JOURNAL OF CRITICAL REVIEWS

ISSN- 2394-5125 VOL 7, ISSUE 15, 2020

FRAUDULENT REVIEWS DETECTION USING MACHINE LEARNING ALGORITHM

ILAKIYASELVAN N[1,*], SHARATH KUMAR J[2], SOURABH VERMA[3]

[1, 2, 3] School of Computer Science and Engineering,

Vellore Institute of Technology, Chennai, Tamilnadu 600127, India.

ilakiyaselvan.n@vit.ac.in[1,*]

Received: 24 April 2020, Revised: 1 June 2020, Accepted: 20 June 2020

Abstract

The development of Ecommerce is one of the rapid growth areas in the internet era. Reviews are playing a significant role in online shopping and likely to have a profound effect on success or survival. The reviewer's actions are collected to define the review based on the interpretation of its review material. Nowadays, people are very much interested in reading reviews before purchasing. Here, we commence on the phony reviews that use the method of decision tree and knowledge gain and to detect spam ratings. The classification criteria of opinion spammer detection are based on the characteristics and recognition of non-genuine reviews dependent on item's information. Let us begin with the crept dataset to assess the highlights of survey information, which shows that the examples of phony reviews for products are indistinguishable in ordinary circumstances. We separate the product review records into constituent or distinct points, and an afterward recognize non-genuine reviews. In this study, reviews for a particular product were extracted from the internet, and the reviews of a few different reviewers' data were gotten to group the phony reviews analysts utilizing choices like "Tree Classifier and Information Gain" and "Random forest," random forest classifier. This further helps in the decision tree classifier to make the number of trees to identify the more relevant and genuine reviews. These techniques are helping in the decision that reviews are genuine or not.

Keywords: Phony Reviews, genuine reviews, decision tree, information gain, random forest, trees.

1. Introduction

Nowadays, many phony reviews are posted on different company's products. The item reviews might be sure or negative; the negative comment would get more consideration than a positive comment to the shoppers. Such phony reviews can affect any product that results in financial loss or gain to any company. Reviews will generally come into view on websites such as Amazon, Flipkart etc. Based on the financial cause, there will be many phony reviews posted on those sites.

Purposefully, the organization proprietors will propel the individuals to make the spam reviews are presented on improving their business towards another item or product. Later, everyone focuses on online reviews to buy a product before purchasing. These reviews can influence the customer's decision, and they can finalize a product in a lack of understanding. The consumer can able to choose only the product that meets their requirements if the reviews are genuine. Phony reviews can mislead the consumer.

Along with this online survey is the origin of data about client feelings for any item. Genuine or non-genuine suggest to any hesitant and insignificant data about the item or administration. Non-genuine reason for existing is to promote and damage the good reputation of the items. Those reviews promote or downgrade the product with positive or negative words and deceive the clients. Those reviews are not identified with product item, not on the various highlights of the item or administrations. The analyst uses brand name over and again to advance a specific brand. Despite that, existing strategies for enormous information preparation cannot attainable take care of the issue of spam reviews.

2. Related Work

The phony reviews are logically impossible to read it. But the following pointers can make the customer believe on the reviews such as:

- Check availability of labeled data set (Most of the user-generated contents are unlabeled).

ISSN- 2394-5125 VOL 7, ISSUE 15, 2020

- Appropriate information selection.
- Cross-site verification (compare reviews of the same product across multiple sites.).
- Question answer with e-commerce sites.

There are many research papers published and distributed in the region of phony reviews verification. The main concentration to find those non-genuine reviews and make a clear understanding of the product in respect of the buyer's perspective. Online buyers are taking preference for reviews that are written by other people or buyers. People visit one or more websites to buy products and finalize the product based on ratings.

