IJNRD.ORG ISSN : 2456-4184

INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG

An International Open Access, Peer-reviewed, Refereed Journal

International Research Journal IJNRD Research Through Innovation

The Fake Review Effect On E-commerce :SKL – Based Fake Review Detection

[1st] Dr. S. Selvakani, [2nd] K. Vasumathi, [3rd] J V Hemakumar,

[1st] Assistant Professor and Head, [2nd] Assistant Professor, [3rd] PG Scholar,
PG Department of Computer Science, Government Arts and Science College,
Arakkonam, Tamilnadu, India – 631051,

Abstract: The significance of reviews on any e-commerce website cannot be overstated, as they often serve as the basis for a potential buyer's purchasing decision. Reviews enable buyers to assess the authenticity and quality of a product based on the feedback provided by other customers. However, some sellers take advantage of reviews by publishing reviews that either promote or discredit a product. These reviews, which do not represent the genuine opinion of an individual, are referred to as fake reviews. The existence of such fake reviews can lead to inaccurate judgments by buyers, which can also damage the platform's credibility. As a result, detecting fake reviews is critical. This paper proposes a method for identifying fake reviews on the platform using a logistic regression model that considers review-specific characteristics and achieves an overall accuracy of 82%. Additionally, our research demonstrates the importance of the "verified purchase" feature in identifying fake reviews.

Keywords **- Fake reviews, k-nearest neighbor (KNN), Machine learning, Natural language processing, Sentiment analysis, Support vector machine (SVM).**

I. INTRODUCTION

In the digital age, customers have the ability to post reviews and opinions about products and services on various websites. These reviews are useful for organizations and potential consumers who want to gather information before making a decision. The number of customer reviews has increased substantially in recent years and they have a significant impact on the purchasing decisions of potential buyers. Consequently, customer reviews offer a valuable service for individuals.

One of the most critical challenges today is the prevalence of fraudulent customer reviews online, which can negatively affect a brand or organization. This type of fraud, known as "Opinion Spamming," has become increasingly sophisticated and organized, leading to substantial profits for those who engage in it. Detecting opinion spam is more challenging than other types of spam, as understanding the context is crucial to identifying deceptive reviews. As a result, researchers are delving into sentiment analysis, opinion mining, and review analysis to address this issue, given that these reviews can have a significant impact on a business's success or failure. It is therefore crucial to determine the authenticity of customer reviews.

As the social web becomes increasingly popular, users continue to spread various types of content that lack a trustworthy external source, making it challenging to authenticate the content being posted. In the business sector, this phenomenon affects individual consumers and undermines their confidence in online shopping. Identifying indicators of fraudulent reviews based on the fraudster's behavior is crucial. Some scholars have used data mining and natural language processing (NLP) techniques, as well as data cleansing and database query processing, to handle raw data. However, these techniques have not effectively solved the spam reviews problem, especially with the increasing number of reviews posted daily. Information cleaning and repair can lead to a flood in high business activity costs.

The novelty of this research lies in several aspects. First, feature selection is based on a multi-level feature extraction system that goes beyond standard NLP techniques to extract various aspects of the reviewer's behavior and reviews. Behavioral features such as review time, writing style, use of punctuation, verb/noun count in a review, and relationship words were extracted to contribute to the overall classification accuracy and authenticity. Secondly, the study aimed to obtain the best fit training and testing dataset samples to obtain the best classifier results.

II. REVIEW OF LITERATURE

D. H. Fusilier, M. M.-Y. Gomez, P. Ross, and R. G. Cabrera [1]

It is highlighted that the internet is not only a vast repository of digital information but also serves as the largest platform for communication. This feature has encouraged businesses of all sizes and types, including hotels, restaurants, television networks, and film makers, to use the web as a crucial marketing platform by creating websites and discussion forums for their products and services. With the proliferation of review sites and blogs, consumers increasingly rely on online reviews to make informed purchase decisions. A recent survey revealed that positive online reviews have strengthened the buying decisions of 87% of consumers, while 80% have changed their minds about purchases based on negative information found online.

D. H. Fusilier, M. M.-Y. Gomez, P. Rosso, and R. G. Cabrera [1]

Supervised text classification techniques have recently been used to successfully detect deceptive opinion spam. The same approach was applied to classify negative opinions and the main conclusion from these works was that standard text categorization techniques using unigrams and bigrams word features are effective at detecting deception in text, with significantly better results than those from human judges.

PU-learning, a semi-supervised technique for building a binary classifier based on positive and unlabeled examples only, was used in the experiments. Two sets of examples were available for training. Naïve Byes (NB) classifier was used for all the experiments, employing the implementation by Wake, and considering all words occurring more than once in the training set as features.

