
Detection of Fake Reviews Using Machine Learning Techniques

Wesam Hameed Asaad

Computer Sciences Department, University of Technology, Iraq
cs.20.48@grad.uotechnology.edu.iq

Yossra Hussain Ali

Assistant Professor, Computer Sciences Department, Computer Sciences
Department, University of Technology, Iraq
yossra.a.ali@uotechnology.edu.iq

Ragheed Allami

Lecturer, Computer Sciences Department, Computer Sciences Department,
University of Technology, Iraq
110046@uotechnology.edu.iq

Abstract:

Online marketing generates vast amounts of digital information, which is then used to advertise millions of goods and services. Finding the best services or goods that meet the need is, therefore, difficult. Customers make decisions based on evaluations or opinions expressed by third parties. In this cutthroat environment. For the past few years, some businesses have employed writers to fabricate favorable reviews of their services or goods or unfairly critical reviews of those of their rivals in order to increase the number of false reviews. Review sites are increasingly having to deal with the distribution of false information that either benefits or hurts particular firms. Different types of opinion spam are used to deceive both human readers and automated sentiment analysis and opinion mining systems. Different strategies have been put forth as a result, allowing for the evaluation of the user-generated content's credibility. The current paper provides a quick introduction to the examination of

opinion mining and the review-centric attributes that have been suggested for use in machine learning techniques that use graphical methods to identify bogus reviews. Along with a review of the related literature and a discussion of it.

Keywords: Fake Review, Opinion Mining, Sentiment, Machine Learning.

1. Introduction

People can, without limits, communicate their opinions and feelings on an internet forum, such as social media, about other people or products, even if occasionally their comments are hurtful or ungenune. Some people dare to express such fake information without hesitation or fear because no standards can restrict them from doing so [1]. Social media and internet publishing have made it simpler to voice comments more confidently. These comments and reviews offer advantages and cons, as they give a real-life experience with that specific product or describe its usability. However, some people might be hired to supply fake information based on certain propaganda and promote a product by allowing such people to submit reviews promoting things that are not worth such a positive rating. Alternatively, diverge from competent services and products without disclosing who they are or their business [2]. E-commerce sectors and internet reviews are ceaselessly developing. Reliance on these internet reviews is also increasing. Due to the exponential expansion in these sectors. There are approaches that have been presented to analyze the information provided by customers on commercial websites in order to verify original evaluations using categorization methods.

The most explored tasks have been the identification of [2]:

1-Review the site investigation.

2- Fake reviews on microblogging sites

Machine learning approaches are making a great contribution to detecting fake evaluations of commercial websites. In general, such procedures and techniques [3] can uncover and extract useful knowledge by utilizing machine data mining concepts. Usually, fake review identification depends not only on the category of reviews but also on some features that are not directly connected to the content. Building features for reviews generally requires text and natural language processing (NLP). However, fake reviews might need to incorporate other qualities linked to the reviewer himself, for example. The word "long" in terms of the longevity of a portable computer battery is a good idea, whereas the same phrase regarding the first time is a negative idea. This shows that the concept-based mining system cannot recognize this type of phrase, which has a distinct meaning in different situations [4]. For example, the word "long" in terms of the longevity of a portable computer battery is a good idea, whereas the same phrase regarding the first time is a negative idea. This illustrates that the concept-based mining method cannot recognize this type of phrase, which has diverse meanings in different contexts. Finally, in some circumstances, people make inconsistent remarks that are impossible to anticipate. To alleviate the significant difficulty that online websites are having owing to the spam of ideas, this project offers to identify any such bogus updates that have been spammed by classifying them as fake and real. This is a supervised learning strategy that uses different machine learning algorithms to detect false or real updates. In this work, we are analyzing the review sites investigation and analysis to discover fraudulent reviews. Both supervised and unsupervised strategies will be evaluated and addressed in this study. However, due to the number of scholarly studies performed in this sector utilizing supervised techniques, we studied extensively the classification methods used for false review detection. On the other hand, the use of supervised procedures has also been witnessed in the field.

