

Fake Opinions on Online Products Detection Using Random Forest Classifier

Nandhakumar C¹, Shanthosh N², Kumaresan M³, Prasanth M⁴

Assistant Professor, Department of Information Technology¹

Students, Department of Information Technology^{2,3,4}

Dr. Mahalingam College of Engineering and Technology, NPT-MCET, Pollachi

Abstract: *The internet-based audits can change choice of the client and they can settle an item by contrasting and the brands of items, the client can choose the item and fulfill their prerequisite provided that the surveys are not phony. Then again in the event that the surveys are phony, it deludes the client to take care of this issue the distinguishing proof of the phony suppositions from the client should be separated. Commentators' ways of behaving are extricated based the semantically investigation of his survey content to distinguish the audit as phony or not. In this work the surveys are removed from the web for a specific item, alongside the audits of a few other data connected with the analysts likewise been extricated to recognize the phony commentators utilizing k-implies bunching Information Gain and Classifier. Data gain is used to validate the meaning of the highlights on the option. Tests were conducted on a large number of surveys taken from the internet, which demonstrated the effectiveness of the suggested method.*

I. INTRODUCTION

In the Modern days many items were ordered or bought through online platforms or websites where are some people who buy the product without viewing the review about the product and some are used to see the reviews about the product before buying the product at the time if there is any fake reviews in the specific product it may affect the product growth and sentiment of the product. So, to overcome these specific issues we have created a specific API on e-commerce site to reduction of review and the reviewers on a product.

II. EXISTING SYSTEM

Fake review detection based on rules Manually creating relevant feature rules and conducting fake review detection based on professional knowledge and expertise is what rule-based fake review detection entails. Review reviewers and stores to form a network structure for nodes and utilises iterative computation to find garbage reviews. Machine learning approaches for detecting fake reviews include the Bayesian classification algorithm, logistic regression, Korder closest neighbour algorithm, support vector machine, and others.

III. LITERATURE SURVEY

ADNAN ISHAQ, SOHAIL ASGHAR, SAIRA ANDLEEB GILLANI are the authors of this work, which is titled Aspect-Based Sentiment Analysis Using a Hybridized Approach. CNN and GA were used. classifier- Opinions are extracted using a convolutional neural network (CNN). A multi-objective function is utilised to tune CNN parameters, and the dataset is classified using a Genetic Algorithm. Fake review prevention has not been completed, and project accuracy has not been reached. The authors of this research are S. Banerjee, A. Chua, and J. Kim, and the title of the study is Using Supervised Learning to Classify Authentic and Fake Online Reviews. 2015 Random Forest, Support vector machine, and Naive bayes are supervised machine learning algorithms that are used to categorise the data set. Linguistic hints, writing style, and level of difficulty are some of the features.

IV. PROPOSED SYSTEM

Propose F2RSpotAPI as a new model. Fake reviews are detected and eliminated using Ensemble Learning. To prevent false reviews, the Genuine Reviewer Protocol is utilised. In this model we proposed the random forest classifier to classify the genuine and fake review system and we proposed new module to prevent the fake review called fake review prevention and detection. This is the application programming Interface can be implemented in the e-commerce sites

and its useful to prevent the fake reviews. Then it can be implemented any other sites to prevent fake reviews and fake opinion about the project. The advantages of our API is to prevention of the fake review and the prevention of the fake reviewer.

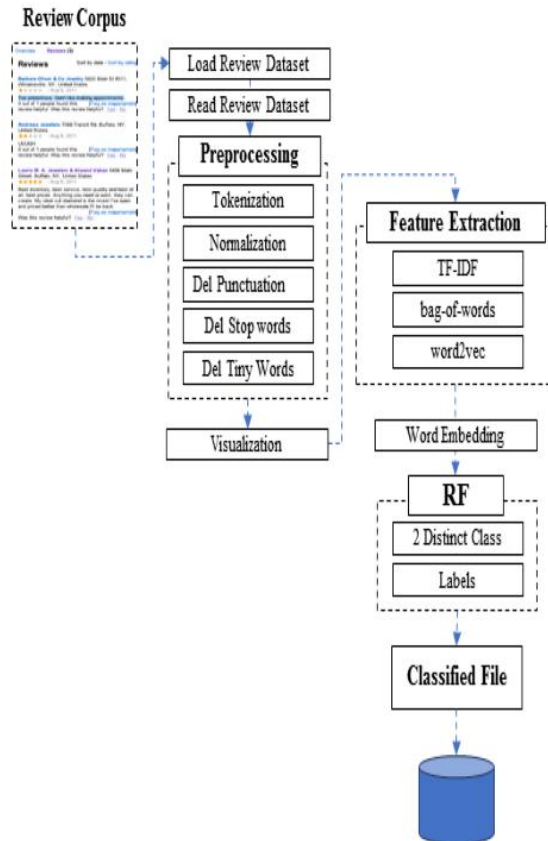


FIGURE 4.1: Proposed Architecture

1. Preprocessing

In this module Preprocessing was carried out to modify the text data appropriately in the experiment. We used decapitalization and did not mark the start and end of the sentences. The system deleted #, two or more spaces, tabs, Retweets (RT), and stop words. Convert text to vectors, our classifier only takes numerical data.

