

An Efficient Hybrid Approach to Detect Spam in Product based User Review

Avishi Kansal

Computer Science Department
Galgotias College of
Engineering and Technology
Greater Noida

Hari Shankar Aggarwal

Computer Science Department
Galgotias College of
Engineering and Technology
Greater Noida

Aatif Jamshed

Computer Science Department
Galgotias College of
Engineering and Technology
Greater Noida

ABSTRACT

All e-commerce sites provide facility to the users for giving views and experience of the product and services they experienced. The customer's reviews are increasingly used by individuals, manufacturers and retailers for purchase and business decisions. As there is no scrutiny over the reviews received, spammers produce synthesized reviews to promote some products/brand and demote competitors' products/brand for profit or publicity. As the amount of spam has been increased tremendously using bulk mailing tools, there is an emerging need for spam detection. In this paper we propose an optimal approach to detect spam reviews based on number of reviews posted per day from a particular IP address and geographic location. In case of spam, it blocks the spammer's IP and also send a mail intimation to give an alert. It performs feature extraction based on the authentic reviews and also provides a star rating system. In our work we have combined LSVD and LSI algorithms to guarantee very high detection rates as well as feature extraction facility. Other concepts like ontology, spam dictionary, sentiment analysis, indexing, decision tree, opinion mining, clustering have also been included to provide the most efficient approach.

General Terms

Support Vector Machine, Review Manipulation, Tokenization.

Keywords

Review, spam-detection, opinion, feature extraction, positive and negative review, spam dictionary.

1. INTRODUCTION

As we already have visually perceived more and more users now preferring online shopping because of convenient platform like internet. This shopping is largely influenced by reviews. But research has shown all these reviews might not be authentic or trustworthy. Many suppliers, vendors or publishers incline to manipulate these reviews to promote their sales or on other hand they might post negative reviews to demote other product. These fake reviews are considered review spam and can have a great impact in the online marketplace due to the importance of reviews. Indicting such fake or erroneous reviews is called as review manipulation.

Reviews manipulation is basically, writers, publishers, or company people or any third-party writing untrue comments or reviews on behalf of customer when needed, to increase the sales of their products. A customer review consists of two parts, one with star rating and other with textual comments. If

unauthentic person posts review, he may either give high rating to the product or can manipulate textual comment. So by analyzing the spam words present in a review or number of postings from the same location and device in a day, we can detect manipulated reviews.

1.1 Motivation

The reviews online are advantageous only if they are posted appropriately without any incorrect intention. Day by day Review Manipulation is getting worse and also more sophisticated. Detecting spam reviews or opinions have become more and more critical. The position is at present quite bad. According to our annual Local Consumer Review Survey, the importance of online reviews in the purchasing decision is growing day by day. Many customers even say that they wouldn't buy product without reading online reviews. Reviews become even more important in the situation where customers are not able to test products practically prior to purchasing it. But these reviews have some drawbacks which give motivation to review spam detection-bad publicity, some reviews are fake given for supporting or devaluing product.

1.2 Objective

To propose an approach that will identify "illegal" activities (e.g., writing fake reviews/ Opinions) that try to mislead readers and to Propose a system that will filter out such spam content from reviews and prevent users from getting mislead. There are two types of review spam, one is manipulated review that deliberately mislead readers by giving undeserving positive reviews to promote particular products and malicious negative reviews to demote others, and other is non reviews (i.e. ads) which contains no opinions to the product.

Our objective of this work is to highlight review spam in order to shed some light on the trustworthiness of on-line reviews and to detect possible spam activities. We hereby propose a hybrid software that detects spam reviews, block spammer's IP, send a mail intimation in case of spam review, provide a star rating of a product based on only authentic reviews, and keeps check on IP address, Latitude-Longitude(Location), account used, and number of reviews posted per day. It also extracts the core features of the products based on the genuine reviews. It uses the concepts of ontology, spam dictionary, sentiment analysis, indexing, decision tree, opinion mining, clustering etc. It encompasses algorithms like LSVD, LSI, etc.

