S. Deva, 2021, 9:2
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

Fake Review Classification from Product Reviews vs Deep Learning

Asst. Prof. V. Kavitha, Asst. Prof. J. Charanya, S. Deva, R. Maniraj, G. Surya, R. Yuvaraj
Department of Information Technology,
Nandha College of Technology, Perundurai,
Tamilnadu, India

Abstract- Fake reviews may change their customer mindset. A large number of reviews are affect the product quality and user damage the product usage. They will number of fake reviews are even cause genuine user opinion and user experience crises. Therefore, it's important to detect and solve to filter the fake reviews and remove all fake reviews. However more existing methods are lower accuracy and most of fake reviews are some algorithms are not detected. To solve thisproblem, we propose a new approach and change that existing approaches, Solve the problem and we addmethod to detect fake reviews and based on multiple features and rolling collaborative training. We use First method to multiple features and sentiment features of reviews and behavior features of reviewers. Second features are data set to training in sample set. We designed all algorithms to extract all the features of review. The classifications of reviews are algorithmic to detect that remove manually. The support vector (SVM) classifier has been proven to be by statistical means. It means multiple features and higher accuracy than the baseline models. Meanwhile to resolve the problem of more lacking algorithm and training samples in fake review detection.

Keywords: - Fake review classification, support vector machine (SVM), Amazon.

I. INTRODUCTION

For online shopping, there are inconsistencies between products information and products that deliver thatoffline, whenthecustomer the read the large amount of product review and we choose the target products and to judgment. There are the product reviews are not affect consumers, purchase intensions, it may affect in enterprises. There will be the positive reviews are attract the customers and sale that product in expensive.

While the negative reviews are defame the product and drain the potential customers. Therefore, the product reviews that positive reviews are sale in failure product and then same time the negative reviews are defame that product and to suppress them. The e-commerce platforms the researches show that fake review are not recognized by customer. The online shopping platform the consumer good shopping experience and opinion and then truthful effect review and then methods to find fake reviews are detect.

II. RELATED WORKS

User review are short text as well as its classification problem into binary values. The task will determine whether reviews are around in sample sets train to another new review to find that same set and automatically to note.

Already using machine learning algorithm in which some of algorithm not used in this type huge amount of dataset not used. At the same time there no public datasets. For fake reviews and they decided the review is fake or not to find the judging review is fake in duplicate review.As well as the SVM classification is used to bigram and LIWC the two

International Journal of Science, Engineering and Technology

An Open Access Journal

mixed features to determine fake reviews used the SVM classifier for classification.Neural network based model to learn represent the documents and its calculate the sentence to detect the fake reviews.Proposed a Deceptive Review Identification by Recurrent Convolutional Neural Network model by using word context and deep learning to identify the fake reviews.

III. EXISTING WORK

1. Identify Fake Reviews from Reviewers Behavior Characteristics:

The analyzed the behavior of reviews and used the scoring behavior to detect fake reviews for the first time. Behavior based on the false reviewer's possibility for identifying the fake review the contribution of weight ranking and positive outlier. An effective iterative algorithm for solving these three concepts. Based on the graph have also developed. We proposed iterative framework with three types of nodes. Reviewer and review and product based on the review graph.

IV. INSUFFICIENCY OF EXISTING RESEARCH

In this time several problems in the related research of fake review detection. Supervised learning algorithm based usually fake review detection in classification method under the method used. The large amount of labeled data as a training samples and labeled data are difficult to obtain. Unsupervised learning unlabeled cluster and analysis to classify though the fusing the detection tasks, the accuracy is not high. Semi supervised learning well balances basic features as part of speech or n-gram are used for modeling factors as interaction between features are ignored to reduce the classification.

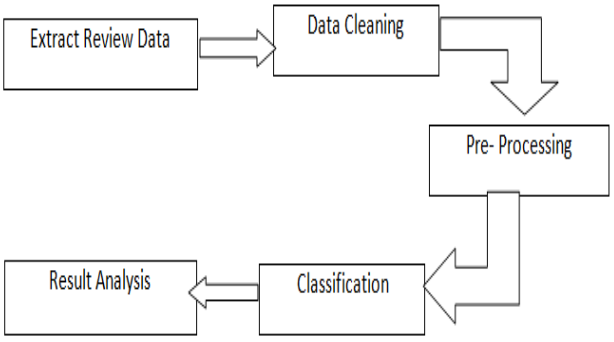

Fig 1. Fake Review Detection Steps.

Existing deep learning models perform well in plain text classification, are not effective in the fake review detection.The main reason is that difficult for fake reviews are features from plain text analyzed multiple angels.User information, business information and other factors. A multiple dimensions method detects fake reviews.

V. PROPOSED SYSTEM

1. The Fake Review Detection Model:

The overall framework of fake review detection in this novels,Build a multi-level- indicator system includes review text and user behavior information, Use crawler technology to collect web reviews and user information. Design a feature extraction algorithm for text and user information, Quantify indicators, delete irrelevant data, and build a classification model for fake review detection based on multi feature fusion and rolling collaborative training.

