ISSN 2616-5902 Vol. 6, Issue 4: 276-281, DOI: 10.25236/AJBM.2024.060440

Research on the identification method of fake reviews in e-commerce

Daming Fan

School of Communication, Qufu Normal University, Rizhao, 276826, China

Abstract: *With the development of e-commerce, many merchants, to lure consumers to buy their own goods, will create the illusion of hot stores and excellent goods by means of false reviews. To identify false reviews, this research paper proposes a false review identification method, which mainly uses entropy weight TOPSIS model, K-Means cluster analysis and logistic regression. The article selects evaluation indexes such as sentiment polarity, text length, review usefulness and rating deviation degree, and calculates the review score by entropy weighting method of TOPSIS model. Subsequently, K-Means clustering was used to categorize the reviews into two groups: customer reviews and machine reviews and validated by logistic regression. The experimental results show that the method has a good ability to recognize false reviews with high accuracy, recall, precision, and AUC value. Taken together, the method provides an effective solution for false comment recognition with practical application potential.*

Keywords: *Fake reviews, TOPSIS model, Cluster analysis, Logistic regression*

1. Introduction

The rise of online shopping, "China Internet Development Status Statistical Report" shows that as of June, China's online shopping user scale reached 884 million people, an increase of 38.8 million people in December 2022, accounting for 82.0% of the overall netizens [1], online shopping has become a common behavior, the commodity reviews have become one of the user's reference bases for shopping, and positive reviews can stimulate the consumer to buy, while negative reviews Positive comments can stimulate consumers to buy, while negative comments can weaken consumers' willingness to buy. To build a good sales brand, many merchants often use false reviews to fake a good sales image. False reviews have a direct impact on users' decision-making and interfere with their willingness to buy.

Currently in the era of big data, natural language processing technology is gradually maturing, the use of artificial intelligence to mine the value of the existence of text data, which helps to analyze the intrinsic information of the text, and the identification of false online shopping reviews has become possible. The main methods for identifying false reviews of goods can be divided into the detection of false review text, the detection of false reviewers, and the detection of false review groups.

In terms of false review text detection, Dongwei Cao et al. proposed a false review detection method based on fused semantic similarity graph convolutional network [2], and Yunmei Shi et al. used text convolutional neural network and BERT pre-training model, multimodal fusion of textual features and visual features, to detect false reviews [3]. Many deep learning detection methods have emerged that do not rely on manual feature engineering and show superiority.

In false commenter detection, Juanjuan Hsing selects comment content features and commenter behavioral features, and then, defines first-order logic predicates and logic formulas based on the features, and introduces the process of weight learning and inference [4]. Many scholars use historical data to mine the potential information of user comments, and analyze the features related to the user's historical comments to determine whether they are false users.

In terms of false comment group detection, a spectral clustering group detection algorithm based on commenter similarity matrix is proposed by Ye Zicheng et al [5], and Zhang Qi et al. propose a sailor group detection algorithm through the selection of features of sailor group counterfeiting behaviors, the construction of comment graphs with weights, the screening of suspicious subgraphs, and clustering of sailor groups based on community discovery algorithms [6]. For the research on false comment group detection, most scholars use frequent item mining and clustering algorithms to obtain candidate groups and manually label them.

ISSN 2616-5902 Vol. 6, Issue 4: 276-281, DOI: 10.25236/AJBM.2024.060440

In summary, different feature mining algorithms are used to select recognition methods based on different research scenarios. There is a missing dataset of false reviews, and although deep learning does not need to distinguish whether the data is labeled or not, the interpretability is poor, and the amount of required data is huge.

Taking Amazon product reviews as a dataset, the method of text detection is chosen. The TOPSIS model and K-Means clustering model are used to quantitatively analyze the reviews, distinguish customer reviews and machine reviews based on the comprehensive quantitative results, label the unlabeled data, and construct a logistic regression model to train the model with the features of sentiment polarity, usefulness of the reviews, degree of deviation of the ratings, and length of the text to test the reasonableness of the evaluation criteria, which provides a user-friendly way of distinguishing false reviews.

2. Entropy-weighted TOPSIS-K-mean clustering model

Entropy weighting TOPSIS model is a multi-criteria decision-making technique designed to help decision makers rank and select alternatives under multiple evaluation metrics. Entropy is a concept in information theory used to measure uncertainty. The weights of the factors are calculated to quantify the metrics to determine the scoring criteria for each metric to obtain a composite scoring criterion.

