

A REVIEW BASED MACHINE LEARNING APPROACH TO DETECT THE FAKE ONLINE REVIEWS

B. VENKATESWARA REDDY¹, D.NARENDRA², CH. NAGASAIKUMAR³, CH.
LAVANYA⁴, D.KALLADAR⁵.

¹Assistant Professor, CSE, Chalapathi Institute of Technology, Guntur, India

²UG Student, CSE, Chalapathi Institute of Technology, Guntur, India

³UG Student, CSE, Chalapathi Institute of Technology, Guntur, India

⁴UG Student, CSE, Chalapathi Institute of Technology, Guntur, India

⁵UG Student, CSE, Chalapathi Institute of Technology, Guntur, India

ABSTRACT: This project deals with detecting the Fake Online Reviews. With the continuous evolve of Ecommerce systems, online reviews are mainly considered as a crucial factor for building and maintaining a good reputation. Moreover, they have an efficient role in the decision making process for end users. Usually, a positive review for a target object attracts more customers and lead to high increment in sales. Nowadays, deceptive or fake reviews are deliberately written to build virtual reputation and attracting potential customers. Thus, identifying fake reviews is a vivid and ongoing research area. Identifying fake reviews depends not only on the key features of the reviews but also on the behaviors of the reviewers.

We propose a machine learning approach to identify fake reviews. In addition to the features extraction process of the reviews, this project applies several features engineering to extract various behaviors of the reviewers. The paper compares the performance of several experiments done on a real Yelp dataset of restaurants reviews, Naive Bayes (NB). The results reveal that Logistic Regression outperforms the rest of classifiers in terms of accuracy achieving best. The results show that the system has better ability to detect a review as fake or original.

1. INTRODUCTION

Nowadays, when customers want to draw a decision about services or products, reviews

become the main source of their information. For example, when customers take the initiation to book a hotel, they read the reviews on the opinions of other customers on the hotel services. Depending on the feedback of the reviews, they decide to book room or not. If they came to a positive feedback from the reviews, they probably proceed to book the room. Thus, historical reviews became very credible sources of information to most people in several online services. Since, reviews are considered forms of sharing authentic feedback about positive or negative services, any attempt to manipulate those reviews by writing misleading or inauthentic content is considered as deceptive action and such reviews are labelled as fake. Such case leads us to think what if not all the written reviews are honest or credible. What if some of these reviews are fake. Thus, detecting fake review has become and still in the state of active and required research area. The rise of social media has blurred the line between authentic content and advertising, leading to an explosion in deceptive endorsements across the marketplace. Fake online reviews and other deceptive endorsements often tout products throughout the online world. Consequently, the FTC is now using its Penalty Offense Authority to remind advertisers of the law and deter them from breaking it.

“Fake reviews and other forms of deceptive endorsements cheat consumers and undercut

honest businesses,” said Samuel Levine, Director of the FTC’s Bureau of Consumer Protection. “Advertisers will pay a price if they engage in these deceptive practices.” To this end, this project applies several machine learning classifiers to identify fake

reviews based on the content of the reviews as well as several extracted features from the

reviewers. We apply the classifiers on real corpus of reviews taken from open-source sites. Besides the normal natural language processing on the corpus to extract and feed the features of the reviews to the classifiers, the paper also applies several features engineering on the corpus to extract various behaviors of the reviewers. The paper compares the impact of extracted features of the reviewers if they are taken into consideration within the classifiers. The papers compare the results in the absence and the presence of the extracted features in two different language models namely TF-IDF. The results indicate that the engineered features increase the performance of fake reviews detection process.

The rapid growth of the Internet influenced many of our daily activities. One of the very rapid growth area is e-commerce. Generally, e-commerce provide facility for customers to write reviews related with its service. The existence of these reviews can be used as a source of information. For examples, companies can use it to make design decisions of their products or services, while potential customers can use it to decide either to buy or to use a product. Unfortunately, the importance of the review is misused by certain parties who tried to create fake reviews, both aimed at raising the popularity or to discredit the product. This research aims to detect fake reviews for a product by using the text and rating property from a review. Machine learning techniques can