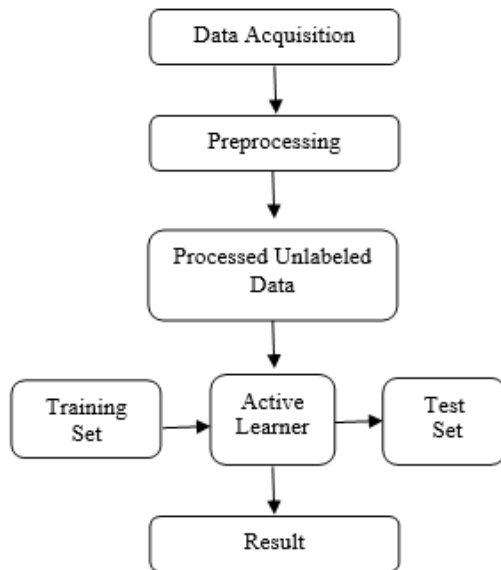


Fig 1 : Data flowchart of input character

2. PREVIOUS WORKS

This paper gives study of identifying manipulated reviews using different methods. As this is very newly introduced problem because of intensive use of e-commerce websites, there are very few works related to this.

In [1], the first attempt for identifying manipulated reviews i.e. detecting review spam was taken. In this paper, they have highlighted that there are two types of review spam, one is manipulated review which will mislead the customer and another is non-review i.e. it is not giving any actual opinion about the product, it can be advertisement of product. The objective of this paper was to perform spam detection based on duplicate findings and classification. They have classified the reviews in two categories as spam and non-spam review. This paper performs spam detection based on two methods.

The next paper [2] introduces use of temporal pattern discovery in identifying manipulated reviews. To address this problem, this paper states that the normal reviewers' arrival pattern is stable and uncorrelated to their rating pattern temporally. In contrast, spam attacks are usually unstable and either positively or negatively correlated to the rating. Thus, it is proposed that to detect such attacks via unusually correlated temporal patterns. [4] Exploits machine learning methods to identify such fake reviews. One way of identifying fake review is that reviewer is anonymous or fake. This paper discussed a framework of product review mining system. This paper provides

Review related features: 1.Content Feature 2. Sentiment features 3. Product features and

Reviewer Related Features: 1. Profile Features 2. Behavior Features Above work uses supervised learning methods and analyze the effect of various features on identifying manipulated reviews and two view semi-supervised method to exploit the large amount of unlabeled data.

In [6], the study employs Decision tree to improve the classification performance of manipulated reviews. This paper introduces eight potential factors for identifying manipulated reviews using correlation analysis and extracted knowledge rules. The eight potential factors are:

1. Text Difficulty:
2. TTR (Type-Token Ratio)
3. Tokens
4. Positive sentiment
5. Negative Sentiment
6. Sentiment
7. Product Characteristics
8. Expertise

Based on these eight potential factors proposed work classifies reviews using Decision Tree, along with that correlation analysis and knowledge rules discovered by decision tree are used to select key review manipulation attribute.

[17] Proposes a hybrid cuckoo with harmony search for feature selection, to select the optimized feature subset from the dataset and Naive Bayes is used for classification. Experimental results show that the hybrid cuckoo with harmony search is capable of identifying good quality feature subsets. The resulting classification accuracy tested and is compared with the whole feature subset. In almost all aspects, the proposed approach delivered considerably better results than binary cuckoo search. The proposed hybrid feature selection method for review spam detection provides better classification accuracy with an optimized feature subset.

3. MATHEMATICAL BACKGROUND

In this section we will discuss all the significant equations and formulas which are necessary to conduct our experiments. There is a little knowledge about the techniques like singular value Decomposition, Term Frequency (tf), Inverse Document Frequency (IDF), Cosine similarity some accuracy methods are very much needed. A. Singular value Decomposition Linear algebra has got a decomposition method in which the matrix factorization will take place. Before we proceed with SVD, Eigenvector, Term vector and Document Vector need to define. From the theory of Linear Algebra, we can define the SVD as. It is a matrix decomposition of the form;

