Comment Sentiment Analysis and Fake Product Review Detection

Jyoti Bist¹, Neha Hulsurkar², Shraddha Bhalerao³, Deepali Narkhede⁴

^{1,2,3}Student, Information Technology, Shivajirao S. Jondhale College of Engineering, Dombivli, Mumbai, India ⁴Assistant Professor, Information Technology, Shivajirao S. Jondhale college of Engineering, Dombivli, Mumbai, India

muna

Abstract—Opinion mining has become very important with the increase of E-commerce websites and gives clear bifurcation to customers about the product reviews and service reviews. Sentiment analysis helps computer to extract emotions from customer's reviews further helping users in decision making process while shopping. Naive Bayes, a machine learning algorithm will be used for sentiment analysis and fake review detection along with other methods. In this paper we propose a system which improves users shopping experience by recognizing emotions behind the reviews and detecting fake or false reviews posted by opponent with wrong intentions

Keywords-naïve bayes, opinion mining, sentiment analysis

1. INTRODUCTION

Now-a-days, due to the advent of technology and internet, shopping is mostly based on reviews or feedbacks. Traditionally when e-commerce websites were not considered. Buying product then using it will reveal its quality. E-commerce have played a vital role in changing this shopping culture. Today everything can be purchased with just a phone and network. As the use of e-commerce websites is increasing, it is more susceptible to fault intensions. Fake product reviews can lead to massive growth or great financial losses. We propose a project which focuses on removing fake reviews and analyzing users review for great shopping experience.

2. EXISTING SYSTEM

Several methods have been suggested to understand and implement opinion exploration and sentiment analysis. Scientists have developed models to identify the polarity of words, sentences, and the entire document. There are now several tools available to explore opinions, analyze feelings, and synthesize opinions. Previously, the basic concept among all the algorithms and models has been identifying emotional words first. These words are used to find where opinion is present inside the document. Then the opinions extracted are analyzed to find out polarity of the opinion.

2.1 Sentiment analysis of product reviews for ecommerce recommendation

This article describes the methods for performing magazine analysis. Here semi-supervised method suggested in which approach is on important opinion words finds using Word Net. This method is beneficial comparing with properly supervised or unsupervised algorithms with the help of advanced learning method as ANN to increase accuracy. Sentiment analysis is used at the sentence level with NLTK with the probability model Naïve Bayes. The representation of the results is done graphically and statistically [1].

2.2 Fake product review monitoring using opinion mining

Product reviews play an important role in deciding to sell a particular product on "e-commerce websites" or apps like "Flipkart", "Amazon", "Snapdeal", etc. In sentiment analysis, the goal is to figure out the opinion of a customer through a piece of text. First it is checked if the review is related to the specific product with the help of Decision tree. Spam dictionary is used to identify the spam words in the reviews. In Text Mining several algorithms are applied and on the basis of these algorithms specific results are obtained[2].

2.3 Fake review detection using opinion mining

As e-commerce grows and becomes more and more popular day by day, the number of comments received from customers about any product increases rapidly. Nowadays, people rely heavily on reviews before buying anything. This leads many people to write unnecessary scams and reviews about other related products or services. Some organizations in the marketplace even hire professionals to write false reviews and promote their products or defame the products of their competitors. Therefore, this article aims to develop a method that detects and records false reviews. The proposed method automatically classifies users' opinions into "suspicious", "clear" and "fuzzy" categories by phase processing. The fuzzy category recursively reveals suspicious or clear elements. This results in richer detection and can be useful to both the business organization and customers. The sales

organization can monitor the sale of its products by analyzing and understanding what customers say about the products. This can help customers buy valuable products and spend their money on quality products. Finally, end users view each individual exam with scored polarity scores and a credibility score [3].

2.4 Amazon review classification and sentiment analysis

Biggest online shopping mall in the world is "Amazon". Before buying the item people always belief items by looking at product reviews. But reviews on "Amazon" are not necessarily items, but a mix of item review and item service review ("Amazon"-linked or item-company-linked). Customer is mistaken as a general feeling (grading ranking) that "Amazon" provides feedback as a collective opinion and that there is no difference between a service review and a item review. The proposed model makes a satisfactory distinction between service review and product review, in addition to that, if the user talks about particular product feature it also classifies the review as a feature review. A featured review is nothing more than a product review, this template also gives an impression of the text about the functionality of the product. For example, if the user writes in their review, "The storage for this phone is excellent," then we also rate the storage function as positive. This was to construct a system that

4. ARCHITECTURE

envision the feel of the magazine in the form of graphics [4].