Well-known supervised and semi-supervised techniques are used for the detection of fake online reviews [1]. A structure has been introduced for phony online reviews recognizing model dependent on include examination [2]. The conduct of survey records, the result of review information, and a technique for the recognition of phony reviews are dependent on a brief component of survey and remarks [3]. The highlights of three perspectives, including metadata include the closeness highlight and viewpoint of customer investigation. A multi-angle include a based neural system model in identifying non-genuine, is considered [4]. For spam, reviews refer to any unaffected and unessential data about the item or company, proposed fake review identification utilizing AI method [5]. A procedure and the framework uses are to recognize the boosted reviews dependent on content examination, and item survey believability investigation is given [6]. An overview of the online survey spam recognition method found this checking on the framework urges a few people to enter their wrong feedback to advance a few items or criticize some others [7]. Honest reviews were accumulated, the spammer was entrusted to compose a beguiling an opinion on one of the 10 Chicago hotels, in the discovery of online hotel feedback [8]. Genuine case investigation on the interpersonal organization of sentiment saw that phony reviews harm buyers' trust to service. Product feedback sites have a focus on the presence of spam reviews.

Table 1. Research Algorithms

Algorithm	Description	Published Year
Decision Tree and Information Gain algorithm	Recognize the phony reviews by utilizing six distinct conditions that are star rating, reaction, answer, customer profile, profile status, layout condition.	2019
Spammer's Behaviour Feature	Observe and detect the review similarity between two feedbacks of reviews.	2017
Isolation Forest algorithm	The location strategy distinguishes the audits item by dissecting the brief patterns of feedback and remarks.	2019

The online locales to utilize clients for posting spam reviews on its site. It is hard to recognize whether they are fake or not fake [10]. In an improved design for include based sentiment mining from online product surveys [11], creators predict viewpoint-based conclusion mining utilizing a help vector machine classifier is utilized for extraction and synopsis of client reviews by utilizing affiliation standards dependent on a need calculation. Viewpoint based assessment mining is a fine-grained idea mining method, which concentrates item includes and predict its rating from the content review. Numerous strategies have been recognized to remove conclusion highlights or viewpoints during the time spent assessment mining in "Model-Based Opinion Mining and Sentiment Analysis" [12]. Spammer's conduct highlights the strategy proposed to illuminate the issue of spam review discovery and execution of different ways to deal with spam review arrangement and recognition [13]. For recognizing spam and fake review and mean reviews with a call, irrelevant words, and criticize words, utilizing hypothesis investigation. Potential highlights are removed from the accumulation of reviews got. At that point, the guidelines for Decision Rule Classifier are built up to recognize irrelevant reviews utilizing the Decision Classifier method dependent on different conditions [14]. Research on Decision Tree Classification Algorithm in Data Mining [15] and a new approach for identifying manipulated online reviews using decision tree" [16] and detection of fake review and brand

ISSN- 2394-5125 VOL 7, ISSUE 15, 2020

spam using data mining technique [17]. Spam detection algorithm related to the paper also given opinion spam detection [18] and in 2017 Research on product review analysis and spam review detection [19] and in 2012 Implementation of decision tree algorithm to the analysis of performance [20] is given in the field of spam detection

3. Methodology

Today web-based purchasing became a milestone in the e-commerce sector, because of phony reviews the evaluations of the marked items are dropping. Concentrating on discovering phony reviews is the principal task. The Technique of Decision Tree Classifiers is utilized in this work wrong feedbacks have been removed and gathered to distinguish the spam reviews utilizing six distinct conditions, to be specific star evaluations, reaction, reaction, customer profile, profile status, format conditions. Observing guidelines are characterized to recognize whether feedback of the item is fake or not.