The TOPSIS method is a commonly used multi-criteria decision-making method that evaluates the proximity of alternatives to the ideal solution. Based on the weights of the indicators and the evaluation values of the alternatives, the distance of each alternative from the ideal solution is calculated, and then ranked according to the size of the distance, with smaller distances indicating that the alternative is closer to the ideal solution, i.e., better.

K-Means is a commonly used cluster analysis algorithm that is widely used in data mining and machine learning tasks. The goal is to divide a set of data points into a predetermined number of K clusters (clusters) such that each data point belongs to its nearest cluster.

Logistic regression is a commonly used classification model for establishing associations between input features and binary categories (e.g. yes/no, success/failure, etc.). The binary nature of the dependent variable has been applied a powerful tool in multivariate techniques. The core idea is to convert the results of linear regression into probabilities by using logistic functions (also known as sigmoid functions) and further make classification decisions.

2.1 Evaluation indicators

To determine whether a product review is a customer review or a machine review, 1000 Amazon reviews were randomly selected as training data. Analyze the raw data of the subject, and each comment contains data information such as reviewerID, asin, reviewer Name, helpful, review Text, overall, summary, unixReviewTime, and review Time. After mining the original data, the evaluation indicators selected for the model are as follows.

The sentiment of machine reviews will be more polarized to a certain extent compared to the sentiment of customers.

Text length: the level of detail and breadth of content of the review. Longer reviews provide more information, shorter reviews may only contain a summary of the evaluation. Customers tend to make comprehensive considerations after using a product, analyzing the experience of using the product from various aspects. The longer the text length, the higher the probability that the review is a customer review.

Review usefulness: Statistics on the usefulness of reviews provide insights into the extent to which reviews are recognized by other users and influence purchasing decisions. Customers tend to post reviews after intensive use of the product, and a higher usefulness may mean that the review is informative to other users and is more likely to be a customer review.

Degree of rating bias: Degree of rating bias is the degree to which an individual review's rating differs from the overall rating. A large deviation means that the review's rating of the product is significantly different from the overall rating of other users, and a higher deviation indicates a more biased sentiment.

ISSN 2616-5902 Vol. 6, Issue 4: 276-281, DOI: 10.25236/AJBM.2024.060440

2.2 Entropy weight TOPSIS modelling

(1) Standardization and normalization

In the case of data with different scales, standardization and normalization avoid the effect of scale and allow data from different units to be compared and analyzed.

(2) Determine the weights of indicators

The TOPSIS entropy weight method is used as the evaluation model. Entropy weight formula:

$$w_i = \frac{1 - \frac{H_i}{\ln n}}{n - \sum_{j=1}^{n} \frac{H_j}{\ln n}} \tag{1}$$

Information entropy formula in entropy weight method:

$$H_i = -\Sigma_{j=1}^{m} \frac{p_{ij}}{\ln p_{ij}} \tag{2}$$

(3) Calculate the positive ideal solution distance D+, negative ideal solution distance D-

Define the matrix as Z, find the optimal and inferior solutions, and calculate the positive ideal solution distance D+ and negative ideal solution distance D-. Based on the normalized decision matrix, calculate the distance to the positive and negative ideal solutions for each evaluation object and weight them. Where ω_j represents the weights.

$$D_i^+ = \sqrt{\sum_{j=1}^{m} \omega_j \left(z_j^+ - z_{ij}\right)^2} \tag{3}$$

$$D_i^- = \sqrt{\sum_{j=1}^{m} \omega_j \left(z_j^- - z_{ij}\right)^2} \tag{4}$$

(4) Composite Score Calculation

$$S_i = \frac{D_i^-}{D_i^- + D_i^+} \tag{5}$$

(5) K-Means clustering

Divide the given dataset into K clusters and define the loss function:

$$J(c,u) = \min \sum_{i=1}^{M} \left\| x_i - u_{c_i} \right\|^2 \tag{6}$$

Repeat the following process in iterations, for each sample, assigning it to the nearest center:

$$c_i^t < -\operatorname{argmin}_k \left\| x_i - \mu_k^t \right\|^2 \tag{7}$$

For each class center k, recalculate the center of the class:

$$\mu_i^t < -\operatorname{argmin}_\mu \sum_{i:c_i^t = k}^{b} \left\| x_i - \mu \right\|^2 \tag{8}$$

