A review of research on detection of fake commodity reviews

A review of research on detection of fake commodity reviews

Haoxing Tang [1 a], Hui Cao [1*]

[1] Laboratory of China's Ethnic Languages and Information Technology of Ministry of Education, Northwest Minzu University, Lanzhou, Gansu 730000, China

[a]y182730607@stu.xbmu.edu.cn [*]Corresponding author's e-mail: 2232694781@qq.com

Abstract. With the development of the Internet era, more and more consumers are shopping online. Product reviews can influence consumers' consumption decisions. Fake review detection is a valuable research. From the perspective of research data content, it can be divided into two categories: review content based and review metadata; from the perspective of fake review detection method, it can be divided into four categories: rule-based, graph-based model, machine-learning based and deep learning based. Common data sets are also introduced.

1. Introduction

With the development of the Internet, people's shopping consumption is no longer limited by space-time conditions, which brings great convenience to life. However, due to the virtuality of the network, consumers also face the risk of information asymmetry while online consumption reduces costs [1]. Buyers and sellers do not know the true identity of the other party. Consumers obtain commodity information through commodity descriptions provided by merchants and user reviews displayed on e-commerce platforms. Criminals may influence the sales of commodities by means of reviews. These fake reviews will undoubtedly interfere with the purchasing decisions of ordinary consumers [2]. Studies have shown that consumers are very sensitive to positive and negative reviews on commodities. 80% of consumers will change their original consumption decisions because of the large amount of negative reviews on commodities, and 87% will decide whether to buy the commodities based on the positive reviews [3]. Because review information can influence consumers' purchasing decisions, affect the sales of goods, and produce huge economic benefits. To some extent, it encourages fake reviews. Because of the volume of commodity review data, manual screening is expensive. Fake review detection has become a valuable research.

2. Fake review detection

In 2007, Jindal and Liu [4-5] first proposed the detection task of spam reviews, and they summarized spam reviews into three main types:

1. Fake review: Refers to positive or negative reviews made to promote sales or discredit reputation without the reviewer's real use experience.

2. Brand-only reviews: Reviews that are positive or negative about a brand and not specific to a product.

3. Advertisements or irrelevant reviews that do not provide user reviews.

As fake reviews are deliberately forged by imitating real consumer reviews, compared with 2 and 3, fake reviews are more difficult to be recognized by ordinary users and more difficult to detect, so fake review detection has become the mainstream of research. With the development of the Internet industry,

new forms of online fraud, such as online mercenaries and part-time job brushing bills, are constantly emerging. The ways and forms of fake review are becoming more and more diversified. The research on fake review detection is also improving and deepening.

2.1. Fake review detection based on review text content

Revolves around the review text content fake reviews research, fake reviews feature extraction based on text content, using the review text characteristic of fake detection way including the use of n-gram model for text categorization, the use of LIWC do cheat detection [6] psycholinguistics feature extracting, the use of statistical part-of-speech tags (POS tag) genre identification, etc. [7].The linguistic features of critical texts can also be applied to the identification of fake reviews. Most real reviews use nouns, adjectives, prepositions, determiners and parallel conjunctions, which are described from specific and sensory aspects. Some fake reviews use verbs, adverbs and pronouns. By studying the distribution of part of speech in review texts, fake review detection can be conducted [8].

2.2. Introduce review metadata for fake review detection

Fake reviewers are mostly involved in creating fake reviews by accepting the task of brushing. Since they have never used goods, they just try to give positive or negative opinions on things they do not know, which must have certain characteristics. In order to improve the accuracy of fake review detection, we can also consider introducing review metadata and extracting the behavioral characteristics of reviewers from it. Review metadata includes reviewer ID, rating information, review Posting time, review thumb up number, review number, commodity attribute, etc.

Table 1. Review data information.

