Research on Analyzing the Emotional Polarity of Malicious Swipe Comments on E-commerce Platforms Based on NPL

Chaoyi Ren

College of Software, Taiyuan University of Technology, 030600, Jinzhong, Shanxi, China

Abstract. In the era of rapid advancements in natural language processing (NLP) models, these technologies have immense potential to detect and address societal issues, enhancing the functioning of the digital society. Online shopping platforms rely heavily on user reviews to influence buyer decisions, yet malicious reviews can significantly degrade user experience. This study focuses on analyzing the emotional polarity of malicious brush-order (falsely generated) reviews in e-commerce product comments, utilizing the Jingdong product review dataset. The methodology involves utilizing the Word2Vec model to vectorize the text data, followed by principal component analysis (PCA) for outlier detection to identify potential malicious reviews based on their unique characteristics. The PCA results are further leveraged for dimensionality reduction, simplifying the dataset. Subsequently, the BERT model is employed to perform semantic similarity analysis, allowing for the screening and expansion of the experimental dataset with similar malicious comments. This enriched dataset is then subjected to sentiment polarity analysis, enabling a deeper understanding of the nature and impact of these malicious reviews. By facilitating buyers in making informed decisions based on genuine reviews, this research underscores the practical value of NLP in addressing real-world challenges in e-commerce.

1. Introduction

China's network society is developing rapidly, and algorithm is the underlying structure of the Internet operation mechanism. Literature [1] mentions that if the algorithm is abused or manipulated, it will have a negative impact on the society. Literature [2] mentions that on the one hand, the company needs positive comments from users, and on the other hand, it needs to maintain the legitimacy of user comments. This approach ensures that the authenticity of user feedback is maintained, fostering trust within the digital marketplace and enhancing the overall quality of decision-making processes for both consumers and businesses alike.

There are already applied studies on e-commerce reviews based on sentiment analysis in China. In this study [3], Work2vec technology is used to make an emotion dictionary applicable to the linguistic environment of online commodity reviews, and the sentiment

Corresponding author: renchaoyi4941@link.tyut.edu.cn

values of corresponding emotion words and the weights of degree words are updated, and the scores of a certain review are obtained by multiplicative of the weights. The reference value of their comments can be seen by comparing the scores of their comments.

Related studies on the use of text similarity to evaluate emotional blackmail discourse [4] mentioned an evaluation method based on parts-of-speech semantic similarity: multiple adjacent parts of speech are formed into a subsequence, and the pronoun with the smallest sum of distance from each part of speech in the word subsequence is defined as the nearest pronoun in the word subsequence. If the nearest pronoun contains a specific pronoun, use 1 to indicate the nearest pronoun; Otherwise use 0 for the nearest pronoun. This method can find out the pronoun related to the part of speech of the subsequence, which can expand the experimental data.

Bidirectional Encoder Representations from Transformers (BERT) is based on Transformers model [5]. Compared with the simpler model of Generative Pre-Trained 1.0 (GPT1.0), the self-attention mechanism of this model can effectively process data sets in parallel. However, there are many models based on BERT model that are better than those dealing with sentiment analysis problems. For example, BERT-MIFN [6], an aspect level sentiment analysis model based on multi-information fusion based on BERT, can use dual paths to carry out multi-information fusion and overall semantic extraction respectively, thus reducing the impact of the same words on text semantics and enhancing the diversity of data sets.

Combined with the above research, this experiment first uses Work2vec to vectorize the data, standardizes the text samples, and creates a basic environment for principal component analysis (PCA). PCA was used to detect outliers, detect vectors far from the offset data gathering point, find their samples, and make statistics as feature vectors. The existing text similarity model is used to compare the feature vector with other samples, and the samples with high similarity are screened to further expand the data set of feature samples. The expanded feature vector data set was used to train the BERT sentiment analysis model [7], and the range of data points with high aggregation of sentiment values was obtained.

