Detection of Fake Online Reviews and Automation in Comments on E-Commerce Platform using Supervised Learning

Gaikwad Ameya Vinod[1], Aditya Shankar Jha[2], Varpe Tejas Kisan[3], Wagh Sakshi Santosh[4], Prof. S. S. Pawar[5]

[1,2,3,4]Student, Dept. of Computer Engineering, Sinhgad Academy of Engineering, Pune, Maharashtra, India
[5]Professor (Guide), Dept. of Computer Engineering, Sinhgad Academy of Engineering, Pune, Maharashtra, India

---***---

Abstract – New technologies are enabling people to have their work across different domains done efficiently. There have been various discoveries or evolution to this effect, and one such evolution is the online e-commerce platforms. These online e-commerce platform is used for online shopping. While buying from these e-commerce platforms reviews are of great importance for the buyer or the user as well as the e-commerce company. These reviews can have a positive or negative impact on the overall business and commerce. Therefore the reviews given by the end user is of great importance and it influence's the other user's when they purchase online products. Sometimes individuals with malicious intentions try to influence the reviews by giving fake reviews. Here we introduce a system to improve detecting of fake online reviews. We introduce some semi supervised text mining models for detection of fake online reviews. E- Commerce platforms helps us to share our views and ideas about the product. We further classify the reviews into Positive or Negative using lexical analysis. Our model also classifies the different words in comments in six classes which are Toxic, Severe Toxic, Obscene, Threat, Insult and Identity-hate. Multi-Label Classification helps us to provide an automatic solution to deal with the problem of abusive words in the review, because abusive or toxic words being used in reviews have a negative bearing on the reputation of the company, so we classify them using Multi-Label Classification.

*Keywords***:** *Fake review, Lexical Analysis, Multi-Label classification, detection.*

1. INTRODUCTION

Technologies are changing rapidly. Old technologies are continuously being replaced by new and complicated ones. These new technologies are enabling people to possess their work done efficiently. Such an evolution of technology is online marketplace. We can shop and make using online websites. Almost, every one of us checks out reviews before purchasing some products or services. Hence, online reviews became an excellent source of reputation for the businesses. Also, they need large impact on advertisement and promotion of products and services. With the spread of online marketplace, fake online reviews are getting great matter of concern. People can make false reviews for promotion of their own products that harms the particular users. Also, competitive companies can attempt to damage each other's reputation by providing fake reviews.

We make some classification approaches for detecting fake online reviews, some of which are semi supervised and others are supervised. For semi-supervised learning, we use statistical Naive Bayes classifier are used as classifiers in our research work to improve the performance of classification. We have mainly focused on the content of the review based approaches. We also use lexical analysis for classifying review into positive and negative comments, and besides this we have used Multi-Label Classification for classifying the abusive words into different classes.

2. LITERATURE SURVEY

Paper Name: Automation in Social Networking Comments with the Help of Robust fastText and CNN

Authors: Suresh Mestry, Hargun Singh, Roshan Chauhan, Vishal Bisht, Kaushik Tiwari

Description: Social networking and online conversation platforms provide us with the power to share our views and ideas. However, nowadays on social media platforms, many people are taking these platforms for granted, they see it as an opportunity to harass and target others leading to cyber-attack and cyber-bullying which lead to traumatic experiences and suicidal attempts in extreme cases. Manually identifying and classifying such comments is a very long, tiresome and unreliable process. To solve this challenge, we have developed a deep learning system which will identify such negative content on online

discussion platforms and successfully classify them into proper labels.

Paper Name: Detection of fake online reviews using semi-supervised and supervised learning

Authors: Rakibul Hassan, Md. Rabiul Islam

Description: Online reviews have great impact on today's business and commerce. Decision making for purchase of online products mostly depends on reviews given by the users. Hence, opportunistic individuals or groups try to manipulate product reviews for their own interests. This paper introduces some semi-supervised and supervised text mining models to detect fake on-line reviews as well as compares the efficiency of both techniques on dataset containing hotel reviews.