(6) Logistic regression

ISSN 2616-5902 Vol. 6, Issue 4: 276-281, DOI: 10.25236/AJBM.2024.060440

Analyzing the linear relationship between the independent and dependent variables can be expressed as:

$$\hat{y} = \omega^T x + b \tag{9}$$

The interval to which the output value is mapped by the sigmoid activation function. The mathematical formulation of the sigmoid function is:

$$\sigma(z) = \frac{1}{1+e^{-z}} \tag{10}$$

2.3 Review Identification and Evaluation Tests

(1) Machine comment recognition

The weights are obtained by entropy weighting method and substituted into TOPSIS model to calculate the score of each review, which is used to find the threshold between the customer review score and the machine review score. The positive and negative ideal solutions are shown in Table 1:

Table 1: Positive and negative ideal solutions

term (in a mathematical formula)	positive ideal solution	negative ideal solution
Comments on usefulness	0.99990002	0.00009998
Text length	0.99999999	1e-8
emotional polarity	0.99989532	0.00010468
Degree of deviation of ratings	0.99996667	0.00003333

After calculating the scores of some of the reviews, K-Means was used as a clustering model and the category clusters were set to 2 i.e., they were divided into customer reviews and machine reviews and customer reviews were labeled as 1 using category 1 and machine reviews were labeled as 0 using category 2. These labels were used for logistic regression validation. Some of the results are shown in Table 2:

Table 2: Partial comment markers

emotional polarity	Degree of deviation of ratings	Text length	Comments on usefulness	Composite score index	markings
1	1	7351	0.961538	0.972174	1
1	0.672	7297	1	0.921572	1
1	1	4656	1	0.791878	1
1	0.672	3990	1	0.735336	1
1	0.672	3419	0.863636	0.667708	1
1	1	2595	1	0.666369	1
1	1	2536	1	0.663248	1
1	0.672	1221	1	0.590571	0
1	1	923	1	0.588315	0
1	0.672	1143	1	0.587308	0

Table 3: Cluster Analysis

	Clustering categories (mean ± standard deviation)		F	P
	Category 1 (n=882)	Category 2 (n=118)		
emotional polarity	0.897±0.174	1.0±0.0	41.889	0.000***
Degree of deviation of ratings	0.673±0.252	0.643±0.296	1.354	0.245
Text length	374.863±346.499	2681.466±1232.72	1948.815	0.000***
Comments on usefulness	0.269±0.376	0.562±0.368	63.209	0.000***
Composite score index	0.282±0.15	0.493±0.17	198.801	0.000***

Note: ***, **, * represent 1 per cent, 5 per cent and 10 per cent significance levels, respectively.

The clustering results were divided into two categories, with a frequency of 882 and a percentage of

ISSN 2616-5902 Vol. 6, Issue 4: 276-281, DOI: 10.25236/AJBM.2024.060440

88.2% for clustering category 2 and a frequency of 118 and a percentage of 11.8% for clustering category 1. The result of the analysis is that category 2 is real comments. The clustering results are shown in Table 3:

Analysis of clustering effects. As shown in Table 4:

Table 4: Evaluation indicators

Contour Coefficient	DBI	CH
0.778	0.52	1948.813

The analysis results show that the clustering effect is better when the contour coefficient is close to 1. DBI (Davies-bouldin) is the ratio of intra-cluster distance to inter-cluster distance for any two clusters, the smaller the DBI, the better the clustering effect. Using CH (Calinski-Harbasz Score): the affinity of the class is measured by calculating the sum of squares of the distances between the points within the class and the center of the class (denominator), and the separateness of the dataset is measured by calculating the sum of squares of the distances between the center of mass of the class and the center of mass of the dataset (numerator), and CH metrics are obtained from the ratio of the separateness to the affinity, and the bigger the CH is, the better the clustering is.

The results show that when sentiment polarity, text length, comment usefulness, and the degree of scoring deviation are selected as evaluation indicators, after the scores are obtained by using the entropy-weighted TOPSIS model evaluation, the comments with scores lower than 0.607916447 have a high probability of being machine comments, and there are only a small number of customer comments with scores below 0.607916447 and below totaled 965 reviews for the 1000 reviews, of which only 83 customer reviews existed, a mere 8.601%.