1.1 Problem Statement

Most of the previous studies dealt with the issue of fake reviews and had some flaws. The aim of this study is to compare it with theirs to show the difference between them such as Low accuracy, the researcher does not use a complete data set, and high execution time implementation.

1.2 Research Hypothesis

Previous studies dealt mainly with the issue of mock reviews, especially in terms of accuracy, but there are deficiencies in the processes to address this problem.

1.3 Aim of the Study

The paper proposal is to improve the classification of consumers' opinions using ML algorithms and techniques to classify comments and reviews as truthful reviews or fake reviews using sentiment analysis (SA), review the previous study, and discuss them.

1.4 The importance of research

Globally, e-commerce is expanding at an unprecedented rate. The significance of internet reviews is growing every day due to their expansion. Reviews have the power to sway consumers' purchase choices. Nowadays, especially among prospective clients, reading product reviews before purchasing an item has become a habit. Customers publish feedback about the goods they buy, which might be favorable or unfavorable. Such reviews offer insightful input on the products, and prospective buyers may utilize them to learn what other users have to say about a product before making a purchasing decision. Customers frequently examine reviews of the current product before making a purchase decision. If most of the

reviews are favorable, there is a good possibility that you will purchase the item. If the reviews are overwhelmingly bad, shoppers will instead choose different products. While relying only on online reviews can be beneficial, doing so puts both buyers and sellers at risk.

1.5 The limits of the study

Studies related to the research topic were covered from 2018 to 2022.

1.6 Terminology of study

We review some of the terms used in this paper.

Term	Defining
FE	Feature Extraction
Light GBM	Light Gradient Boosting Machine
ML	Machine learning
NLTK	Natural Language Toolkit
SA	Sentiment analysis
TFIDF	Term frequency-inverse document frequency
XGboost	Extreme Gradient Boosting
DM	Data mining
SVM	Sport vector machine
Dt	Decision tree
SVC	Support Vector Classification.
KNN	K-nearest neighbors
LR	Logistic regression
NB	Naive Bayes

2. Related work

Supervised learning is a method that involves training the system into pre-determined knowledge. the supervised classification system. So, this research paper collected many studies on this subject.

1- Elmurngi, A.Gherbi, et al. (2018). proposed an open software tool named "Weka Tool" to construct machine learning algorithms utilizing sentiment analysis. The categorization is undertaken to differentiate fair and unfair reviews from Amazon reviews based on three categories: positive, negative, and negative neutral terms. In this research work, the spam reviews are discovered by just including the helpful votes chosen by the customers and the rating deviation evaluated, which limits the system's overall performance. Also, as per the researcher's observations and testing results, the existing system uses a Naive Bayes classifier for spam and non-spam classification. The precision is relatively low, which may not deliver correct results for the user [1].

2- Aishwarya Pendyala (2019). The authors attempt to classify the reviews obtained from freely available datasets from various categories, including service-based, product-based, customer feedback, experience-based, and the crawled Amazon dataset, with greater accuracy using Naïve Bayes, linear SVC, SVM, random forest, and decision tree algorithms. In order to increase accuracy, other features, including a comparison of the sentiment of the review, verified purchases, ratings, emoji count, and product category with the overall score, are employed in addition to the review specifics. A classifier is developed based on the discovered properties. In addition, such traits are assigned a likelihood factor or a weight depending on the classified training sets. This is a supervised learning strategy that applies different machine learning algorithms to detect fraudulent or real reviews. [6].

3 - M.Bansode. et al (2021). Processed the hotel ratings with the use of DM methods and reviews acquired from negative or positive users which could be utilized for customer sentiment for neutralizing a product and algorithms were used [7]

4- Elmogy, Ahmed M. et al. (2021). employed engineering features in ML to detect fake reviews, extract features from reviews, and operate on a real data set for restaurants without features derived from user behaviors, where the methodologies were LR, NB, KNN, and SVM [3].

3. Implementation Study

1) Data collection: We will upload the Yelp dataset to the program using this module.

2) Data Pre-processing: In this module, we will read every review, normalize stemming, remove stop words, and then pre-process every review to extract features.