Paper Name: SPR2EP: A Semi-Supervised Spam Review Detection Frame-work

Authors: Cennet Merve Ylmaz, Ahmet Onur Durahim

Description: Authenticity and reliability of the information spread over the cyberspace is becoming increasingly important. This is especially important in e-commerce since potential customers check reviews and customer feedbacks before making a purchase. Although this information is easily accessible through related websites, lack of verification of the authenticity of these reviews raises concerns about their reliability. Besides, fraudulent users disseminate misinformation to deceive people into acting against their interest. So, detection of fake and unreliable reviews is a crucial problem that must be addressed by the security researchers.

Paper Name: Revisiting Semi-supervised Learning for Online Deceptive Review Detection

Authors: Jitendra Kumar Rout, Anmol Dalmia, Kim-Kwang Raymond Choo, Sambit Bakshi, and Sanjay Kumar Jena

Description: With more consumers using online opinion reviews to inform their service decision making, opinion reviews have an economic impact on the bottom line of businesses. Unsurprisingly, opportunistic individuals or groups have attempted to abuse or manipulate online opinion reviews (e.g. spam reviews) to make pro ts, etc, and that detecting deceptive and fake opinion reviews is a topic of ongoing research interest. In this paper, we explain how semi-supervised learning methods can be used to detect spam reviews, prior to demonstrating its utility using a dataset of hotel reviews.

Paper Name: A SURVEY ON ONLINE REVIEW SPAM DETECTION TECHNIQUES

Authors: S. P. Rajamohana, Dr. K. Umamaheswari, M. Dharani, R. Vedackshya

Description: The development of Internet influenced many of our day-to-day activities. Ecommerce is one of the rapid growth areas in the Internet era. People are eager to buy products from online sites like Amazon, e-bay, Flipkart etc. Online sites also provide facility for customers to write review on products they buy. These reviews help consumers and vendors for making decision on marketing strategies, and the improvement of products and services. Nowadays people are very much interested to read reviews before purchasing any product and getting services.

3. Project Overview

For detection of fake online reviews, we start with raw text data. We have used a dataset which was already labeled by the previous researchers. We remove un-necessary texts like article and prepositions in the data. Then these text data are converted into numeric data for making them suitable for the classifier.

3.1 Proposed System:

In this system, we make some classification approaches for detecting fake online reviews, some of which are semi supervised. For semi-supervised learning, we use Statistical Naive Bayes classifier and are used as classifiers in our research work to improve the performance of classification. We have mainly focused on the content of the review based approaches. As feature we have used word frequency count, sentiment polarity. In this system, input as comments will be fed from social sites which will be analyzed and sent to word embedding phase. In this phase, sentences are broken into words and embedded into vectors.

3.2 Analysis Model:

The Waterfall model is used in this project.
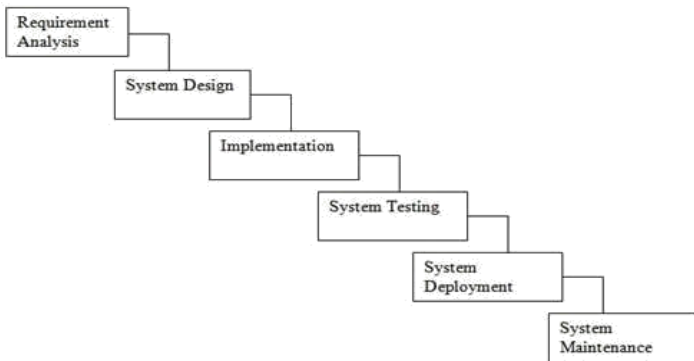
1. Requirement Gathering and Analysis: Hardware and Software requirements are analysed and gathered in this step of Waterfall model.