2. Construction of commodity review data set on e-commerce platform

Before conducting the experiment, it need a certain amount of data for reference. Here, Jingdong product web review data set is used, and Scrapy framework is used to implement the data crawling of Jingdong product reviews. According to literature [8], it can be found that in the data set used in this literature experiment, negative comments account for 39% of the total comments, while the number of positive comments considered useless by users is about ten times that of negative comments. However, there is no clear judgment criteria for the identification of malicious swipe reviews, and different cognition will result in different characteristics of the data set. This study aims to objectively identify invalid malicious swipe reviews that affect users' judgment of product quality.

In order to ensure the objectivity of the research, malicious single comment is used to affect the correct judgment of e-commerce users on the product, which is a small number of special characteristics. The data set is analyzed by anomaly, the vector that deviates from the sample gathering point is found, and its text is recorded as the characteristic value.

3. Malicious swipe review feature sample screening

Firstly, vectorial preprocessing of text is performed using Work2vec technology. Work2vec is a deep learn-based tool used to calculate the similarity of word vectors. Proposed by

Google in 2013 [9], it can convert words into vectors. This transformation not only preserves the semantic relationship between words, but also enables similar or related words to have similar positions in the vector space. Here, it does not use this technology to predict synonyms of Words, but only use the Continuous Bag of Words (CBOW) model of work2vec technology to vectored the text.

PAC is used to detect outliers [10]. PCA technology can reduce the dimensionality of text vectors and make the features of feature samples independent of each other, so as to facilitate the separation of outliers and take the separated data points as feature samples.

The PCA technique first calculates the covariance matrix of the standardized data, which describes the linear correlation between the data features, and then performs the eigenvalue decomposition of the covariance matrix to obtain the eigenvalues and corresponding eigenvectors. The eigenvector represents the principal component direction of the data in the new coordinate system, while the eigenvalue represents the variance of the data in the corresponding eigenvector direction. The eigenvalues are sorted in order from largest to smallest, and the first k largest eigenvalues and their corresponding eigenvectors are selected as the principal components. The original data is projected onto the selected k principal components to obtain the data set after dimensionality reduction. The original data is projected onto the selected k principal components to obtain the data set after dimensionality reduction. Calculate the difference between the original data and the reconstructed data, that is, the reconstruction error. For most normal data, reconstruction errors are relatively small because they are more consistent with the overall data distribution. As for the outliers, because they are different from the overall data distribution, the reconstruction error will be relatively large. The outliers are identified according to the size of the reconstruction error. It is usually possible to set a threshold and treat data samples with a reconstruction error greater than this threshold as outliers.

The BERT model was used to compare the semantic similarity between feature samples and other samples. The feature samples were taken as independent variables and the other samples as dependent variables. The cosine similarity of their vectors was compared, and the recorded data was shown in Table 1.

Table 1. Semantic similarity ratio

independent variable	dependent variable	similarity ratio
really speechless	Die speechless	95.94%
This thing ran out of power in the morning, and it didn't have much else	It doesn't work very well, it only has a morning's worth of power and the model is pretty average	54.55%
Don't think about it. It's good	Not recommended to buy, cost-effective is not high	0.00%

It can be found that in a text content, when there are the same words, the similarity will be relatively high, and the impression factors of semantic bias such as word count, key words, and format, etc., but the semantic similarity detected by samples with the same sentence pattern but opposite semantics is 0.00%. Considering that there are many factors affecting semantic similarity and the model scores are polarized, the paper extracted samples with semantic similarity exceeding 50.00% and combined the newly extracted samples with the independent variable samples in the above content as new feature samples. The purpose of this step is to avoid that PCA anomaly detection can only screen out too small number of samples, resulting in increased chance of the experiment. Use semantic similarity to filter more samples and expand the dataset for study.

4. Emotion analysis of feature samples

Using the data set obtained in the above experiment, BERT model is used to analyze the data set, and the model judges the emotional polarity of the text sample with scores between [-1, 1], where the positive and negative values represent positive and negative emotions, and the absolute values represent the intensity of emotions. The greater the absolute value, the stronger the emotion of the text. On the contrary, the smaller the absolute value, the more emotionally stable the text. Record the emotion polarity value corresponding to the sample and observe its corresponding emotion value, as shown in Table 2.