- **Expert response to the buyer:** It involves confirming if the buyer has gotten any reaction for his report. Generally, when a buyer composes a phony survey of a specific item, individuals in the organization will react rapidly to the feedback.
- **Useful profile of user:** This is to test whether the profile of the buyer is helpful on the off chance that the analyst's profile is certified. If the buyer offers the incorrect announcement, then the profile would not care for it, and the profile is not trustable.
- **Template:** This is to test whether the client is utilizing the standard template. On the off chance that there is a model in the reviews, at that point, the review would be fake. On the off chance that inside one passage, the observer kept in touch with some negative or positive reviews without referencing the explanation, it is fake.
- **Stars rating of buyer less than 2(<2):** This condition specifies that if the buyer gives less than (<2) star rating in any product and do not mention the reason, so there is a chance that buyer's review is non-genuine.
- **Buyer replies to company's response:** This condition is to check and verify the reviewer's response to the company. If the buyer has responded to the company's response to the buyer then we can say that buyer is certified and gave genuine feedback.
- **User/Buyer profile contains proper and real information:** This condition is to check the buyer profile's information. The buyer profile contains full data about him/her—expert demonstrating that the buyer is certifiably not a suspected. If the profile does not contain any useful information, so it is known to be a fake profile.

In this procedure, we set up a dataset by gathering the surveys of the product name Samsung LED TV from the e-commerce site like Amazon. From the site, 200 online reviews are gathered for the experiment. By using the classifier, we observed that, from 200 reviews, 80 reviews had been distinguished as a phony or fake.

3.1. Decision Tree and Information Gain algorithm

Input: Online reviews of products from an E-commerce site.

Step 1. Initiate classifier,
Step 2. First, read the reviews of any e-commerce site product and check conditions in the manner that user's response to company, user maintain a convenient profile, any template in feedback, user's stars, rating to product and company's reply in buyer feedback,
Step 3. If the occurrence that the reaction of the organization is given and there is no answer for a reaction then it is a fake, otherwise not a fake review,
Step 4. In the feedback that the buyer profile is not valuable and not maintains appropriate information, at that point it is a fake,
Step 5. Check if, in the review, buyer utilizes the short format of words or template, so at that point, it is a fake,
Step 6. If the buyer gives less than two stars without mentioning the reason so it will be a fake review,
Step 7. Result (Fake or No Fake),

ISSN- 2394-5125 VOL 7, ISSUE 15, 2020

So, Decision Tree and Information Gain based on informative based reviews.

The below figure 1 shows a pictorial form of the working principle of algorithm and flow of the proposed model.

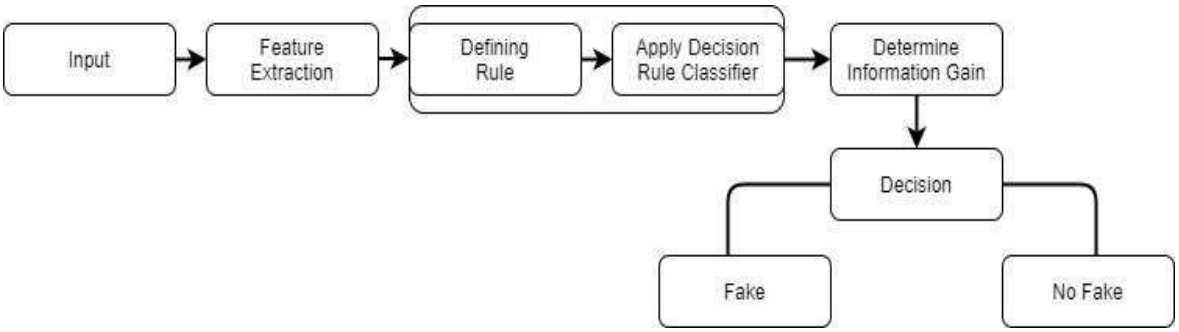

Figure 1 Flow of proposed model

Our algorithm is extracting and splitting the class label until it finds the best split for the dataset. After applying the rule, it calculates the information gain of each split and our data divide in the best splits. Table 2 is showing a sample of reviews, and table 3 is showing the result and accuracy of the algorithm.