Multilayer perception (MLP) and convolutional neural network (CNN) are used to learn feature vectors, and language features and behavioral features are weighted through attention mechanisms.

VI. FEATURE EXTRACTION

ASVM based model that classifies the reviews detection model. Used both the review text and the additional features contained in the data set to build a model that predicted with over 75% accuracy. Review text the field of e-commerce are used as the experimental data set Sentiment intensity and the validity of the Doc2 vect text representation network model. The optimal classifier combination is selected through experiments to complete the classification model construction. Test and verify the effectiveness of the multi-feature fusion rolling collaborative train in method proposed in this research, which is more accurate than traditional text classification methods.

This article obtains the original experimental data set amazon-dataset/from the yelp review website. The data includes user ID, the total number of reviews, review content, review level, review time, etc. A total of 5854 records were used as the data set for this experiment, and the fake reviews were marked with the help of the fake reviews filtering system of the yelp review website. The experimental data set is

International Journal of Science, Engineering and Technology

An Open Access Journal

shown in Table 2. The number of training sets and test sets are divided according to the ratio of 8: 2.

The final evaluation index of the experimental results adopts the comprehensive index F1 value and accuracy. Since the reviews are published are unknown the credibility is unknown, the index system to approach to find fake reviews are uncertainly unknown. For example the fake review stars the 4. 1 star reviews and 5 star reviews are like than 3 star reviews. A review is a fake review is improved the star rating and the star rating influence to choose to desire the product in consumers. Reviews the positive emotions will increase preference of the product and negative review will reduce our interest this product. For to change adds the sentiment analysis task. A fake review detection method based multi features and rolling and collaborative training is proposed in the present study.

The method based article lies in two aspects. We proposed method intensity of emotions, to analysis the emotion and to sample set in behavior reviews for judgment fake reviews. Second, its data asset model learning, in which is rolling and then coordinate training data and features extracted by dynamically updated, the detection performance of the classification model.

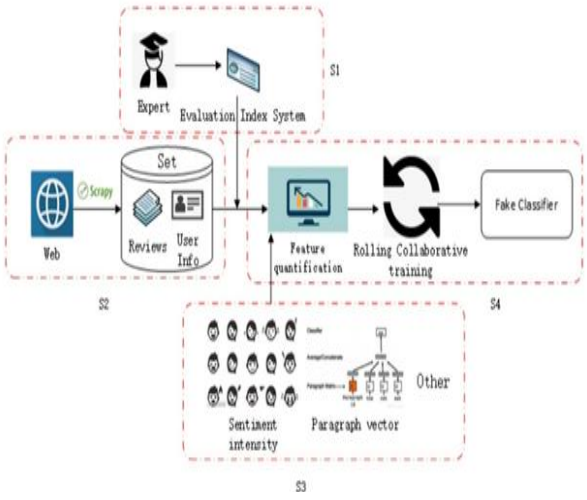

Fig 2. Proposed System Diagram.

The text representation model and sentimental analysis are used to text features, and the users abnormal reviews are the user's behavior. In index system established to users characteristics, and extraction algorithm designed this step. The extracted data is feature are divide into two step according to sets of two attributes, The two feature sets used training and set of the basic classifier in classification model. The integrated learning model in which implements to fake review detection.

Fake review detection model including index selection, feature extraction classifier selection and collaborative training, when there will sample set to execute then rolling and then classifier to automatically to repeat and new set then, fake review detection and article gives possible future research directions.

VII. CONCLUSION

In order to solve the problem the large-scalable data sets are difficult to obtain under the full supervision framework, this study proposed a fake review detection model based on the combination of multi-feature fusion and rolling collaborative training. Experimental results show that this method is more effective than traditional algorithms.

It uses unlabeled data to improve the performance of the classification system, and has better classification accuracy. At the same time, the consistency of sentiment and score is analyzed, and the feature extraction of the review is carried out through the text representation model.

The experimental results in the review of amazon shopping website show that the accuracy of the proposed method for detecting fake reviews is 84.14%, which 3.9% higher than the baseline methods.We also got a novel discovery that the characteristics of reviews will change with time.

The main reason is that the writing methods of professional fake writers will be changed according to the update of the detection mechanism of the e-commerce platform, and we will try our best to bypass. Through the detection mechanism, it is more difficult to be found, and corresponding counter measures can be made based on consumer psychology and times like shopping festivals. We hope that the next researchers can continue their research from the direction of dynamic update detection strategy.

In the future, we will strive to find a more effective and accurate detection method that can detect false information in multiple fields.

International Journal of Science, Engineering and Technology

An Open Access Journal

REFERENCES

[1] G. Lackermair, D. Kailer, and K. Kanmaz, ''Importance of online product reviews from a consumer's perspective,'' Adv. Econ. Bus., vol. 1, no. 1, pp. 1–5, 2013.