H. A. Najada and X. Zhu [2]

Discuss in their paper the increasing dependence on the internet for daily life activities, including shopping. While purchasing items or services previously used is easy, it can be challenging when buying new products. Customers often seek detailed information about the products, particularly from past customer experiences. Some advantages highlighted in the paper include.

The availability of e-commerce websites enables people to purchase various products online. Sentiment analysis and opinion mining are essential in extracting subjective information from written or spoken words, as they are crucial for many relevant applications. Their method involves adding a randomly sampled set from the minority class and replicating examples to increase the number of minority examples, ultimately achieving a balanced degree. To validate their method for detecting review spasm, they used the WEKA data mining tool. Sampling was used to create training datasets with different levels of data imbalance in order to evaluate the performance of their algorithm.

A. Mukherjee, B. Liu, and N. Glance [3]

Emphasize that in today's world, people tend to read reviews of products before making a purchase. Positive reviews often lead to increased sales and popularity, while negative reviews can steer customers away. Unfortunately, this creates a strong incentive for individuals to engage in opinion spamming, where fake reviews are written to deliberately mislead readers. This is done in order to write numerous reviews about multiple products and earn more money. However, such collective behavior can reveal groups of spammers. The focus of this paper is on identifying such groups. Content similarity among their reviews. With the 8 group behavioral features, we can separate spam and non-spam groups with a large margin, • Group tight coupling measures how closely a member works with the other members of the group. If a member, m, posts almost at the same date as other group members, then m is considered to be tightly coupled with the group. There are very few groups that obtain significant feature values before the jump abscissa, as indicated by the results. From our analysis, we found that there are very few spam groups with GSUP < 0.25, and for GCS, we found that approximately 90% of non-spam groups in our database have minimal along with the labeled data from Sec. The classic approach to detecting spammer groups is to use a supervised classification, regression, or learning to rank algorithm to classify or rank candidate groups.

W. Etaiwi and G. Naymat [4]

Explain that business owners may encourage reviewers to write positive reviews of their products or services or negative reviews of their competitors' products or services. Such reviews are considered spam and can have a significant impact on the online marketplace. However, detecting spam reviews is a challenging task since there are no clear indications in the review text that the review is genuine or fake. Additionally, spam reviews often appear to be normal. Therefore, it is impractical and time-consuming to manually distinguish between genuine and fake reviews. While some websites allow users to label reviews as helpful or not, this approach still requires human judgment. The authors proposed categorizing spam reviews into three main types. It is mentioned that business owners may influence reviewers to write positive reviews about their own products or services, or negative reviews about their competitors' products or services. To develop the proposed model for truthful positive reviews of hotels found on Trip Advisor, several features were used, including the frequency distribution of part-of-speech (POS) tags, Linguistic Inquiry and Word Count (LIWC) features, and n-gram feature sets. The model was trained using NB and SVM algorithms. This section briefly discusses the pre-processing methods, dataset used for evaluation, feature selection approach, classification algorithm, and evaluation measures employed in this work. Feature selection involves extracting features from data to use them as classification criteria.

A.Mukherjee, V. Venkataraman, B. Liu, and N. Glance [5]

Discuss the use of online reviews in making purchase and business decisions. They note that positive reviews can lead to significant financial gains and increased popularity for businesses and individuals. Unfortunately, this also creates an incentive for individuals to write fake reviews that sound genuine and are not easily detected by readers. If a real customer is incentivized to write a fake review, they may be more careful in their writing to avoid detection. Similarly, if a paid reviewer is writing a fake review, they may not know the business well but may have experience in writing fake reviews. However, if a Turker writes a review anonymously, they are less likely to know the business well and do not need to write as carefully to avoid detection since they are being paid for their review as part of research. A new set of behavioral features combined with n-gram features significantly improves accuracy of fake review detection. Surprisingly, behavioral features alone outperform n-grams in detection. The paper reports comprehensive classification experiments using both Yelp and AMT datasets, highlighting a significant difference in accuracy between them. All experiments are conducted using 5-fold Cross Validation (CV) and preparing training and test data accordingly, as using highly imbalanced data may lead to poor models. The words used in AMT fake reviews differ significantly from those in genuine reviews, indicating poor fake reviews by Turkers who have little incentive to write them. The study includes crawling the profiles of all reviewers in the hotel and restaurant domains to obtain and analyze behavioral features about reviewers. The effectiveness of these features is also analyzed on a per-reviewer basis.