Table 2. Sample for Genuine & Non-Genuine Reviews of the Product

S. No.	Details	Review 1	Review 2
1	Reviews based on Product	Samsung LED TV	Samsung LED TV
2	Reviews have given for the product	Not a worthy product in this budget	Poor quality waste of money
3	Star rating is given to the product	1	1
4	Service expert Response	No	Yes
5	Buyer's reply to expert	No	No
6	Buyer profile Usefulness	1%	1%
7	Buyer's profile Status	Less informative	Less informative
8	Using templates	Yes	Yes

Table 3: Experimental result of reviews

Reviews are gathered	200
Product has taken	LED TV
E-commerce site	Amazon
Reviews identified as fake by the method	80
The review identified as fake Manually	86
Favorable outcome of the method	94%

3.2 Spammer's Behavior Features

a. **Certified User:** The users on an e-commerce site are partitioned into a few positions based on their complete utilization. We have gotten the user rank data on a couple of well-known web-based business stages through the overview.

ISSN- 2394-5125 VOL 7, ISSUE 15, 2020

b. Divergence ratio: Sensible surveys are reliable with the product's quality and do not veer off from all reviews. As indicated by this element, we can pass judgment on whether a review is a spam or fake. Admittedly, we do not make it impossible for a case where the customer has purchased a low-quality item. The web-based business site will enable the client to discount his cash or replace the issue item with another item. In this way, the clients will even now give a generally sensible review.

c. Unfairness Rate: Various reviews of a similar user on an item may not be predictable. The client's first feedback of an item may not speak to the item's genuine experience, and the subsequent review regularly mirrors the user's experience of the item. If a user has composed at least 3 review for a similar item, those reviews were probably going to contain the unpleasant tendency to the item.

d. Feedback Duplication Rate: A few buyers regularly duplicate other buyer reviews, utilizing them as their very own feedback without or with a couple of slight changes. These copied feedbacks originate from a similar item or comparative items.

e. Feedback applicability Rate: The feedback some of the time has nothing to do with the item itself, for example, a notice, or a connection, or a pre-arranged unimportant substance. To recognize this sort of review, for example, a survey. We have to examine the connection between the reviews and the subject of the item. The review significance rate suggests the relevant between the feedback content and the item's subject.

f. Feedback information length: The review's length is additionally a significant pointer to distinguish spam reviews. At the point when the feedback substance is excessively short, we figure the commentator did not consider the item's experience truly. In this way, this sort of review does not predict well for information examination.

g. Feedback demonstration: A large portion of the online business sites presently gives a component of transferring the item's photos underutilize. To extra review time, by and large, the spammer doesn't offer an image of the item as the reference section of the reviews. Accordingly, we can utilize this element to identify spam reviews.

h. Cursed word using in feedback: Bad survey alludes to the situation wherein a customer has made a high volume out of feedback for different things inside short of time. Table 4 is showing expressions of fake or phony reviews,

Table 4. Word Expression of Phony Reviews

S. No	Reviews
1	Great mobile
2	Excellent mobile
3	Excellent product on amazon.
4	I could not resist using mobile

ISSN- 2394-5125 VOL 7, ISSUE 15, 2020

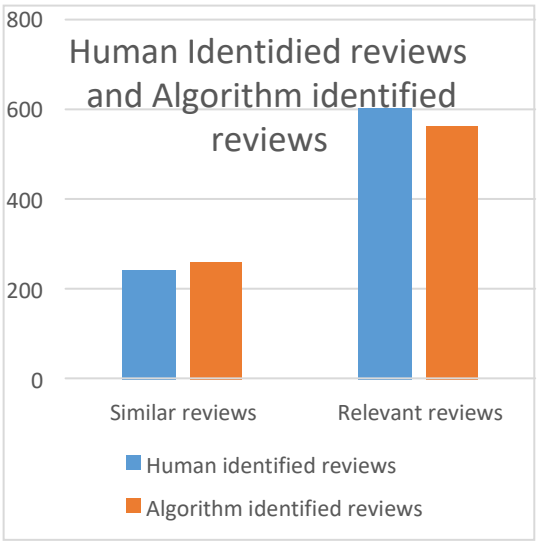

Figure 2. Compare: Human Identified Decision Tree Algorithm identified

Table 5. Isolation Forest Algorithm for Some Non-Genuine Reviews

S. No	Review
1	I am pleased to have this product.
2	Nice product in this budget
3	Excellent feature
4	Low cost to people
5	It came with the Arabian keyboard
6	Awesome