3) Features Extraction: We will apply the TF-IDF (term frequency-inverse document frequency) technique to convert text reviews into numeric vectors. The words will be replaced by vectors that include each word count.

4) Run the XGBoost algorithms: We'll train the XGBoost method using the TF-IDF vector, and then test data will be used to the trained XGboost model to see how well it predicts the future.

5) Run the LightGBM algorithm: We'll train the algorithm using the TF-IDF vector, and then test data will be applied to the trained model to see how well it predicted the future.

Table 1. Fake reviews detection using the supervised methods

Author name	Year publish	Dataset used	Method used	Performance
Elshrif Ibrahim Elmurungi and Abdelouahed Gherbi	2018	Amazon.com	LR	70% of the LR model, which is the base performance model, is accurate.
Pendyala Aishwarya	2019	Amazon.com	Naïve Bayes	Accuracy-80.542 F1 score-77.888 Precision Score-80.612 Recall-79.001
M.Bansode. et al.	2021	'Hotel Reviews' dataset 'Deceptive opinion spam' dataset 'Yelp Labelled Review Dataset with Sentiments and Features	(KNN), LR, Multinomial Naïve Bayes,(SVM),Stochastic Gradient Descent (SGD),	Accuracy 60.54% 77.58% 75.38%76.25% 77.81%
Elmogy, Ahmed M.et al	2021	Yelp	SVMNBKNN Logistic Regression and Random forest	accuracy was achieved using the SVM classifier, which scored 86.90%.

4. Fake review detection

Manually reviewing reviews is a primary means of detecting bogus reviews. This method is built on the notion that humans are capable of detecting hazards. when other humans engage in deception—i.e., understanding the deception's "psychology of lies" [9]. The Benefits of Careful Planning The advantage of reading false reviews is that it allows you to build heuristic guidelines.

5. Datasets

A dataset is a group of data that could be saved in any type of file format, including CSV, text, and others. This data might be kept in tabular form, with each row representing a record for a specific dataset column and each column designating a variable corresponding to that dataset. The dataset is used to give the model particular training data. The Kaggel [7] dataset in Excel file format is used, as shown in Table (2).

Table (2): The Yalp.com dataset.

Variable	Description
User_id	Index review
Product_id	The Product ID is unique within a data set
Rating	(5-star) overall satisfaction rating
Date	Date of review
Review	fakeopinion
Fake (1) and real(0) opinions	Opinion Spam Corpus Positive opinions and negative
Features	Features of fake review

6. Summary and research gap

As fake reviews are such a frequent and damaging problem, helping customers and organizations discern authentic evaluations from fraudulent reviews is a vital and demanding endeavor [10]. A bogus review Manual technique, supervised machine learning, and algorithms can all be used to identify anything. techniques [10]. Some approaches in the literature focus simply on features retrieved from the review text. Linguistic features span from quantifying the frequency of words or n-grams to datasets that may not be correctly conceived, contain mislabeled examples, or are not made publicly available. A crucial insight from past studies is that automatic fake review identification has been only partially successful. While one research project cannot cover all gaps.

7. Discussion

People write unworthy, favorable reviews about things to promote them. On some occasions, malicious, unfavorable reviews of other (competitive) products are given in order to destroy their reputation. Some of them consist of non-reviews (e.g., advertising and promotions) that contain no review of the product. The first problem here is that a term might be positive in one context while being bad in another [10]. For example, the word "long" in terms of a laptop's battery life being lengthy is a favorable comment, whereas the same phrase concerning the start time being long is a negative one. This illustrates that the text mining algorithm trained on words from reviews cannot understand the essence of the word, giving it a varied meaning in different scenarios. Another problem is that people don't always express their reviews in the same way. Most of the standard text processing approaches presume that a slight variation in the text doesn't greatly impact the meaning. However, phony review detection, e.g., the service was excellent" and the service wasn't great," does make a tremendous impact. Finally, in some circumstances, people make

contradictory comments, which makes it impossible to foresee the nature of the comment. There could be a hidden positive in a poor review. Moreover, occasionally there is both a positive and negative impression regarding the product [11]. Thus, finding false reviews in a large dataset is tough enough to become an important research subject. Business organizations, professionals, and academics are striving to create the best approach for fake review analysis. A single algorithm cannot tackle all the difficulties and challenges presented by today's age with improvements in technology, yet a few are quite efficient in their analysis. More future work and expertise are needed to further boost the performance of the opinion review analysis and produce one that is consistently efficient across all types of data [12].