provide a big contribution to detect fake reviews of web contents. Generally, web mining techniques find and extract useful information using several machine learning algorithms. One of the web mining tasks is content mining. A traditional example of content mining is opinion mining which is concerned of finding the sentiment of text (positive or negative) by machine learning where a classifier is trained to analyze the features of the reviews together with the sentiments. Usually, fake reviews detection depends not only on the category of reviews but also on certain features that are not directly connected to the content. Building features of reviews normally involves text and natural language processing NLP. However, fake reviews may require building other features linked to the reviewer himself like for example review time/date or his writing styles. Thus, the successful fake reviews detection lies on the construction of meaningful features extraction of the reviewers. Usually, fake reviews detection depends not only on the category of reviews but also on

certain features that are not directly connected to the content. Building features of reviews normally involves text and natural language processing NLP. However, fake reviews may require building other features linked to the reviewer himself like for example review time/date or his writing styles. Thus, the successful fake reviews detection lies on the construction of meaningful features extraction of the reviewers.

2. LITERATURE SURVEY

“Evaluating Machine Learning algorithms for Fake News Detection”

In this article, the author introduced the concept of the importance of NLP in stumbling across incorrect information. They have used time frequency-inverse document frequency (TF-IDF) of bigrams and probabilistic context-free grammar

detection. Shloka Gilda introduced the concept of the importance of NLP in stumbling over incorrect information. They used Bi-Gram Count Vectorizer and Probabilistic Context-Free Grammar (PCFG) to detect deceptions. They examined the data set in more than one class of algorithms to find out a better model. The count victimizer of bi-grams fed directly into a stochastic gradient descent model which identifies non credible resources with an accuracy of 71.2%.

3 EXISTING SYSTEM

There exists a large body of research on the topic of machine learning methods for deception detection, most of it has been focusing on classifying online reviews and publicly available social media posts. Particularly since late 2016 during the American Presidential election, the question of determining ‘fake news’ has also been the subject of particular attention within the literature. Conroy, Rubin, and Chen outlines several approaches that seem promising towards the aim of perfectly classify the misleading articles. They note that simple content-related n-grams and shallow parts-of-speech (POS) tagging have proven insufficient for the classification task, often failing to account for important context information. Rather, these methods have been shown useful only in tandem with more complex methods of analysis. Deep Syntax analysis using Probabilistic Context Free Grammars (PCFG) have been shown to be particularly valuable in combination with n-gram methods. Feng, Banerjee, and Choi are able to achieve 85%-91% accuracy in deception related classification tasks using online implemented a semantic analysis looking at ‘object: descriptor’ pairs for contradictions with the text on top of Feng’s initial deep syntax model for additional improvement. employ language pattern similarity networks requiring a pre-existing knowledge base.

4. PROPOSED SYSTEM

In this we have introduced some supervised machine learning classification techniques to detect fake online reviews with a good accuracy in the proposed system. each review goes through tokenization process first. Then, unnecessary words are removed and candidate feature words are generated. Each candidate feature words are checked against the dictionary and if its entry is available in the dictionary then its frequency is counted and added to the column in the feature vector that corresponds the numeric map of the word.

Methodology

This project explains the system which is developed in three parts. The first part is static which works on machine learning classifier. We studied and trained the model with 4 different classifiers and chose the best classifier for final execution. The second part is dynamic which takes the keyword/text from user and searches online for the truth probability of the news. The third part provides the authenticity of the URL input by user.

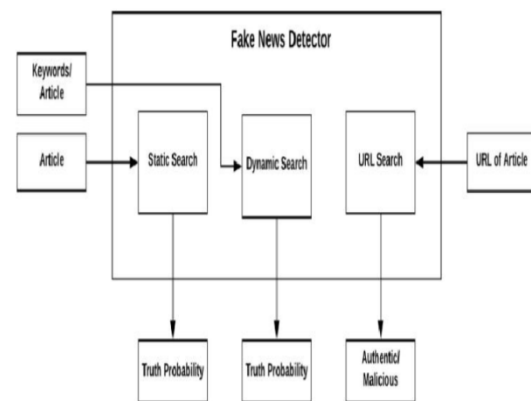


Fig: System Architecture

5. IMPLEMENTATION

1. Data Collection:

This module performs tasks related to gathering the information for this purpose we design a web crawler which extracts all the links from the page. Once parsed, the information is stored in the MySQL

database.

2. PROPOSED FRAMEWORK:

In this project, we propose this application that can be considered a useful system since it helps to reduce the limitations obtained from traditional and other existing methods. The objective of this study is to develop a fast and reliable method which detects and estimates accurately. To design this system, we used a powerful algorithm in a Python-based environment with Django framework.

