International Journal of Innovative Research in Science and Engineering
Volume 9, Issue 04, 2023
www.ijirse.com
ISSN: 2454-9665

Detection of Fake Online Reviews Using Semi supervised and Supervised Learning

E. Bharathi Devi, A.Ramya, A.Sarath, K.Yaswanthi

*(Under the guidance of **Mrs. V. Prathima,** Assistant Professor, B. Tech , M. Tech..,*
Department of Information Technology, Tirumala Engineering College)

ABSTRACT

Consumers review on ecommerce websites , onlineservices rating and experience stories are useful for the user as well as the vendor.The reviewer can increse their brands loyalty and help other customers understand their experience with the product. Similarly reviews help the vendors gain more profiles by incresing their sales of products, if consumers leave positive feedback on their product review. But unfortunately, these review mechanisms can be misused by vendors.

For example ,one may create fake positive reviews to promote brands reputation or try to demote competitors products by leaving fake negative reviews on their product .Existing solutions with supervised include application of different machine learning algorithms and different tools like weak.

Unlike the existing work, instead of using a constrained dataset I chose to have a wide variety of vocabulary to work on such as different subjects of datasets combined as one big data set. Sentiment analysis has been incorporated based on emojis and text content in the reviews. Fake reviews are detected and categorized. The testing results are obtained through theapplication of Naïve Bayes, Linear SVC, Support Vector Machine and Random forest algorithms. The implemented (proposed) solution is to classify these reviews into fake or genuine. The highest accuracy is obtained by using Naïve Bayes by including sentiment classifier.

INTRODUCTION

Everyone can freely express his/her views and opinions anonymously and without the fear of consequences. Social media and online posting have made it even easier to post confidently and openly. These opinions have both pros and cons while providing the right feedback to reach the right person which can help fix the issue and sometimes a con when these get manipulated These opinions are regarded as valuable. This allows people with malicious intentions to easily make the system to give people the impression of genuineness and post opinions to promote their own product or to discredit the competitor products and services, without revealing identity of themselves or the organization they work for. Such people are called opinion spammers and these activities can be termed as opinion spamming.

There are few different types of opinion spamming. One type is givingpositive opinions to some products with intention to promote giving untrue or negative reviews to products to damage their reputation. Second type consists of advertisements with no opinions on product. There is lot of research work done in field of sentiment analysis and created models while using different sentiment analysis on data from various sources, but the primary focus is on the algorithms and not on actual fake review detection. One of many other research works by E. I. Elmurngi and A. Gherbi [1] used machine learning algorithms to classify the product reviews on

ISSN: 2454-9665

Amazon.com dataset [2] including customer usage of the product and buying experiences. The use of Opinion Mining, a type of language processing to track the emotion and thought process of the people or users about a product which can in turn help research work.

One of the biggest applications of opinion mining is in the online and e-commerce reviews of consumer products, feedback and services. As these opinions are so helpful for both the user as well as the seller the e-commerce web sites suggest their customers to leave a feedback and review about their product or service they purchased. These reviews provide valuable information that is used by potential customers to know the opinions of previous or current users before they decide to purchase that product from that seller. Similarly, the seller or service providers use this information to identify any defects or problems users face with their products and to understand the competitive information to know the difference about their similar competitors' products.

LITERATURE SURVEY

we briefly review several related works on fake review detection, including classification methods, approaches for addressing data imbalances, and feature selec- tion methods. We also present the problems with existing studies.Regarding classification methods, machine learning methods are the most frequently used subtype for fake review detec- tion. Machine learning can be classified into two categories: supervised learning and unsupervised learning. Supervised learning is the dominant approach in the field of fake review detection.

There are many super- vised learning algorithms, and it isnot easy to decide which one is the best. Apart from supervised learning, some researchers use unsupervised learning methods or deep learning methods to identify fake reviews due to the difficulty of labeling data. However, considering the lack of large scale datasets, deep learning methods might not be effective. In addition,ensemble learning methods have also been proposed. Ruan *et al.* proposed an ensemble model using the geolocation information of users.

The results showed that the ensemble model could enhance the stability over those of the base models. However, ensemble strategies are still understudied.

Proposed System:

1. System can intimate the driver by making some huge amount of buzz sound
2. Measures and detects distance of moving vehicles
3. Ensures total safety

Hardware requirements:

` 1.Processor: Intel(R)Core i5-7200U@2.7GHz

2. RAM: 8GB
3. System: 64bits,x64 processor
4. 512 SSD Storage

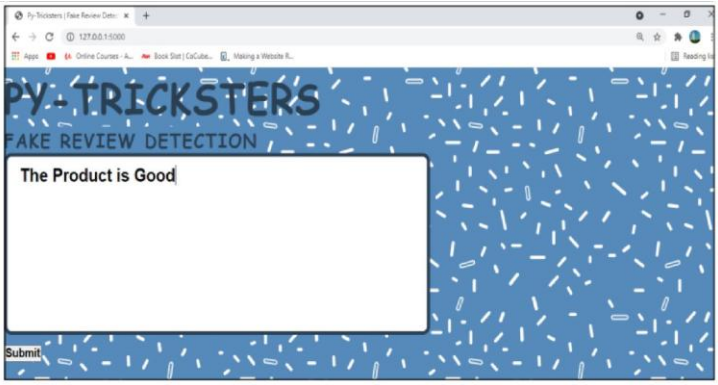
Software configuration:

1. Windows 10

ISSN: 2454-9665

2. python 3.5.2
3. Different libraries are available in python that helps in machine learning,classification Projects .Several of those libraries have improved the performance of this project. Few of themare mentioned in this section .
4. First,"Numpy" that provides with high-level math function collection to supportmulti- Dimensional matrics and arrays .This is used for fast computations over the weights(gradients) in neural networks.
5. Second, "scikit-learn" is a machine learning library for python which features differentAlgorithms and machine learning functions packages
6. NLTK, natural language toolkit is healpful in word processing and tokenization.

RESULTS



CONCLUSION

The fake review detection is designed for filtering the fake reviews. In this research work SVM classification provided a better accuracy of classifying than the Naïve Bayes classifier for testing dataset. On the other hand, the Naïve Bayes classifier has performed better than otheralgorithms on the training data. Revealing that it can generalize better and predict the fake reviews efficiently. This method can be applied over other sampled instances of the dataset. The data visualization helped in exploring the dataset and the features identified contributed to the accuracy of the classification. The various algorithms used, and their accuracies show how each of them have performed based on their accuracy factors.

Also, the approach provides the user with a functionality to recommend the most truthful reviews to enable the purchaser to make decisions about the product. Various factors such as adding new vectors like ratings, emojis, verified purchase have affected the accuracy of classifying the data better.

REFERENCES

1) E. I. Elmuring and A.Gherbi,"Unfair Reviews Detection on Amazon Reviews usingSentiment Analysis with Supervised Learning Techniques," Journals of Computer Science,vol.14,no.5,pp.714-726,june 2018.
2) J. Leskovec," WebData Amazon reviews," [online].

ISSN: 2454-9665

3) J. Li, M.Ott, C. Cardie and E.Hovy,"Towards a General Rule for Identifying DeceptiveOpinion Spam," in Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, Baltimore, MD, USA, 1566-1576,November 2014.

4) C. Hill, "10 Secretsto Uncovering which Online Reviews are Fake," [Online].