2. System Design: Some UML diagrams and data flow diagrams are designed in this phase so that end user can understand the flow of system and modules of system and sequence of execution.

3. Implementation: In this phase, various modules are implemented for the expected outcomes at different module levels. Different inputs from system design is developed into small units which are integrated in next phase. Every units are fist developed and then tasted for its functionality.

4. Testing: In implementation phase all units are integrated for testing. Later in this phase, the whole system is tested to detect any fault and failures that may arise.

5. Deployment of System: After the successful testing of functional and non-functional requirements, the software is successfully deployed to the customers.

6. Maintenance: In this phase maintenance is done to deliver change in the user environment. The different phases are cascaded to each other and appear like a waterfall .In this model the different phases are not overlapped and one phase is started only after the desired goals are achieved for the previous phase.
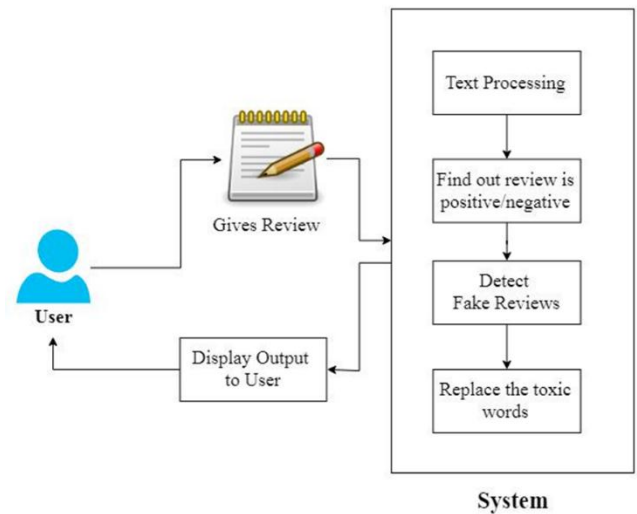
3.3 System Architecture:

Fig. System Architecture

3.4 Advantages:

1. User friendly system i.e. this system that we have proposed in this paper is very easy to use for the user.
2. Through this paper, we enable a way for e-commerce platforms to weed out fake reviews so that the reviews have no impact on the sales of an e-commerce platform.
3. Through this paper, we provide an automated solution for the abusive or toxic words problem people are facing while shopping on these platforms as we successfully removed the toxic or abusive words.
4. This paper provides a technique to classify positive and negative comments.

4. CONCLUSION

We have shown several supervised text mining techniques for detecting fake online reviews in this research. We have combined features from several research works to create a better feature set. Also we have tried some other classifiers that were not used on the previous work. Thus, we have been able to increase the accuracy of previous semi supervised techniques done by Jiten et al. Our model based on Naïve Bayes and Support Vector Machine classification algorithms used to detect fake reviews and classify it into positive or negative and replace vulgar words with asterisk.

5. REFERENCES

[1] Chengai Sun, Qiaolin Du and Gang Tian, Exploiting Product Related Review Features for Fake Review Detection, Mathematical Problems in Engineering, 2016.

[2] A. Heydari, M. A. Tavakoli, N. Salim, and Z. Heydari, Detection of review spam: a survey, Expert Systems with Applications, vol. 42, no. 7, pp. 36343642, 2015.

[3] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, Finding deceptive opinion spam by any stretch of the imagination, in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL-HLT), vol. 1, pp. 309319, Association for Computational Linguistics, Portland.

[4] J. W. Pennebaker, M. E. Francis, and R. J. Booth, Linguistic Inquiry and Word Count: Liwc, vol. 71, 2001.

[5] S. Feng, R. Banerjee, and Y. Choi, Syntactic stylometry for deception detection, in Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers, Vol. 2, 2012.

[6] J. Li, M. Ott, C. Cardie, and E. Hovy, Towards a general rule for identifying deceptive opinion spam, in Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL), 2014.

[7] E. P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw, Detecting product review spammers using rating behaviours, in Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM), 2010.