N. J. Conroy, V. L. Rubin, and Y. Chen [6]

State that in their paper, the focus is on news verification, which involves using technology to detect intentionally deceptive news content online. This is an important issue within certain areas of library and information science (LIS). The task of detecting fake news involves predicting the likelihood of a news article, whether it is a news report, editorial, expose, or any other format, being intentionally deceptive. The goal of these tools is to perform filtering tasks that were previously done by journalists and traditional news publishers. With the rise of user-generated content and Computer Mediated Communication (CMC) technologies such as blogs, the nature of online news publication has changed. This makes it difficult to perform traditional fact-checking and vetting to prevent deception, given the flood of content from various sources, formats, and genres. Most deceivers use language strategically to avoid detection, but certain aspects such as pronoun usage can "leak" their deception. However, relying solely on isolated n-grams divorced from context information is a limitation of this approach, and ambiguity in word sense remains unresolved. Deep Syntax: To improve deception prediction, deeper language structures (syntax) have been analyzed through Probability Context Free Grammars (PCFG). NETWORK APPROACHES: Innovative approaches using network properties and behavior can complement content-based approaches that rely on deceptive language and leakage cues. Linked data: Utilizing knowledge networks may provide a significant advancement towards scalable computational fact-checking methods.

P. Kaghazgaran, J. Caver lee, and M. Alfifi [7]

Emphasize the importance of user reviews in our decision-making process for various products and services. Online review aggregators such as Amazon, Netflix, and Yelp heavily influence our choices, but unfortunately, these reviews can be easily manipulated, which poses a threat to the credibility of the platforms and the products/services they offer. To address this issue, previous studies have attempted to identify and uncover this manipulation through the use of machine learning or graph-based algorithms. One approach to identifying fake reviews is manual labeling, where judges evaluate individual reviews to determine if they are genuine or fake. These judges can be either researchers or a team of labelers from a review site. The study also investigates the changes in these reviews over time, which may aid in identifying reviewers who are trying to remain hidden. The research focuses specifically on tasks related to Amazon posted on the crowd-sourcing site Rapid Workers, but there are many other sites and targets. The researchers also analyzed the content of the reviews and determined the proportion that contained first-person pronouns.

III. PROPOSED METHODOLOGY

In our system, we have implemented the Naive Bays algorithm for prediction. Naive Bays classifies past data to generate prediction output. Our goal is to create an efficient and robust model for detecting fake reviews in e-commerce product reviews using the Naive Byes approach, which models the probability of classifying reviews as fake or genuine. We have also developed a consensus strategy for feature extraction and text preprocessing. The advantage of the Naive Byes classifier is its simplicity and quick convergence. Our system is easy to implement and provides accurate predictions.

