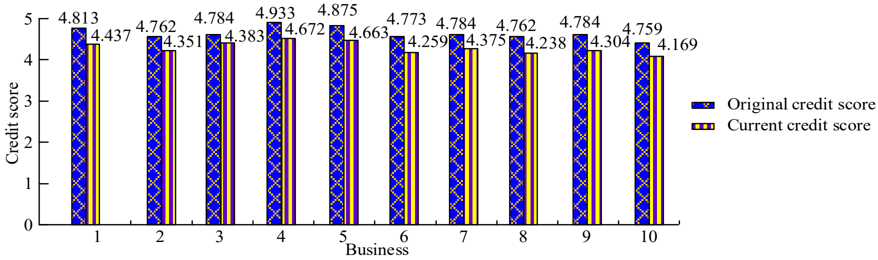


Fig. 4. Credit ratings of different merchants

As can be seen from Figure 4, in the original merchant reputation score, merchant 2 and merchant 8 have the same reputation score, both 4.762 points. Merchant 3, Merchant 7, and Merchant 9 all have reputation scores of 4.784. When consumers shop at these merchants, it is difficult to make shopping decisions based on reputation scores. In the current merchant reputation scores, the credit scores of 10 merchants are different, and consumers can make decisions faster based on the reputation scores. It further proves that the design method can more accurately calculate merchant reputation scores, improve consumers' shopping experience, and reduce transaction risks.

4 CONCLUSION

In order to more accurately reflect the real reputation of merchants, the research first establishes an indicator system according to the characteristics of merchants and commodities, then combines LSTM with CNN, and introduces the attention mechanism to design an emotion analysis algorithm to quantitatively process the acquired merchants and commodities. The outcomes denote that the accuracy rates of the design method, BiLSTM, LSTM and CNN are 94.3%, 89.6%, 83.7% and 79.4% respectively, and the recall rates are 89.1%, 80.5%, 73.6% and 70.9% respectively. The accuracy and recall rate of the design method are significantly higher than the other three methods, which identify that the feature recognition effect of the design method is better. However, the study only used data from one e-commerce platform for testing, and the data was relatively single. Future research will use data from more sources for reputation assessment to verify the universality of the design method.

REFERENCES

1. Gonçalves M J A, Pereira R H, Coelho M A G M. User Reputation on E-Commerce: Blockchain-Based Approaches[J]. Journal of Cybersecurity and Privacy, 2022, 2(4): 907-923.
2. Salamai A A, Ageeli A A, El-Kenawy E M. Forecasting e-commerce adoption based on bidirectional recurrent neural networks[J]. Computers, Materials & Continua, 2022, 70(3): 5091-5106.
3. Pranto T H, Noman A A, Rahaman M, et al. A blockchain, smart contract and data mining based approach toward the betterment of e-commerce[J]. Cybernetics and Systems, 2022, 53(5): 443-467.
4. Hu J. Research on Credit Rating of E-commerce Sellers based on K-means Algorithm-Taking Taobao Children's Clothing Store as an Example[J]. International Core Journal of Engineering, 2022, 8(10): 230-236.
5. Alsharif N. Fake opinion detection in an e-commerce business based on a long-short memory algorithm[J]. Soft Computing, 2022, 26(16): 7847-7854.
6. Kausar M A, Fageeri S O, Soosaimanickam A. Sentiment Classification based on Machine Learning Approaches in Amazon Product Reviews[J]. Engineering, Technology & Applied Science Research, 2023, 13(3): 10849-10855.
7. Ramadhan F A, Ruslan R R P, Zahra A. Sentiment Analysis Of E-Commerce Product Reviews For Content Interaction Using Machine Learning[J]. Cakrawala Repositori IMWI, 2023, 6(1): 207-220.
8. Yang Q. Research on e-commerce customer satisfaction evaluation method based on pso-lstm and text mining[J]. 3c Empresa: investigación y pensamiento crítico, 2023, 12(1): 51-66.
9. Karn A L, Karna R K, Kondamudi B R, et al. Customer centric hybrid recommendation system for E-Commerce applications by integrating hybrid sentiment analysis[J]. Electronic Commerce Research, 2023, 23(1): 279-314.
10. Venkatesan R, Sabari A. Deepsentimodels: A novel hybrid deep learning model for an effective analysis of ensembled sentiments in e-commerce and s-commerce platforms[J]. Cybernetics and Systems, 2023, 54(4): 526-549.