A few mobile reviews from the Amazon site have been chosen. Reviews apply to the subject for the cell phone shopper, and review is of the most significant importance relative to other reviews. But the connection between analysis and the subject is challenging to decide. Also, some feedback out of total reviews was relevant to the topic of the mobile phone service. Non-genuine feedback can build the expense of settling on choices as well as influence the exactness of deciding. Human and Algorithm analysis on dataset and algorithm is extracting more reviews from the dataset. The procedure should be improved to discover increasingly applicable feedbacks. Generally speaking, our procedure discovers a new 288 fake review on 800 reviews. Figure 2 shows the relevancy of the algorithm, and Table 5, showing the type of comments.

3.3. Isolation forest algorithm

- The fundamental unit of information in the strategy is an item or product. Concentrate on the feedback examples of all items to get exceptional feedback.
- Collect the data and building the temporal data.
- Isolation forest model divides feedbacks into the time slots and then calculates the score concerning the time slot.
- According to the score, do a probability of phony reviews.

Isolation forest algorithm consists of the following steps: taking Amazon dataset as input

- **Collect data:** Collect dataset from various e-commerce sites.
- **Building a temporal feature:** Gather the reviews from the dataset.
- **Isolation forest to Model data:** Apply the model to divide reviews based on their time slot periods.
- **Calculate outlier score:** Measure the method's output
- **Output outlier sample score:** Showing relevant results based on the reviews time slot.

ISSN- 2394-5125 VOL 7, ISSUE 15, 2020

Introduce a technique to recognize phony reviews dependent on feedback records related to the product. We initially examine the qualities of review information utilizing an extracted Amazon dataset that identifies that the examples of review records for products are comparative in ordinary circumstances. Concentrate the feedback of the product to a short component vector and apply an isolation strategy to recognize spam reviews by concentrating on the contrasts between the examples of product reviews to distinguish anomaly reviews—this sophisticated spam discovery strategies utilizing the Amazon dataset.

4. Proposed Work

4.1 Adding New Attributes in Decision Tree and Information Gain

This method is used to detecting the phony reviews by adding the two more attributes verified user and helpful votes along with the response, reply, template, star rating, profile useful to the algorithm to find the phony reviews. These two attributes are much more helpful in classifying the dataset with higher accuracy.

4.2 Integrate Decision Tree with Random forest

In the proposed work, the Random Forest is exclusively new implementation to find out the phonier reviews accurately by making the decision trees of the same dataset by which we can predict the dataset more strongly. And form a different tree with the help of Bootstrapping of my training data set by randomness, it will select the random example from the training dataset and create a bootstrap dataset**.** We simply consider the random subspace method. So, we have a random subspace with the different datasets by which we can create several decision trees. Then we used those decision trees to classify our training dataset the random forest consists of trees with different results by which we can make the results, and the conditions are if trees are 4:

- If all four trees are showing the same results (true or false) so we will take it as the final decision.
- If three trees are showing the same result (true or false) and one tree is showing different so we will take a higher number of appearances of decision.
- If the two trees are showing accurate results and the other two trees are showing false, so we take the false as the final result.