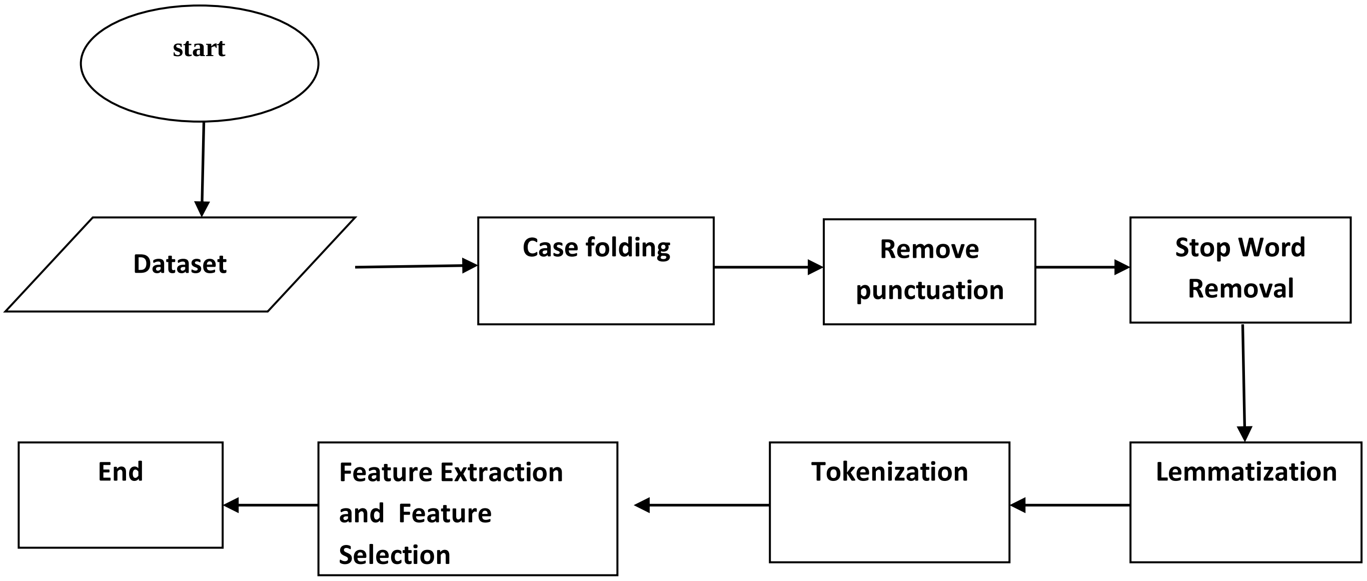

Figure : 1 System architecture

3.1 MODULES

- Data collection
- Data pre-processing
- Feature Extraction
- Evaluation model

IV. EXPERIMENTAL RESULTS

We presented an algorithm for detecting fake reviews in the e-commerce sector based on Support Vector Machine, K-Nearest Neighbor, and Linear Regression (SKL). Our aim was achieved by examining a dataset of hotel reviews and employing machine learning techniques and text classification methods to identify reviews that were not genuine.

4.1 PREPROCESSING

The pre-processing phase involves filtering data, which involves removing less valuable parts of the text, such as punctuation symbols. Eliminating punctuation marks, such as etc., is necessary as they can lower the accuracy of the classification process, and their removal can lead to better output by the algorithm used.

Table 1 : Datasets Details

Dataset	Yelp	Trip Advisor
DATA CONSTRUCTION METHOD	FILTERING ALGORITHM	AMAZON MECHANICAL TURK
TOTAL	1900	800
DOMANID	RESTAURANTS AAND HOTEIS	HOTELS

4.2 CLASSIFICATION ALGORITHMS

A survey report was conducted, and it was concluded that machine learning algorithms show better performance on medium-sized datasets compared to Artificial Neural Network (ANN) and deep learning models, which are more suitable for large datasets. Another reason why ANN and deep lea/rning models do not provide better results for fake review detection is feature engineering.

4.3 DETECTION PROCESS

Upon completion of the training phase, the model is tested using the dataset to predict the output. The SVM, KNN, and Logistic Regression algorithms are used to train the model. A comparison table is generated to determine which algorithm performs the best for the given process. The Friedman test chart is based on statistical analysis. The horizontal axis represents the average order value, while each algorithm is displayed on the vertical axis. A dot is used to represent the average order value of each algorithm.

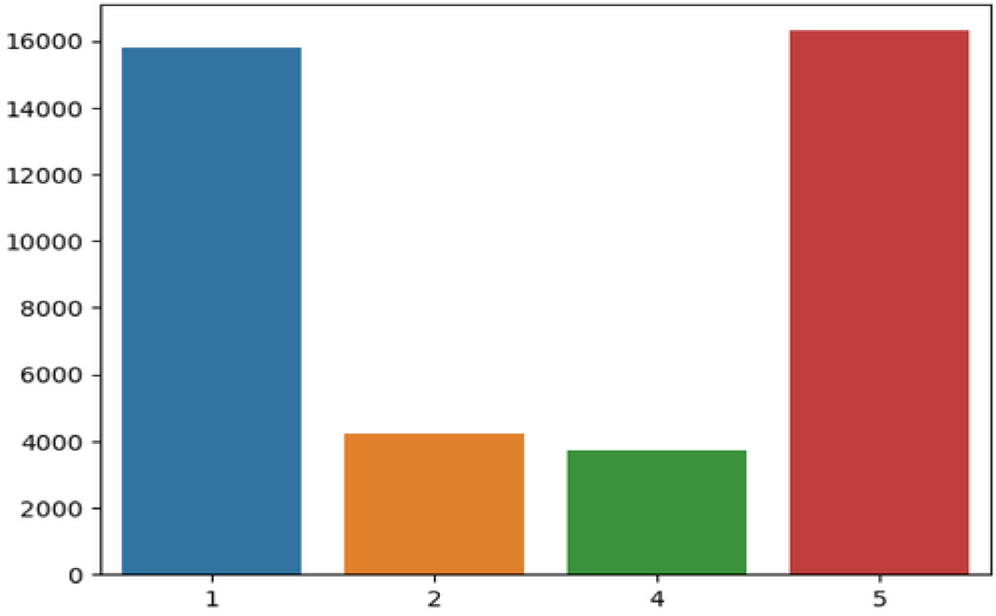

Figure 1: Data Visualization

4.4 DATA COLLECTION

Data collection involves collecting and measuring information from various sources. To create effective solutions using artificial intelligence (AI) and machine learning, data must be collected and stored in a way that aligns with the business problem. Labeled data refers to data where the target answer is already known.