2. Feature Extraction

Tf-Idf Vectoriser from sklearn module to transform the matrix into a tf-idf representation. This is representation is commonly used in document classification and information retrieval. Sentiment score classification

3. RF Classification

A random forest is a meta estimator that employs averaging to increase predicted accuracy and control over-fitting by fitting a number of decision tree classifiers on various sub-samples of the dataset. We create a Random Forest Classifier instance, train it on our training data, and then test it on our test data. Then, using the heatmap function from seabornLabelling, create a confusion matrix.

```
#RFC=RandomforestClassifier
import from sklearn.ensemble RFC=RandomForestClassifier(random state=0) RandomForestClassifier fit(xv train, y train) RFC RandomForestClassifier(random state=0) RFC.predict(xv test) = pred rfc
```

V. PREVENTION ARCHITECTURE

According to the base paper where the prevention of the of fake review has not done but we have implemented the specific module to prevent the fake review in our project with the help of BOT API like OTP to prevent.

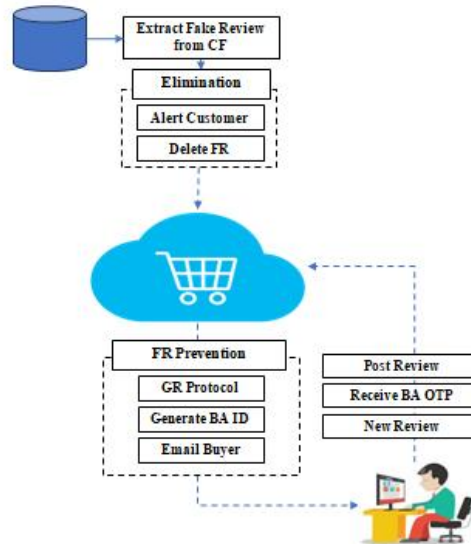


FIGURE 5.1: Prevention Architecture

1. Fake Review Prediction

In this module, user login to the E – Commerce sites, purchase or not purchase products and provide reviews about the products. And those reviews are tested for fake contents with the Amazon review datasets that has been trained.

2. Fake Review Elimination

Fake reviews are eradicated in this module by simply sending an email to the account holder of an ecommerce website to whom they did not purchase products but had left reviews. Simpler, more efficient, and faster convergence. A sentiment categorization strategy using both a supervised deep neural network and an unsupervised probabilistic generative model was proposed. Long and short text review datasets are used to assess performance. Improve your performance.

VI. CONCLUSION

In the Modern days many items were ordered or bought through online platforms or websites where are some people who buy the product without viewing the review about the product and some are used to see the reviews about the product before buying the product at the time if there are any fake reviews in the specific product it may affect the product growth and sentiment of the product. So, to overcome these specific issues we have created a specific API on e-commerce site to reduction of review and the reviewers on a product.

REFERENCES

- [1] J. Wang, H. Kan, F. Meng, Q. Mu, G. Shi, and X. Xiao, "Fake review detection based on multiple feature fusion and rolling collaborative training," *IEEE Access*, vol. 8, pp. 182625182639, 2020.
- [2] R. Barbado, O. Araque, and C. A. Iglesias, "A framework for fake review detection in online consumer electronics retailers," *Inf. Process. Manage.*, vol. 56, no. 4, pp. 12341244, Jul. 2019.
- [3] W. Zhang, Y. Du, T. Yoshida, and Q. Wang, "DRI-RCNN: An approach to deceptive review identification using recurrent convolutional neural network," *Inf. Process. Manage.*, vol. 54, no. 4, pp. 576592, Jul. 2018.
- [4] GurneetKaur and AbhinashSingla "Nostalgic Analysis of Flipkart audits utilizing Naïve Bayes and Decision Tree calculation" *International Journal of Advanced Research in Computer Engineering and Technology*, Volume 5 Issue 1, January 2016, ISSN: 2278 - 1323.



- [5] Kollishivagangadhar, Sagar H., SohanSathyan, Vanipriya C.H., "Fraud Detection in Online Reviews Using Machine Learning Techniques," International Journal of Computational Engineering Research, Volume, 05, Issue 05, May - 2015, ISSN (e): 2250 - 3005.Michael.
- [6] Manqing Dong, Linayao, Xianzhi Wang BoualemBenatallah, Chaoran Huang XiaodongNing "Assessment Fraud Detection through Neural AutoencoderDecision Forest" a School of Computer Science and Engineering, University of New South Wales, Sydney 2052, Australia, 2018, www.elsevier.com.