[2] D. S. Kostyra, J. Reiner, M. Natter, and D. Klapper, ''Decomposing the effects of online customer reviews on brand, price, and product attributes,'' Int. J. Res. Marketing, vol. 33, no. 1, pp. 11–26, Mar. 2016.

[3] S. Ullrich and C. B. Brunner, ''Negative online consumer reviews: Effects of different responses,'' J. Product Brand Manage., vol. 24, no. 1, pp. 66–77, Mar. 2015.

[4] S. Deng, C.-X. Wan, A.-H. Guan, and H. Chen, ''Deceptive reviews detection of technology products based on behavior and content, ''J.Chin. Comput. Syst., vol. 36, no. 11, p. 2498, 2015.

[5] D. Radovanovic and B. Krstajic, ''Review spam detection using machine learning,'' in Proc. 23rd Int. Sci.-Prof. Conf. Inf. Technol., Feb. 2018, pp. 1-4.

[6] H. Deng, L. Zhao, N. Luo, Y. Liu, G. Guo, X. Wang, Z. Tan, S. Wang, and F. Zhou, ''Semi-supervised learning based fake review detection,'' in Proc. IEEE Int. Symp. Parallel Distrib. Process. Appl., Dec. 2017, pp. 1278–1280.

[7] H. Ahmed, I. Traore, and S. Saad, ''Detecting opinion spams and fake news using text classification,'' Secur. Privacy, vol. 1, no. 1, p. e9, 2018.

[8] M. Crawford, T. M. Khoshgoftaar, J. D. Prusa, A. N. Richter, and H. Al Najada, ''Survey of review spam detection using machine learning techniques,'' J. Big Data, vol. 2, no. 1, p. 23, Dec. 2015.

[9] N. Jindaland B. Liu, ''Opinion spam and analysis, 'in Proc Int .Conf. Web Search Data Mining, 2008, pp. 219–230.

[10] K. H. Yoo and U. Gretzel, ''Comparison of deceptive and truthful travel reviews,'' in Proc. ENTER, 2009, pp. 37–47.

[11] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, ''Finding deceptive opinion spam by any stretch of the imagination,'' 2011, arXiv: 1107.4557. Online. Available: http://arxiv.org/abs/1107.4557

[12] C. Tang and L. Guo, ''Digging for gold with a simple tool: Validating text mining in studying electronic word-of-mouth (eWOM) communication,'' Marketing Lett., vol. 26, no. 1, pp. 67–80, Mar. 2015.

[13] S. Feng, R. Banerjee, and Y. Choi, ''Syntactic stylometry for deception detection,'' in Proc. 50th Annu. Meeting Assoc. Comput. Linguistics, vol.2, 2012, pp. 171–175.

[14] C. Yanfang and L. Zhiyu, ''Research on product review attribute-based of emotion evaluate review spam detection,'' Data Anal. Knowl. Discovery, vol. 30, no. 9, pp. 81–90, 2014.

[15] J. Li,. C. Cardie, and S. Li, ''Topic spam: Atopic-model based approach for spam detection,'' in Proc. 51st Annu. Meeting Assoc. Comput. Linguistics, vol. 2, 2013, pp. 217–221.

[16] W. Sun, Y. Guo, Z. Fan, and X. Xu, ''False comment recognition based on the combination of content features and user behavior characteristics,'' in Proc. 3rd IEEE Int. Conf. Comput. Commun. (ICCC), Dec. 2017, pp. 1510–1514.

[17] Y. Ren and D. Ji, ''Neural networks for deceptive opinion spam detection: An empirical study,'' Inf. Sci., vols. 385–386, pp. 213–224, Apr. 2017.

[18] L. Li, B. Qin, W. Ren, and T. Liu, ''Document representation and feature combination for deceptive spam review detection,'' Neuro computing, vol. 254, pp. 33–41, Sep. 2017.

[19] W. Zhang, Y. Du, T. Yoshida, and Q. Wang, ''DRI-RCNN: An approach to deceptive review identification using recurrent convolutional neural network,'' Inf. Process. Manage. vol. 54, no. 4, pp. 576–592, Jul. 2018.

[20] E.-P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw, ''Detecting product review spammers using rating behaviors,'' in Proc. 19th ACM Int. Conf. Inf. Knowl. Manage, 2010, pp. 939–948.

[21] F. Wu, J. Shu, Y. Huang, and Z. Yuan, ''Co-detecting social spammers and spam messages in micro blogging via exploiting social contexts,'' Neuro computing, vol. 201, pp. 51–65, Aug. 2016.

[22] G. Wang, S. Xie, B. Liu, and P. S. Yu, ''Review graph based online store review spammer detection,'' in Proc. IEEE 11th Int. Conf. Data Mining, Dec. 2011, pp. 1242–1247.

[23] Z. Wang, S. Gu, X. Zhao, and X. Xu, ''Graph-based review spammer group detection,'' Knowl. Inf. Syst., vol. 55, no. 3, pp. 571–597, Jun. 2018.