3. PROBLEM DEFINITION

The main objective of this software is to help users gather the best details about a particular item they want, based on the opinions of other consumer, and to help make a choice about any product. This system detects the hidden sentiments like good or bad reviews in comments and rate them accordingly and also it detects the fake reviews of product. The system uses opinion mining methodology in order to achieve desire functionality. This system is beneficial to the user which provides the analysis sentiment and fake product review.



Fig-1 System Architecture of Comment Sentiment Analysis and Fake Product Review Using Opinion Mining

IRIET VOLUME: 07 ISSUE: 05 | MAY 2020

WWW.IRJET.NET

System works as follows:-

1. User registers using email id.

2. User logins to the system

3. User once access the system, can see product and can post review about theproduct.

4. Overall rating will be calculated by the system for a particular product.

5. Fake reviews are detected by the system and will be removed by the admin

5. METHODOLOGY

5.1 Comment Sentiment Analysis

• Naïve Bayes Algorithm

This algorithm is also known as Bayesian theorem. It is mostly used when the dimensionality of inputs are high.



Consider example displayed in the illustration above. Let the red color indicate apples and green color indicates grapes. Our aim is to classify the new cases on arrival, that is, to decide which class label they belong to, according to the objects that currently come out.

Given that there are twice as many grapes as apples, it is obvious to trust that a new case (which has not yet been observed) is twice as likely to have homemade grapes instead of apples. In Bayesian analysis, this belief is called an earlier probability. The above probabilities are based on previous experience.

Data and a billion for		Number of Grapes
Grapes	00	Total number of fruits
Prior probability for	α –	Number of Apples
Apples		Total number of fruits

Since there is a total of 60 objects, 40 of which Are Grapes and 20 Apples, our prior probabilities for class membership are:

After formulation of prior probability, we are now ready to classify a new object (WHITE circle).

Since more likely that the new cases belong to that particular fruit.



From the illustration, it is clear that the likelihood of X given grapes is smaller than apples as the circle encompasses 3 apples and 1 grapes.

Probability of X given	$\alpha \frac{1}{40}$
grapes	40
Probability of X given	$\frac{3}{20}$
apples	20

In "Bayesian analysis" the concluding classification is made by combining the two sources of data, i.e. the previous one and the probability, to form a posterior probability using the so-called "Bayes"rule.

```
Posterior probability of X being grapes _{\infty}

Prior probability of grapes x likelihood of X given grapes

=\frac{4}{6} \times \frac{1}{40} = \frac{1}{60}

Posterior probability of X being apples ^{\infty}

Prior probability of apples x likelihood of X given apples

=\frac{2}{6} \times \frac{3}{20} = \frac{1}{20}
```

Lastly, we classify X as apple since its class membership reaches the largest posterior probability.

Note. The above probabilities are not standardized. However, this does not change the result of the classification since their normalization constants are the same. IRIET VOLUME: 07 ISSUE: 05 | MAY 2020 WWW.II

5.2. Fake product review detection

• IP address method

In this method, IP address of the user is tracked. If same IP address is used for posting multiple reviews then the Reviews are considered Spam.

• Timestamp Method

Timestamp is a digital proof of the time of occurrence of a particular event. The system detects if the timestamp of the reviews are same.

• Is product purchased

System check if the user posting the review about the particular product has purchased the product or not.

• User reporting

System keeps record of reviews which are reported by other users. Admin can delete the reported comments.