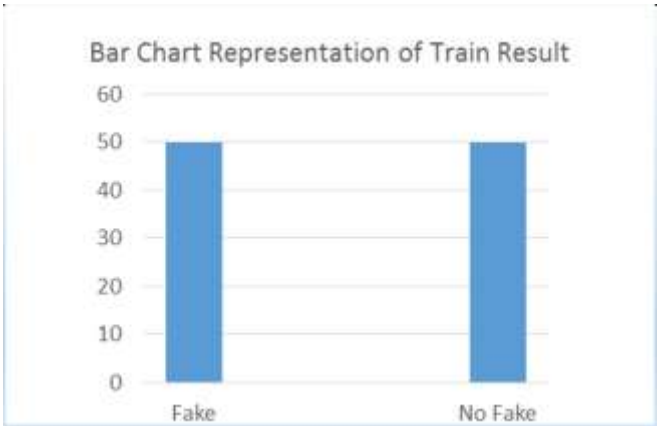

Figure 3. Training result of fake and no fake

5. Results and Discussion

5.1 Experiment Data and Evaluation

We collected the dataset from Kaggle and determined the features of the dataset along with the label. The dataset contains all the feature's values which we will use in classification. The class label contains two classes one is 'Fake' denoted as 'F', and the other one is 'No Fake' denoted as 'NF'. We used the Decision Tree algorithm to create a tree then we create four different trees of same the dataset with the help of Random Forest. Now the Random Forest algorithm is implemented on the same dataset to create the multiple decision trees, and this is done by the bootstrapping of the training dataset. In this decision trees, it is classified our dataset with more knowledge and information, so we can say that the proposed method can do better than the previous methodologies used in spam detection or fake reviews detection by which it achieves the best accuracy. So, the classification along with tree representation of decision tree and random forest are shown. The classification of all decision trees and predictions of those trees are mainly shown. In figure 3 it shows the train data with fake or no fake reviews:

ISSN- 2394-5125 VOL 7, ISSUE 15, 2020

Table 6 is showing the random forest algorithm's result. The final result or classification of our test dataset by random forest, and we classify the dataset with higher accuracy is 93.33% and form a reliable prediction.

Table 6. Results of Random Forest Algorithm

S. No	Tree 0	Tree 1	Tree 2	Tree 3
24	F	F	F	F
26	F	F	F	F
2	NF	NF	NF	NF
16	F	F	NF	F
32	F	F	F	F
31	NF	NF	F	NF
25	F	NF	F	F
19	F	F	NF	NF

5.2 Experiment Result

The result of our algorithm is much more accurate than previous algorithms, and it provides more information and determines the best splits based on the class label. First, it finds potential split then best splits among them now it calculates information gain for all potential splits afterward, take only the best splits. The item called "Samsung LED TV" from the e-commerce site "www.amazon.com," online reviews are gathered for experimentation. We selected 200 reviews in which 80 reviews have been identified as phony reviews. The choice of the examination approach is approved by manually reviews inspection, in which 86 reviews are distinguished as fake. As of now, the achievement footstep 93.33% has been accomplished. In the dataset, we split the dataset into a training dataset, and second is the test dataset and bootstrap the training dataset with shuffling the attributes and create the decision of test dataset. After the creation of the decision tree, it got 80% accuracy. With random forest classification, we got 93.33%

Table 7. Train Dataset with Decision Tree Algorithm

S. N0.	Response	Reply	Profile Useful	Thickness	Template	Star Rating	Verified	Helpful Votes	Label
0	Y	Y	>=20%	Thick	N	>=2	False	>=50	NF
1	Y	Y	>=20%	Thick	N	<2	True	>=50	NF
2	Y	Y	>=20%	Thick	Y	>=2	False	>=50	NF
3	Y	Y	>=20%	Thick	Y	<2	True	<50	NF
4	Y	N	>=20%	Thick	N	>=2	True	<50	NF