Data categories	Data content	Data characteristics
Text data	Review text content	text characteristic
Metadata	Reviewer ID Rating information Review Posting time Thumb up number Review number Commodity attribute	Commenter behavior characteristics

A reviewer ID is a unique identifier of an individual reviewer and can be used for reviewer identification. The rating information can directly reflect the critic's emotional attitude towards the product. By making statistics on the time of review publication, we can judge whether the review is generated in random time. The number of reviews thumb up and the number of reviews can reflect how attractive the review is to other users. The commodity attribute can be used to identify whether the review content is related to the item purchased by the reviewer.

After comprehensive analysis of the above metadata information, the behavioral characteristics of reviewers can be extracted, including the reviewer activity window, the maximum number of reviews per day, the total number of reviews, the proportion of positive reviews, and users' abnormal ratings.

3. Fake review detection technology

With the development of computer technology, various technologies have been introduced into the research of fake review detection.

3.1. Rules-based fake review detection

Rule-based fake review detection means to manually make corresponding feature rules and conduct fake review detection based on professional knowledge and experience. Gilbert [7] found similar reviews and repeated users through comprehensive evaluation of content and user characteristics. Lim [9] et al. analyzed and studied the rating behavior patterns of e-commerce platform users, and identified true and

fake reviewers by identifying abnormal rating behavior patterns. Hu Nan[11-12] et al. studied the review data of book sales websites and used the rule that fake reviewers may evaluate a large number of books in a short time to identify fake reviews posted by merchants to influence consumers' purchasing decisions. The efficiency and accuracy of rule-based methods are dependent on the rules made.

3.2. Fake review detection based on graph model

Wang[13] uses review reviewers and stores to construct a graph structure for nodes and uses iterative calculation to identify junk reviews. Li [14] et al. used social graph to study abnormal behaviors of reviewers and conduct fake review detection, with the detection accuracy reaching 87%.

3.3. Fake review detection based on machine learning

Common machine learning methods include Bayesian classification algorithm, logistic regression, K-order nearest neighbor algorithm, support vector machine, etc.

Zhao[15] et al studied the relationship between fake reviews and consumer credibility by changing the ratio of real and fake reviews, and established a learning model of user online reviews based on the framework of Bayesian model and credibility. Feng[16] studied the syntactic styles of fake reviews, and established the grammar style characteristic rules with context-free grammar to conduct the research on fake reviews detection. By designing a supervised linear regression model. Qiu [17] et al. found out the producer of fake reviews, and the model they constructed was mainly the fake reviews of commodities based on user behaviors. Mukherjee[18] et al. used Bayesian algorithm to detect fake reviews by analyzing the difference between malicious reviewators' behavior characteristics and normal reviewators. According to the interaction between reviewers, merchants and reviews, Wang [19] used Logistic regression model to detect fake reviews. Ren [20] et al.proposed a PU(positive and unlabeled) learning framework, and labeled reviews in a data set composed of a few real reviews and a large number of unlabeled reviews were labeled with LDA topic generation model and clustering method. These conclusions were achieved using SVM classifier. Zhao [21] et al. proposed a text-based emotion analysis method to identify fake reviews, and gave different weights to adverbs of degree according to their emotional degree and emotional degree. Then, the reviews were classified according to the emotional polarity before and after the related words, and finally according to the Logistics model.

3.4. Fake review detection based on deep learning

The deep learning method can solve the problems encountered by traditional machine learning methods in feature engineering. The neural network is used for feature extraction, which reduces the time, labor cost and dependence on expert knowledge brought by artificial feature design.