6. RESULT

🗄 Super Market an	Econr X +				-	.	6	X
0 @) localhost 8080	(FakeproductReview(Admin)Commer	tanalysisjep	0 Å	ţ	L	Ê	
Product ID.	USER Name	Email Address	Statement of user	Get Review Time	Status			
1	tijo	jyoti.07@gmail.com	Amazing camera	2019-10-14 15:06:42.0	5			
1	tijo	jyoti.07@gmail.com	Amazing camera	2019-10-14 15:06:42.0	5			
7	tijo	jyoti.07@gmail.com	Amazing camera	2019-10-14 15:06:42.0	5			
8	Neha	neha.hulsurkar@gmail.com	Its an amazing phone.	2019-10-16 18:59:20.0	5			
1	tijo	jyoti.07@gmail.com	Amazing camera	2019-10-14 15:06:42.0	5			
8	Neha	neha.hulsurkar@gmail.com	Its an amazing phone.	2019-10-16 18:59:20.0	5			
1	tijo	jyoti.07@gmail.com	Amazing camera	2019-10-14 15:06:42.0	5			
8	Neha	neha.hulsurkar@gmail.com	Its an amazing phone.	2019-10-16 18:59:20.0	5			
9	Neha	neha.hulsurkar@gmail.com	Its liked this product.	2019-10-16 19:39:57.0	5			
1	tijo	jyoti.07@gmail.com	Amazing camera	2019-10-14 15:06:42.0	5			
8	Neha	neha.hulsurkar@gmail.com	Its an amazing phone.	2019-10-16 18:59:20.0	5			
9	Neha	neha.hulsurkar@gmail.com	Its liked this product.	2019-10-16 19:39:57.0	5			
1	tijo	jyoti.07@gmail.com	Amazing camera	2019-10-14 15:06:42.0	5			
8	Neha	neha.hulsurkar@gmail.com	Its an amazing phone.	2019-10-16 18:59:20.0	5			
9	Neha	neha.hulsurkar@gmail.com	Its liked this product.	2019-10-16 19:39:57.0	5			
7	alla.	kuti 179anail aan	Intrins company	2010 10 11 15 06 10 0	ç		t	

Fig-2 Comment Sentiment Analysis Result



Fig.-3 User Report Method

👌 Super Narket an Ecommerce Cri 🗴	+	- 0 X
€ → C O locahost8880F	akeprodustkeinen/Khinn TimeStampRevien/Fiteringjep	x 🖲 🖯 :
\Xi Apps M Gmail 🟮 YouTube () Schare Engineeri.	

DATE AND TIME FAKE REVIEW TABLE

1 1pt	PTCOUCE ID.	USER Name	Email Address	Statement of user	Get Review Time	Status
tip jpt. Digral.com ido ot lie fits poduci; jam ot satified 2019-10-34 1511-6.80 Toe shari shari shari shari topod topod 2019-10-34 15248.00 Toe n mka seka.Mourkarsgmal.com in not satified with tip poduci at the lot of bag. 2010-09 10-34 15248.00 Toe n mka seka.Mourkarsgmal.com tot tog 2010-09 10-34 152.00 Toe s mka seka.Mourkarsgmal.com tot tog 2000-09 142150.00 Toe s mka seka.Mourkarsgmal.com tot tog 2000-09 142150.00 Toe s mka seka.Mourkarsgmal.com tot tog 2000-09 142150.00 Toe s mka seka.Mourkarsgmal.com Koz Pore 2000-09 142150.00 Toe s mka seka.Mourkarsgmal.com Koz Pore 2000-09 1495151.00 Toe s mka seka.Mourkarsgmal.com Koz Pore 2000-09 1495151.00 Toe	1	tijo	jyoti.07@gnail.com	Amazing camera	2019-10-14 15:06:42.0	True
shari shorell Higgmail.com to good 2019-10-26-10-26-06-00 Twee 1 nela nela.kludwarsgmail.com in not attrifiel with thi product at this to dribug. 2000-109-11-84-70.0 Twee 9 nela nela.kludwarsgmail.com texting! 2000-109-11-84-70.0 Twee 5 nela eelasjmail.com texting! 2000-109-11-12-10.0 Twee 2 tip tytic biologingmail.com Nice Rivere 2000-109-10-156-16.0 Twee 2 nela sela.kludwarsgmail.com Nice Rivere 2000-109-10-156-16.0 Twee	1	tijo	jyoti.07@gnail.com	i do not like this product, i am not satisfied	2019-10-14 15:11:46.0	True
neta neta.bluskratgration in not table with the product at the lab of dags. 20040-09 1246170 The 9 neta neta.bluskratgration setingt 20040-09 1246170 The 5 neta neta.bluskratgration textingt 20010-09 1246170 The 2 tip jot ibstilligeral.com Kite Anore 200010-09 1266100 The 2 neta neta.bluskratgration Kite Anore 200010-19 1266100 The 2 neta neta.bluskratgration Kite Anore 200010-19 1266100 The	1	shari	slove1119@gmail.com	its good	2019-10-26 10:39:09.0	True
9 nela nela.bludrategradiom tettngl 00040341421:0.00 The 5 nela delaignadiom tettngl 00040341421:0.00 The 2 tip jott bio010gnadiom Nice Ricce 00040341051051:0.00 The 2 nela ela-bludrategradiom Nice Ricce 00040341051053:5.00 The	1	neha	neha.huburkar@gmail.com	im not satisfied with this product as it has lots of bugs.	2020-03-09 13:46:17.0	True
5 nela nelasignation toting 20040741421x401 Tue 2 tip juti bit016gradi.com Kice Alore 200407411621x401 Tue 2 nela nela.bitudire/gradi.com Kice Alore 200407411695540 Fale	9	neha	neha.huburkar@gmail.com	testing1	2020-03-09 14:21:32.0	True
2 tip jyth bioD10gmal.com Nice Rione 200140-51 0556:58.0 Time 2 mela mela.blubulerogmal.com Nice Rione 200140-51 0556:58.0 Fale	5	neha	neha@gmail.com	testing	2020-03-09 14:26:40.0	True
2 nela nela.bubuta@gmal.com Kice.Piore 2004/01/3105454.0 Fale	1	tijo	jyoti.bist07@gnail.com	Nice Phone	2020-03-13 10:56:18.0	True
	2	neha	neha.huburkar@gmail.com	Nice Phone	2020-03-13 10:56:54.0	Fale