3. DATA PRE-PROCESSING

The data which is collected is not consistent using various machine learning algorithms. The data is trained and presented in a particular format. The first step in the proposed approach is data pre-processing; one of the essential steps in machine learning approaches. Data pre-processing is a critical activity as the world data is never appropriate to be used. A sequence of pre-processing steps has been used in this work to prepare the raw data of the Yelp dataset for computational activities.

4. Prediction of Disease:

Feature extraction is a step which aims to increase the performance either for pattern recognition or machine learning system. Feature extraction represents a reduction phase of the data to its important features which yields in feeding machine and deep learning models with more valuable data. It is mainly a procedure of removing the unneeded attributes from data that may actually reduce the accuracy of the model.

6. RESULT



7. CONCLUSION:

In this project, we showed the importance of reviews and how they affect almost everything related to web-based data. It is obvious that reviews play a crucial role in people's decision. Thus, fake reviews detection is a vivid and ongoing research area. In this project, a machine learning fake reviews detection approach is presented. In the proposed approach, both the features of the reviews and the behavioral features of the reviewers are considered. The Yelp dataset is used to evaluate the proposed approach. Different classifiers are implemented in the developed approach. We have successfully developed a system to detect fake reviews in this application. This is created in a user-friendly environment with Python programming and Django framework. The

system is likely to gather data from the user in order to determine whether the review is fake or not.

Future work may consider including other behavioral features such as features that depend on the frequent times the reviewers do the reviews, the time reviewers take to complete reviews, and how frequent they are submitting positive or negative reviews. It is highly expected that considering more behavioral features will enhance the performance of the presented fake reviews detection approach.

FUTURE SCOPE:

The future scope indicate that fake Robert can successfully predict both computer and human-generated reviews. However, ML models face the general caveat of dataset specificity, so as the nature of human-generated reviews evolves over time, the only way to maintain a highperformance is frequently updating the classifiers. Therefore, future work needs to pursue the creation of trustworthy baseline datasets, especially of large human generated fake reviews. Also, because the nature of communication (e.g., micro-text reviews in Twitter versus longer reviews in Amazon) differs by platform, the applicability of fake detection classifiers across platforms should be examined. Again, this requires that not only e-commerce product reviews but also other forms of reviews taking place in social media are included in fake review datasets.

REFERENCES:

1. R. Barbado, O. Araque, and C. A. Iglesias, "A framework for fake review detection in online consumer electronics retailers," *Information Processing & Management*, vol. 56, no. 4, pp. 1234 – 1244, 2019.
2. S. Tadelis, "The economics of reputation and feedback systems in e-commerce marketplaces," *IEEE Internet Computing*, vol. 20, no. 1, pp. 12–19, 2016.
3. M. J. H. Mughal, "Data mining: Web data mining techniques, tools and algorithms: An overview," *Information Retrieval*, vol. 9, no. 6, 2018.
4. C. Aggarwal, "Opinion mining and sentiment analysis," in *Machine Learning for Text*. Springer, 2018, pp. 413–434.
5. A. Mukherjee, V. Venkataraman, B. Liu, and N. Glance, "What yelp fake review filter might be doing?" in *Seventh international AAAI conference on weblogs and social media*, 2013.
6. N. Jindal and B. Liu, "Review spam detection," in *Proceedings of the 16th International Conference on World Wide Web*, ser. WWW '07, 2007.
7. E. Elmurugi and A. Gherbi, *Detecting Fake Reviews through Sentiment Analysis Using Machine Learning Techniques*. IARIA/DATA ANA- LYTICS, 2017.
8. V. Singh, R. Piryani, A. Uddin, and P. Waila, "Sentiment analysis of movie reviews and blogposts," in *Advance Computing Conference (IACC)*, 2013, pp. 893–898.
9. A. Molla, Y. Biadgie, and K.-A. Sohn, "Detecting Negative Deceptive Opinion from Tweets." in *International Conference on Mobile and Wireless Technology*. Singapore: Springer, 2017.
10. S. Shojaee et al., "Detecting deceptive reviews using lexical and syntactic features." 2013.
11. Y. Ren and D. Ji, "Neural networks for deceptive opinion spam detection: An empirical study," *Information Sciences*, vol. 385, pp. 213– 224, 2017.
12. H. Li et al., "Spotting fake reviews via collective positive-unlabelled learning." 2014.

13.