Jing[22] on the basis of previous studies, this paper proposes a fake reviews recognition method based on the deep learning, through the study of the feature extraction and feature selection of data sets, using common neural network, the DBN-DNN network and LBP compared three methods recognition experiment, the last in a mixed data sets using the neural network algorithm for identifying the purpose of fake reviews. Wang[23] proposed a neural network based on attention mechanism and combined linguistic and behavioral characteristics to detect fake reviews. Two neural network models, multi-layer perceptron (MLP) and convolutional neural network (CNN), are used to learn feature vectors, and attention mechanisms are used to give weight to linguistic and behavioral features. Good classification results are obtained and the validity of the model is verified. Zhang [24] used the CNN-LSTM model with the integrated Attention mechanism to conduct research on fake review detection, and verified the effectiveness of his deep learning model in the task of fake review detection. Ren[25] used the neural network model to learn the document-level representation, and integrated the features of the neural network and the discrete features to detect the fake reviews at the document level. In this model, the convolutional neural network CNN is used to learn the sentence representation, then the LSTM neural network is used to learn the document representation, and finally the document representation is directly used as the feature to identify fake reviews. Zhang [26] used three deep learning networks, CNN, DBN and LSTM, respectively, to extract features, and combined with user behavior characteristics, used SVM

classifier to detect fake reviews. Li [27] trained the BiGRU neural network model with the addition of Attention mechanism for fake review identification, and achieved better results than the random forest and LSTM models. The accuracy of fake review identification reached 93.78%.

4. Review data sets

The main sources of review data are all kinds of review websites. The methods to mark fake reviews can be divided into two types: using the filtering algorithm of review sites to mark and manual marking.

Most of the experiments used hotel and restaurant review data from Yelp, hotel review data from TripAdvisor, and Amazon's product review data set.

In 2011, Ott [10] et al. constructed the data set by crowdsourcing to collect fake reviews. Taking the credibility of the reviews as an indicator, the researchers selected 400 real reviews from the reviews crawled by TripAdvisor and collected 400 fake reviews through crowdsourcing platform, forming the gold standard data set of fake reviews. Because the fake reviews collected by crowdsourcing are mainly based on subjective imagination, there is still a gap between them and the fake reviews in real scenes.

Yelp has launched a review filtering algorithm, which separates the spam reviews determined by the algorithm from the normal reviews. The review data of the website can be used to conveniently conduct research related to the detection of fake reviews. Amazon has a wide variety of products and rich review data. Consumers can evaluate the usefulness of the review on the website. By using other consumers to mark the proportion of helpful or unhelpful reviews, fake review labeling can be made.

Table 2. Comparison of common data sets.

Data source	Commodity category	fake reviews marking mode
Amazon	Clothing, books, toys, computers, food	Filtering algorithm of review sites
TripAdvisor	Hotels, restaurants	Manual marking
Yelp	Restaurants, shopping centers, hotels	Filtering algorithm of review sites

5. Conclusion

The detection of fake reviews, especially the detection of Chinese language's fake reviews, still needs further research. With the development of big data, deep learning, natural language processing and other computer technologies, the methods of fake review detection are bound to become more and more abundant.

The research on fake review detection is of great value, which can guarantee the authenticity of reviews, reduce the cost of cleaning up fake reviews on e-commerce platforms, and provide consumers with a better shopping experience. Good evaluation of goods will make consumers more willing to buy goods and protect the legitimate rights and interests of merchants and consumers. Manufacturers can also use fake review detection technology to obtain real review data, analyze users' real feelings, constantly update products, and improve service quality.

References

[1] Chen, L. (2019) On Detection of Fake Online Consumer Review Discourse[D]. Guangdong University of Foreign Studies.
[2] Chen, R.Y., Guo, J.Y. and Deng, X.L. (2014) Detecting Fake Reviews of Hype About Restaurants by Sentiment Analysis[J]. Web-Age Information Management, Springer:22-30.
[3] Cone Research.Game changer: cone survey finds 4-out-of-5 consumers reverse purchase decisions ased on negative online reviews. Available at: http://www.conecomm.com/contentmgr/showdetails.php/id4008.20.
[4] Jindal, N and Liu, B. (2007) Review spam detection[C]. In: Proceedings of the 2007 International conference on the World Wide Web:1089-1090.