Fig-4 Fake Product Review Detection Using Timestamp Method



🖇 Supe Holder an Accornered Dr. X 🕴	- 0 X
← → C © localkost000 feleproduct/leview/kdmin ProductpurchaseReview.jpp	‡ ₫ θ :
🧮 kaps M Grail 🔹 toulute 🕐 Schweizigneeil.	

IS PRODUCT PURCHASED REVIEW TABLE

Sr. No.	Product ID	USER Name	Email Address	Statement of user	Get Review Time	Status
1	1	tījo	jycti.07@gnail.com	Amazing camera	2019-10-14 15:06:42.0	Fake
2	1	tījo	jyoti.07@gmail.com	i do not like this product, i am not satisfied	2019-10-14 15:11:46.0	Fake
3	1	stari	slove1119@gmail.com	its good	2019-10-26 10:39:09.0	Fake
4	1	neha	neha.hulsurkar@gmail.com	in not satisfied with this product as it has lots of bugs.	2020-03-09 13:46:17.0	True
5	9	neha	neha.hulsurkar@gmail.com	testing!	2020-03-09 14:21:32.0	Fake
6	5	neha	neha@gmail.com	testing	2020-03-09 14:26:40.0	Fake
1	1	tijo	jyoti.bist07@gmail.com	Nice Phone	2020-03-13 10:56:18.0	Fake
8	2	neha	neha.hulsurkar@gmail.com	Nice Phone	2020-03-13 10:56:54.0	Fake

ł	, ^p Type here to search	0	×.	Ê	1	L	0	đ	-	Ç	9	- A	ÔD @¢ ^{BNG} N	641 PM 3/19/2020	Ę