(2) Results testing

The model is trained using labeled samples and then predicted using unlabeled datasets. The validity of the evaluation criteria was verified by calculating the prediction accuracy of the model. In the test, only three of the evaluation metrics were selected as features of the logistic regression in order to test the robustness of the model. They are the degree of rating bias, sentiment polarity, and review usefulness.

The ROC curve was used to measure the classification effect of logistic regression. As shown in Figure 1:

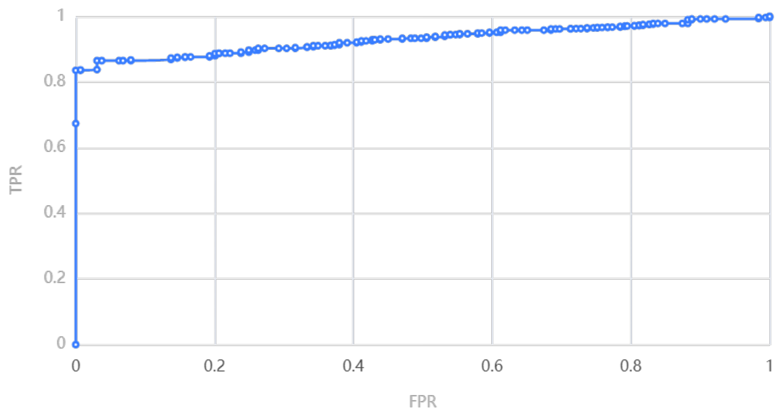

Figure 1: ROC curve.

Measuring the Predictive Effectiveness of Logistic Regression. As shown in table 5:

Table 5: Logistic regression effects.

accuracy	recall rate	accuracy	F1	AUC
0.89	0.89	0.91	0.889	0.93

The accuracy of 0.89 indicates that the model correctly predicted 89% of all samples. A high accuracy rate means that the model predicts better overall. A recall of 0.89 indicates that the model correctly identifies 89% of the true comments. An accuracy of 0.91 indicates that the model can identify 91% of all samples predicted to be positive examples as true comments. An F1 value of 0.889 combines the performance of the check accuracy and the check completeness. The closer the F1 value is to 1, it means that the model performs better in combination. An AUC of 0.93 indicates that 93% of the area is under

ISSN 2616-5902 Vol. 6, Issue 4: 276-281, DOI: 10.25236/AJBM.2024.060440

the ROC curve. The higher the AUC, the better the classification ability of the model.

In summary, the values of Accuracy, Recall, Precision, F1, and AUC indicate that the model predicts better, accurately identifies true cases, and has a strong overall classification ability.

3. Conclusions

Construct entropy weight TOPSIS- k -mean clustering-logistic regression model, the results show that after determining the evaluation indexes, after using the entropy weight TOPSIS model evaluation to get the score, the score is lower than 0.607916447 for the machine reviews, that is, false reviews, which contains a small number of real reviews. Using the logistic regression model to test the labeled data, the model accuracy is 0.89, recall is 0.89, precision is 0.91, F1 value is 0.889, AUC is 0.93. It shows that the model has a good prediction effect, can accurately identify the real samples, and has a strong comprehensive classification ability.

References

[1] China Internet Network Information Center. The 52nd statistical report on China's Internet development [EB/OL]. [2023-08-28]. https://www.cnnic.cn/n4/2023/0828/c88-10829.html.

[2] Cao Dongwei, Li Shaomei, Chen Hongchang. A GCN-based method for false comment detection [J]. Computer Engineering and Application, 2022, 58(03):181-186.

[3] Shi Yunmei, Yuan Bo, Zhang Le et al. IMTS: Fusion of image and text semantics for fake comment detection [J]. Data Analysis and Knowledge Discovery, 2022, 6(08):84-96.

[4] Hsing Juanjuan. A false comment recognition method based on Markov logic network [J]. Journal of Chinese Information, 2016, 30(05):94-100.

[5] Ye Zicheng, Wang Banghai. Spectral clustering based false comment group detection [J]. Computer Applications and Software, 2021, 38(08):175-181.

[6] Q. Zhang, S. Ji, Q. Fu et al. Buffer group detection and characterization based on weighted comment graph [J]. Computer Applications, 2019, 39(06):1595-1600.