ISSN- 2394-5125 VOL 7, ISSUE 15, 2020

Table 8. Classification of Fake Vs. No Fake

S. No	Respons E	Reply	Profile Useful	Thick ness	Templat e	Star Ratin	Verified	Helpful Votes	Label	Classificati on	Classification Correct
24	Y	N	<20%	Thin	N	<2	False	>=50	F	F	True
26	N	N	<20%	Thin	Y	<2	True	>=50	F	F	True
2	Y	N	>=20%	Thick	Y	>=2	False	>=50	NF	NF	True
16	N	N	<20%	Thin	N	<2	True	<50	F	F	True
32	Y	N	>=20%	Thin	Y	<2	False	<50	F	F	True
31	Y	Y	>=20%	Thin	N	>=2	True	>=50	NF	NF	True
25	Y	N	<20%	Thin	N	>=2	True	>=50	F	F	True
19	Y	Y	<20%	Thin	n	<2	False	<50	NF	NF	True
30	Y	Y	<20%	Thin	Y	>=2	True	>50	NF	NF	True
11	N	N	>=20%	Thick	Y	<2	True	<50	NF	NF	True
18	Y	Y	<20%	Thin	N	>=2	False	>=50	NF	NF	True
28	N	N	>=20%	Thin	N	>=2	False	<50	F	F	True
6	Y	N	>=20%	Thick	Y	>=2	False	<50	NF	NF	True
36	Y	N	<20%	Thick	N	<2	False	>=50	F	N	False

6. Conclusion

The recognizable proof of spam reviews from e-commerce site feedback has been characterized. Henceforth to recognize whether the feedback is genuine or not, but the integrated random forest creates 4 more decisions tree through bootstrapping with shuffling the accuracy of the test dataset. So, it is clear that our algorithm classifies any type of dataset with more proficiency. Algorithms also found non-similar reviews, related reviews; the algorithm is using a dataset from Amazon.com of LED TV reviews. The experimental results indicate that our algorithms are more effective than the conventional algorithm in the detection of phony reviews.

Table 7 shows the train dataset result, and Table 8 is showing the final output. Attributes and algorithms made it more suitable—the choice of the classifier by which the data increase to recognize the most noteworthy reviews. Our decision tree classifier is applied for different potential highlights, for example, the response of the company, user's profile information, template using by the user, star rating, reply to company's response by the user and verified user by the company and helpful votes given to the user or reviewer by other people. Here the methodology analyzes reviews taken from the Amazon site of LED TV. The proficiency of our methodology has made a 93.33% accuracy of the test dataset. Based on profile information, how we can assume it is fake. If a profile is not containing proper information, but the user is verified, and the rating is real. In this, the star rating is less than 2 is not considered as a strong point to determine fake reviews, and mostly we do not have time to give furthermore reply to company's response. Based on the review template, we cannot justify that it is fake, so they are the future works. Hence, it builds marking the proof markers for the identification of wrong feedbacks on the online shopping sites of products on the internet. Counterfeit reviewer has some essential conduct highlights, to execute the proposed discerning proof pointers, present procedure to perceive both comparable and significant reviews, individually. The test perception shows that our decision tree and random forest have higher effectiveness in recognizing non-genuine feedback.

7. References

[1] Rakibul Hassan, Md. Rabiul Islam "Detection of fake online reviews using Supervised and semi-supervised learning" 2019 International Conference on Electrical, Computer and Communication Engineering ISSN 1856-6224

[2] Li Jing "Online fake comments detecting model based on feature analysis" 2018 International conference on smart grid and electrical automation ISSN 1816-8542

[3] Wenqian Liu, Jingsha He,Song Han, Fangbo Cai,Zhenning Yang, Nafei Zhu "A Method for the Detection of fake reviews Based on Temporal Featuresof Reviews and Comments" 2019 IEEE Engineering Management Review ISSN 1937-4178

[4] Ning Luo, Huaxun Deng, Linfeng Zhao, Yuan Liu∗, Xingwei Wang, Zhenhua Tan "Multi aspect feature based neural network model in detecting fake reviews" 2017 4th International Conference on Information Science and Control Engineering ISSN 17376084