Fig.-5.Fake Product Review Detection Using Product Purchased Method

	X 10 Stretter	•					-	٥	
C O b	cahost:081/Fikep	nodu:tReview/Adm	nin/AlgorthmFilteringisp				4 0	θ	
M Grai	🛛 Yau'lube 🍈 S	othane Engineeri.							
		N	AIVE BAYES	ALGORTITHM REVIEW 1	TABLE				
	Product ID.	USER Name	Email Address	Statement of user	Get Review Time	Status			
	Product ID. 7	USER Name tijo	Enall Address jysti.078gmail.com	Statement of user Amazing camera	Get Review Time 2019-10-14 15:06:42.0	Status 5			
	Product ID. 7 1	USER Name tijo tijo	Email Address jysti.078gmail.com jysti.078gmail.com	Statement of user Amazing camera i do not like this product, i am not satisfied	Get Review Time 2019-10-14 15:06:42.0 2019-10-14 15:11:46.0	Status 5 2			
	Product ID. 7 1	USER Name tijo tijo shari	Enall Address jysti. J79gmail.com jysti. J79gmail.com dove11198gmail.com	Statement of user Anazing camera 1 do not like this product,) am not satisfied 1/s good	Get Raview Time 2019-10-14 15:06:42.0 2019-10-14 15:11:46.0 2019-10-14 15:11:46.0 2019-10-26 10:39:09.0	Status 5 2 3			
	Product ID. 7 1 1 1	USER Name tijo tijo shari neha	Enall Address jysti. U7@gmail.com jysti. U7@gmail.com dove1119@gmail.com neha.hulsurkan@gmail.com	Statement of star Arracing cames I do not like this product i am not satisfied Its good Im not satisfied with this product as it has late of hugs.	Get Review Time 2019-10-14 15:06:42.0 2019-10-14 15:11:46.0 2019-10-16 10:39:09.0 2019-10-26 10:39:09.0 2020-03-09 13:46:17.0	Status 5 2 3 1			
	Product D. 7 1 1 1 9	USER Name tijo tijo shari neha	Enall Address jyrti J78gmail.com jyrti J78gmail.com done11198gmail.com nelta.huburkantgmail.com nelta.huburkantgmail.com	Redeneet of our Anazing cames ide not like this product, i are not catabilied the good in not catabilied with this product as it has late of hugs. Recting!	Get Raview Time 2019-10-14 15:06:40.0 2019-10-14 15:11:46.0 2019-10-14 15:11:46.0 2019-10-26 10:35:09.0 2020-03:09 13:46:17.0 2020-03:09 14:21:32.0	Status 5 2 3 1 5			
	Product ID. 7 1 1 1 9 5	USER Name tijo tijo shari neha neha	Enall Address jysti J78gmail.com jysti J78gmail.com daret1198gmail.com neha.huburkarigmail.com neha.huburkarigmail.com	Robenet of our Anazing cames ide not like this product; iam not catofiel this good in not catofiel with this product is it has lots of hags, secting tecting	Get Review Time 2019-10-14 15:06:42.0 2019-10-14 15:01:42.0 2019-10-26 10:39:09.0 2019-10-26 10:39:09.0 2020-03-09 13:46:17.0 2020-03-09 14:21:32.0 2020-03-09 14:21:32.0	Status 5 2 3 1 5 5			
	Product D. 7 1 1 1 9 5 2	USER Name tijo tijo shari neha neha tijo	Enall Address jorti D78gmail.com jorti D78gmail.com dene1198gmail.com neka.hdoxAnsgmail.com neka.hdoxAnsgmail.com neka.igmail.com jorti biot078gmail.com	Balaneet of our Anazing cames ide not like this product; an not catofied this good in not catofied with this product as it has list of bags, acting acting like Prove	Get Review Time 0019-10-14 (5:01-40.0 0019-10-14 (5:01-40.0 0019-10-16 (0:38:08.0 0019-10-26 (0:38:08.0 0019-10-26 (0:38:08.0 0019-10-26 (0:38:08.0 0019-10-26 (0:38:08.0 0019-10-26 (0:38:08.0 0019-10-26 (0:38:08.0 0010-43-09 (3:44:17.0 0010-43-09 (4:24:32.0 0010-43-09 (4:24:42.0 0010-43-09 (4:24:42.0 0010-43-13 (0:56:18.0	Status 5 2 3 1 5 5 5 5 5			

ł	ρ Type here to search	0	XXX	Ê	h	۵	0	đ	Î	¢	1	^ 0 ∎∉¢ ⁸	NG 62 N 311	8911 92120	- Harrison