[5] Jindal, N and Liu, B. (2008) Opinion spam and analysis[C]. In: Proceedings of the 1st International ACM Conference on Web Search and Data Mining, ACM: 219-230.
[6] DANDIBHOTLA, T.S. (2018) Finding and classifying the deceptive spam reviews using LIWC dictionary variables and decision tree classifier[J]. International Journal of Engineering & Technology, 10(1): 308-315.
[7] Gilbert, E., Karahalios, K. (2010) Understanding deja reviewers[C]. Proceedings of the 2010 ACM Conference on Computer Supported Cooperative work. ACM:225-228.
[8] Mayzlin, D., Dove R.Y., Chevalier J. (2014) Promotional reviews: an empirical investigation of online review manipulation[J]. American Economic Review, 104 (8): 2421-2455.
[9] Lim, E.P., Nguyen, V.A., Jindal, N., Liu, B. and Lauw, H. W. (2010) Detecting product review spammers using rating behaviors. Proceedings of the 19th ACM international conference on Information and knowledge management, ACM:939-984.
[10] Ott, M., Choi, Y., Caridie, C. and Jeffrey, T. H. (2011) Finding Deceptive Opinion Spam by Any Stretch of the Imagination[C].In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies:303-319.
[11] Hu, N.,Bose, I., Gao, Y. and Liu, L. (2011) Manipulation in digital word-of-mouth: A reality check for book reviews[J]. Decision Support Systems,50(3):627-635.
[12] Hu, N., L.Liu and V.Sambamurthy. (2011) Fraud detection in online consumer reviews[J].Decision Support Systems,50(3):614-626.
[13] Wang, G., Xie, S.L., Liu, B.and Yu, P.S. (2011) Identify Online Store Review Spammers via Social Review Graph. ACM Transactions on Intelligent Systems and Technology, 3(4):1-21.
[14] Li, Y.Q.,Fu, H.G. (2014) Fake comments recognition based on social network graph model [J]. Journal of Computer Applications,34(S2):151-153+158.
[15] Zhao, Y., Yang, S., Narayan, V. and Zhao, Y. (2013) Modeling Consumer Learning from Online Product Reviews[J], Marketing science,32(1):153-169.
[16] Feng, S., Banerjee, R., Choi, Y. (2012) Syntactic stylometry for deception detection[C]//Meeting of the Association for Computational Linguistics: Short Paper. Association for Computational Linguistics:171-175.
[17] Qiu, Y.F., Wang, J.K., Shao, L.S., et al. (2012) Research on Product Review Spammer Detection Based on Users' Behavior[J]. Computer Engineering,38(11):1000-3428.
[18] Fei, G., Mukherjee, A., Liu, B., et al. (2013) Explotinng Burstiness in Reviews for Review Spammer Detection[J].ICWSM, 13:175-184.
[19] Wang, Y. (2014) Detecting Fake Review Based on Relationship[D].Nanjing University of Posts and Telecommunications.
[20] Ren, Y.F.,Ji, D.H., Zhang, H.B., et al. (2015) Deceptive Reviews Detection Based on Positive and Unlabeled Learning [J]. Journal of Computer Research and Development,52(3):639-648.
[21] Zhao, J., Wang, H. (2016) Detection of Fake Reviews Based on Emotional Orientation and Logistic Regression [J]. CAAI Transactions on Intelligent Systems,2016,11(3):336-342.
[22] Jing, Y.P. (2014) Research of Deceptive Opinion Spam Recognition Based on Deep Learning[D].East China Normal University.
[23] Wang, X., Liu, K., Zhao, J. (2017) Detecting Deceptive Review Spam via Attention-Based Neural Networks[C]. National CCF Conference on Natural Language Processing & Chinese Computing. Springer, Cham:866-876.
[24] Zhang, H. (2017) Research on Deceptive Opinion Recognition Based on Deep Learning[D].Harbin Institute of Technology.
[25] Ren, Y., Ji, D. (2017) Neural networks for deceptive opinion spam detection: An empirical study [J].Information Science, 385-386: 213-224.
[26] Zhang, S.N. (2018) Research and Design of Fake Reviews Detection Based on Deep Learning[D].Chongqing University.
[27] Li, C.Y. (2020) Research on Review Spam Identification Methods Based on Bi-GRU[D].Shangha i Normal University.