[5] Nidhi A. Patel, Prof. Rakesh Patel "A Survey on Fake Review Detection using Machine Learning Techniques" 2018 4th International Conference on Computing Communication and Automation (ICCCA)

[6] Anusha Prabakaran, Min Chen "Product Review Credibility Analysis" 2019 workshop on computing networking and communication ISSN 18600759

[7] SP.Rajamohana, Dr.K.Umamaheswari, M.Dharani, R.Vedackshya "A SURVEY ON ONLINE REVIEW SPAM DETECTION TECHNIQUES" International Conference on Innovations in Green Energy and Healthcare Technologies ISSN 17333507

[8] Anna V. Sandifer, Casey Wilson, Aspen Olmsted "Detection of fake online hotel reviews" The 12th International Conference for Internet Technology and Secured Transactions

[9] Chih-Chien Wang, Min-Yuh Day, Yu-Ruei Lin "A Real Case Analytics on Social Network of Opinion Spammers" 2016 IEEE 17th International Conference on Information Reuse and Integration

[10] K.Manoj, T.S.Sandeep, Dr. N Sudhakar Reddy, P.M.D.Alikhan **"**Genuine ratings for mobile apps with the support of authenticated users' reviews"

[11] A. Angelpreethi, Dr. S. Britto Ramesh Kumar "An Enhanced Architecture For Feature Based Opinion Mining From Product Reviews" World Congress on Computing and Communication Technologies ISSN 17280104

[12] Vamshi Krishna. B, Dr. Ajeet Kumar Pandey, Dr. Siva Kumar A. P "Topic Model Based Opinion Mining and Sentiment Analysis "2018 International Conference on Computer Communication and Informatics (*ICCCI* -2018), Jan. 04 – 06, 2018, Coimbatore, INDIA

[13] Pan Liu, Zhenning Xu,Jun Ai, Fei Wang "Identifying Indicators of fake reviews Based on Spammer's Behaviour Features" 2017 IEEE International Conference on Software Quality, Reliability and Security Companion, ISSN 1709-8576

[14] Sanjay K.S, Dr. Ajit Danti "Detection of fake opinions on online products using Decision Tree and Information Gain" 2019 3rd International Conference on Computing Methodologies and Communication, ISSN 1895-8282

[15] Qing-yun dai,Chun-ping Zhang and Hao wu "Research of Decision Tree Classification Algorithm in Data Mining" International Journal of Database Theory and Application, Vol.9, No.5 (2016), pp.1-8, ISSN: 2005-4270 IJDTA.

[16] Rajashree S. Jadhav and Deipali V. Gore "A New Approach for Identifying Manipulated Online Reviews using Decision Tree" International Journal of Computer Science and Information Technologies, Vol. 5 (2), 2014, 1447-1450, ISSN: 0975-9646.

[17] Rashmi Gomatesh Adike and Vivekanand Reddy "Detection of Fake Review and Brand Spam Using Data Mining Technique" International Journal of Recent Trends in Engineering and Research, Volume 02, Issue 07; July - 2016, ISSN: 2455-1457.

ISSN- 2394-5125 VOL 7, ISSUE 15, 2020

[18] Salma Farooq and Hilal Ahmad Khanday "Opinion Spam Detection: A Review" International Journal of Engineering Research and Development, Volume 12, Issue 4, e-ISSN: 2278- 067X, p-ISSN: 2278-800X, April 2016.

[19] Shashank Kumar "Research on Product Review Analysis and Spam Review Detection" Research gate- Conference Paper, February 2017, 317932754.

[20] Pooja Sharma and Rupali Bhartiya"Implementation of decision tree Algorithm to analysis of performance" International Journal of Advanced Research in Computer and Communication Engineering", Vol. 1, Issue 10, December 2012, ISSN : 2278-1021.