Fig-6. Fake Product review detection using naïve Bayes algorithm

lęps M Grail 🚦 Ibu	iute 🍵 Sotivarei	Engineeri			- •						
IP BASED ANALYSIS											
Si	r. No.	USER Name	P	Total Review From same IP	Status						
3	0	neha	0:0:0:0:0:0:0:1	1	Not Fake						
2	9	tijo	0:0:0:0:0:0:0:1	1	Not Fake						
3		subhash	114.79.143.106	1	Not Fake						
5		shailesh	114.79.143.107	1	Not Fake						
2	6	neha	192.168.0.102	1	Not Fake						
2	4	neha	192.168.0.103	1	Not Fake						
2	3	shari	192.168.0.103	1	Not Fake						
2	1	tijo	192.168.0.103	1	Not Fake						
2	0	yogini123	192.168.0.103	1	Not Fake						
2	1	neha	192.168.0.105	1	Not Fake						
21	8	neha	192.168.0.108	1	Not Fake						
9		pia	192.168.0.11	1	Not Fake						
13	3	poa	192.168.0.11	1	Not Fake						
					· · · · ·						

Fig-7. Fake Product Review detection using IP based analysis

👌 Super Marke	tan Econmece Ori X	+		-	٥	χ
$\leftrightarrow \ {\tt C}$	0 localhost8080,63	aleproductReview/Admin/OveralReviewAnalysist.jpp	Å	ł	θ	ł
🚦 Apps - M G	imai 🖪 YouTube 🕻	Software Engineer L.				

OVERALL RESULT TABLE

Millional 🖷 conce. 🖸 provespiljestr

Product ID.	USER Name	Email Address	Statement of user	Algorithm Analysis	Purchased Analysis	Date Time Wise Analysis	ip Analysis	Overall	
7	tijo	jyoti.07@gmail.com	Amazing camera	40%	OK	10%	10%	60.0 Fake	
1	tijo	jyəti.07@gmail.com	i do not like this product, i am not satisfied	40%	0%	10%	105	60.0 Fake	
1	shari	slove1119@gmail.com	its good	40%	0%	10%	105	60.0 Fake	
1	neha	neha.hulsurkar@gmail.com	in not satisfied with this product as it has lots of bugs.	40%	OK	10%	105	60.0 Fake	
9	neha	neha.hulsurkar@gmail.com	testing1	40%	OK	10%	OX	50.0 True	
5	neha	neha@gmail.com	testing	40%	OK	10%	0%	50.0 True	
2	tijo	jyoti.bist07@gmail.com	Nice Phone	40%	C%	10%	0%	50.0 True	
2	neha	neha.hulsurkar@gmail.com	Nice Phone	40%	OK	05	0%	40.0 True	

Fig-8. Overall Result Table



1

7. CONCLUSIONS

Now a day's technology is growing day by day and there are so many website and application available in the online market by which they sell their product. Every product contains millions of reviews and on basis of these reviews user buy the product most of the time. There are some organizations which post fake reviews on genuine product for some reasons. Reasons are defamation of the organization by the opponent company also for promoting its own products organization often pay its employees to comment on its products. So our software will do analysis of these comments which display the overall rating of the product which will help customer for choosing the product and provide only genuine reviews by detecting the fake review.

8. REFERENCES

- Abhyankar, Bhavarthi, Gaidhar, Bangare ,"SentimentAnalysis Of Product Reviews For Ecommerce Recommendation"Mali,Conference Paper • January 2016
- [2] Anusha Sinha, Nishant Arora, Shipra Singh, MohitaCheema,AktharNazir, "Fake Product Review Monitoring Using Opinion Mining", International Journal of Pure and Applied Mathematics, Volume 119 No. 122018
- [3] Dhairya Patel, Aishwerya Kapoor, SameetSonawane, "Fake Review Detection Using Opinion Mining "International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 05 Issue:12|Dec2018
- [4] Aashutosh Bhatt, Ankit Patel, Harsh Chheda, KiranGawande," Amazon Review Classification and Sentiment Analysis",International Journal of Computer Science and Information Technologies (IJCSIT), Vol. 6 (6) 2015
- [5] Madhura N Hegde ,Sanjeetha K Shetty , Sheikh Mohammed Anas , Varun K, "Fake Product Review Monitoring" International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056, Volume: 05 Issue: 06 | June2018
- [6] SanthoshKumar K L Jayanti Desai Jharna Majumdar "OpinionMiningand SentimentAnalysisOnOnlineCustomerReview"Confere nce Paper · December [2016]