8. Conclusions

Because of the fast improvements in internet commerce, the volume of reviews on products and purchases rose. These vast quantities of information are produced on the Internet. There isn't any analysis of the quantity of consumer reviews. Any form of comment or review may be posted that conclusively leads to phony reviews, and some organizations are recruiting people to publish reviews. Some of the false reviews have been purposely manufactured to seem authentic; the power to identify fake online evaluations is vital. In this research, we have studied numerous fake review detection strategies in terms of supervised methodology. We have also examined several ways to identify bogus reviews. We have also covered the major issues of bogus review detection.

In this research, we discussed light GBM, XGbooster, Naïve Bayes, linear SVC, SVM, LR, NB, KNN, random forest, and decision trees. All these methods were proven to be effective when the supplied data were labeled, i.e., they could do the classification. The distinctions between them are relative. Many of the examined publications reported sentiment classification algorithms and carried out experiments

using three separate datasets of Amazon reviews with stop words eliminated. Yalp dataset the experiments investigated the accuracy, precision, and recall of categorization. The authors were effective in detecting unfair negative reviews, unfair neutral reviews, and unfair positive reviews by utilizing the detection techniques of this system. The results of the current study compared to previous studies are more accurate than previous studies.

References

- 1 Elmurngi, E. I., & Gherbi, A. (2018). Unfair reviews detection on amazon reviews using sentiment analysis with supervised learning techniques. *J. Comput. Sci.*, 14(5), 714-726.
- 2 Patel, Nidhi A., and Rakesh Patel. "A survey on fake review detection using machine learning techniques." 2018 4th International Conference on Computing Communication and Automation (ICCCA). IEEE, 2018.
- 3 A. M. Elmogy, U. Tariq, M. Ammar, and A. Ibrahim, "Fake reviews detection using supervised machine learning," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 1, 2021.
- 4 Shinde, M., Shingate, R., Badgular, T., Zingade, G., & Kadam–Rawte, M. N. FAKE REVIEW DETECTION SYSTEM.
- 5 R. Dewang, A.Singh, "Identify of Fake review using new set of lexical and syntactic feature", *Proceedings of the Sixth International Conference on Computer and Communication Technology 2015*, ACM, pp. 115-119, 2015
- 6 Aishwarya Pendyala, FAKE CONSUMER REVIEW DETECTION, 2019
- 7 Bansode, M. and Birajdar, A. (2021). Fake Review Prediction and Review Analysis. *International Journal of Innovative Technology and Exploring Engineering*, 10(7), 143–151
- 8 Salminen, J., Kandpal, C., Kamel, A. M., Jung, S. G., & Jansen, B. J. (2022). Creating and detecting fake reviews of online products. *Journal of Retailing and Consumer Services*, 64, 102771.

- 9 Li, H., Fei, G., Wang, S., Liu, B., Shao, W., Mukherjee, A., & Shao, J. (2017, April). Bimodal distribution and co-bursting in review spam detection. In Proceedings of the 26th international conference on world wide web (pp. 1063-1072).
- 10 Choo, E., Yu, T., & Chi, M. (2015). Detecting opinion spammer groups through community discovery and sentiment analysis. In Data and Applications Security and Privacy XXIX: 29th Annual IFIP WG 11.3 Working Conference, DBSec 2015, Fairfax, VA, USA, July 13-15, 2015, Proceedings 29 (pp. 170-187). Springer International Publishing.
- 11 Wang, Z., Hou, T., Song, D., Li, Z., & Kong, T. (2016). Detecting review spammer groups via bipartite graph projection. The Computer Journal, 59(6), 861-874.