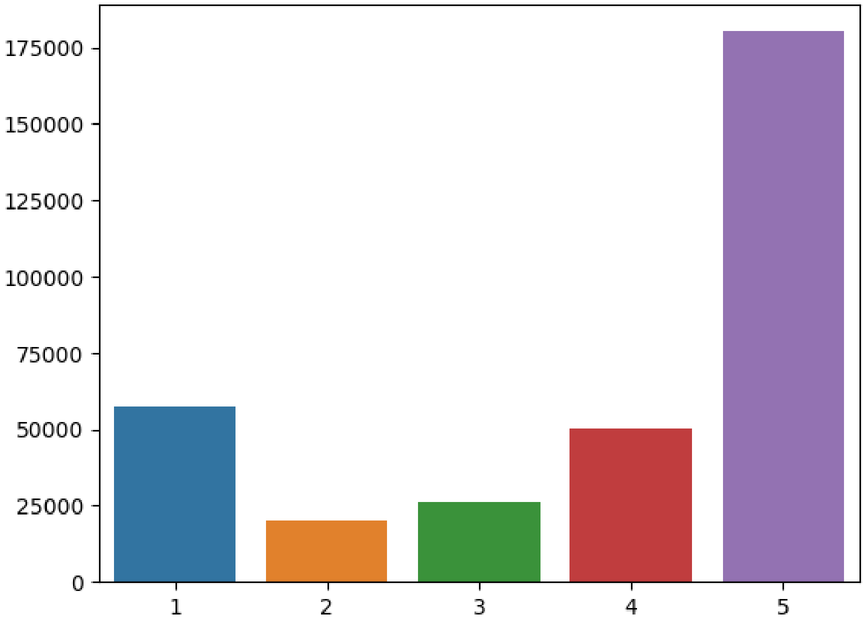

Figure 2: Data Visualization

4.5 DATA PRE-PROCESSING

Organize your selected data by formatting, cleaning and sampling from it. Three common data pre-processing steps are: Formatting, Cleaning, Sampling.

V. CONCLUSION:

In this paper, the use of logistic regression model for detecting fake reviews through review-centric features is discussed. Besides the review content, a set of review-centric features is provided for classification of fake reviews. One such feature proposed is the "verified purchase". The research shows that the inclusion of "verified purchase" as a feature for classifying fake reviews has a significant impact. Additionally, two feature extraction techniques, namely Tf-idf and Count Vectorizer, are proposed. It is concluded that the implementation of logistic regression with Count Vectorizer on the given dataset achieves an accuracy of 82%, whereas an accuracy of 81% is obtained with Tf-idf.

REFERENCES:

[1] A. Mukherjee, B. Liu, and N. Glance, ‘‘Spotting fake reviewer groups in consumer reviews,’’ in Proc. 21st Int. Conf. World Wide Web (WWW), 2012, pp. 191–200.

[2] A. Mukherjee, V. Venkataraman, B. Liu, and N. Glance, ‘‘What yelp fake review filter might be doing,’’ in Proc. Int. AAAI Conf. Web Social Media, vol. 7, 2013.

[3] D. H. Fusilier, M. M.-Y. Gomez, P. Rosso, and R. G. Cabrera, ‘‘Detecting positive and negative deceptive opinions using PU-learning,’’ Inf. Process. Manage, vol. 51, no. 4, pp. 433–443, 2015

[4] H. A. Najada and X. Zhu, ‘‘ISRD: Spam review detection with imbalanced data distributions,’’ in Proc. IEEE 15th Int. Conf. Inf. Reuse Integer. (IEEE IRI), Aug. 2014, pp. 553–560.

[5] N. J. Conroy, V. L. Rubin, and Y. Chen, ‘‘Automatic deception detection: Methods for finding fake news,’’ in Proc. 78th ASIST An nu. Meeting, Inf. Sci. Impact, Res. Community, vol. 52, no. 1, 2015, pp. 1–4.

[6] P. Kaghazgaran, J. Caverlee, and M. Alfifi, ‘‘Behavioral analysis of review fraud: Linking malicious crowd sourcing to Amazon and beyond,’’ in Proc. Int. AAAI Conf. Web Social Media, vol. 11, 2017.

[7] W. Etaiwi and G. Naymat, ‘‘the impact of applying different preprocessing steps on review spam detection,’’ Proc. Computer. Sci., vol. 113, pp. 273–279, Jan. 2017. J. Mach. Learn. Res., vol. 7, pp. 1–30, Jan. 2006.