Table 2. Emotional polarity values of feature samples

eigenvalue	Affective polarity value
Speechless dead.	-0.92
It doesn't work very well, it only has a morning's worth of power and the model is pretty average.	-0.93
Don't think about it. It's good.	+0.85
This thing works really well!	+0.87
There are some flaws in the battery but they are negligible.	+0.84
Cost-effective, good value for money, with some flaws	+0.95

Due to the immaturity of the experimental model, there are errors in the experimental data, and the comments that are subjectively not malicious in the data set are also screened and the emotional polarity value is close to 1. For example, in Table 2, "Cost-effective, good value for money, with some flaws" The affective polarity reached +0.95, and the statistical data was removed at the end of the experiment to reduce the influence of the error value on the experimental results.

By statistical analysis of the emotional polarity values obtained from the above experiments, it is found that the emotional polarity values of most characteristic samples are close to -1 and 1, and a few samples are relatively close to -0.5 and 0.5. The value range of negative emotion polarity is [-0.98, -0.90], and the value range of positive emotion polarity is [0.78, 0.90].

5. Conclusion

Through the experiment, it can be found that the emotion of malicious comments is very distinct, but it is not negative emotion or positive emotion in common sense, but there are two extreme emotional characteristics of negative emotion and positive emotion at the same time. Such comments are brief and have distinct emotional characteristics, and they lack the objectivity that should be paid attention to when commenting on products.

There are also aspects that can be optimized and improved in this experiment. First, the data set of this experiment can be further expanded, and the product review data set of Jingdong e-commerce platform used in this experiment can be further expanded to the product review data set of multiple e-commerce platforms. Different platforms have different user composition, which may lead to large deviations in the data set. The language processing experiment has many variables and is not easy to control. With the iteration and upgrade of the model in the future, the sentiment analysis of malicious brushing comments will be more accurate.

According to the latest report of the Cyberspace Administration of China, the current development of the Internet should clarify the value orientation and behavior orientation of

Internet users, as well as the management system of the form and content of comments. At present, the major e-commerce platforms do not have much supervision on shopping reviews, especially malicious swipe reviews will cause bad social problems, which will constitute bad emotional consumption guidance for buyers by merchants, but the supervision completely depends on user reports and platform review, and the review of comments will also consume a lot of manpower. Therefore, this study can be used as a reference for society and enterprises, and this experimental method can be used to deal with malicious comments and create a good network environment.

References

1. H. L. Cai, Media availability horizon, trill electricity live emotional communication research, (The northeast university of finance and economics, 2023)
2. C. Y. He, Research on User Malicious review Detection based on Machine learning, Beijing University of Posts and Telecommunications (2019)
3. C. Y. Lu, Z. Z. Dai, Y. H. Wang, Application research of e-commerce platform review based on Sentiment analysis. Science and Technology Entrepreneurship Monthly, 36(10):150-154, (2023)
4. W. S. Lin, G. C. Yang, S. H. Zhong, Based on Chinese text similarity assessment of emotional blackmail discourse system, (Computer application research, 1-9, 2024)
5. T. Y. Lin, Y. X. Wang, X.Y. Liu, X. P. Qiu, A survey of transformers, AI Open, Volume 3, Pages 111-132, ISSN 2666-6510, (2022)
6. X. F. Zhang, H. C. Tao, Multi-information Fusion aspect level Sentiment Analysis Model based on BERT, (Journal of Chengdu University of Information Technology, 2024)
7. A. Chinnalagu and A. K. Durairaj, "Comparative Analysis of BERT-base Transformers and Deep Learning Sentiment Prediction Models" 2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART), Moradabad, India, (2022)
8. M. Y. Lv, X. L. Wang, K. X. Li, Negative comments to guide the higher dimensional multivariate data visualization analysis system for, (Journal of computer-aided design and graphics, 1-11, 2024)
9. Y. Yao, X. Li, X. Liu, P. Liu, Z. Liang, J. Zhang, K. Mai, Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model, *International Journal of Geographical Information Science*. **31**, 4, 825–848, (2017)
10. A. Maćkiewicz, W. Ratajczak, Principal components analysis (PCA), Computers & Geosciences, Volume **19**, 3, 303-342, (1993)