$$X = U\Sigma VT \quad (1)$$

Where, U contains the Eigenvector of XX^T and V contains Eigenvector of XTX i.e., U matrix is formed by right singular vectors like $u_1 u_2 u_3$ and V matrix is of the left singular vectors like $v_1 v_2 v_3$ and then Σ represents the singular values like $\sigma_1, \sigma_2, \sigma_3$. Columns of U matrix are formed in such a way that all needy vectors which will be kept on the left most side that is the property how they are built with, i.e., essentially the first column of the matrix represent the terms which tends to occur in all of the documents, which is called as the Term vector. Rows of the V matrix are arranged in such a way that, vectors represent the all the documents in the data text corpus into the Vector Space are called as the Document vectors.

3.1 Vector Space Model

In machine learning, we have an algebraic model which represents the textual data from the document corpus as a vector, necessarily the components of this vector represent the value or the importance of the term. Usually the text is represented as a vector of numbers. It shows the features which are to be extracted from the text source. In general, the

queries and documents are represented as the vectors.

$$d_j = (w_1, j, w_2, j \dots w_t, j) \quad (2)$$

$$q = (w_1, q, w_2, q \dots w_t, q) \quad (3)$$

Where d represents the documents and q represent the query. In here, every dimension relates to a term. The pictorial representation of the vector space model is given in the below figure 1.

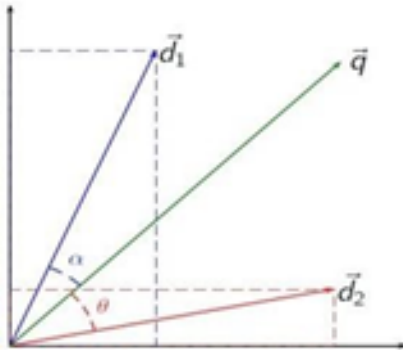


Fig 2 : Vector Space model

The term is generally a keyword or phrase; the dimensionality here in a vector space is the total number of the words presents in the vocabulary. This vector operation application comes when there is a need for comparison of two documents.

$$C.Tf - Idf \quad (4)$$

Term frequency (Tf), is the number of times the particular word appears in a document [13]. The weight of the term in a document is directly proportional to that of the term count or frequency, which is represented as tf (t, d);

$$f(t, d) = \sum fr(x, t) \quad (5)$$

Where its value will be 1 if $x=-1$ and otherwise 0. Inverse document frequency shows that the importance of word the particular document, it can be treated as the log scale function which contain the word, we can formulate it as;

$$idf(t, d) = \frac{\log idf t D(.) \log N}{|\{d \in D t \in d\}|} \quad (6)$$

3.2 Cosine Similarity

By considering the inner product in the space between two vectors as the documents, we find the similarity between these documents is ideally called as the cosine similarity. Usually this is done at the vector level with which the documents are represented [14].

It can be calculated by using the dot product formulas like Euclidean (7) and (8):

$$Similarity = cosine(\alpha) = \frac{A \cdot B}{||A|| \cdot ||B||} \quad (7)$$

4. PROPOSED WORK

The process starts with the collection of reviews from various Internet sources. To find out the spam reviews, we carry out a procedure which is given in the below algorithms.

Initially as the user submits the review, all his details are recorded for the further procedure.

IP Address, Geographic Location, Date, Category of Product, User Name, Product Name, E-mail Id is submitted along with the review to the system. Keeping track of user details is necessary for the next step of the procedure. Also it helps the system to take suitable action as required by the situation.

4.1 Spam Detection

Temp Review Table keeps the records of all the reviews submitted to the system along with the user details. Before any review is posted to the actual website, it is first checked whether or not more than 5 reviews from that same IP address and Geographic location are available in the table. This condition leads us to two possibilities:

- Already 5 reviews, from the same IP Address and Geographic location on that day have been posted and are available in the Temp Review Table.
- Less than 5 reviews from that IP Address and Geographic location have been posted on that day.

4.2 Blocking of IP Address

If former one is the case then the review is validated for IP blocking and e-mail sending. An alert e-mail is sent to the user and it's IP address is blocked to send any more reviews. The submit button is disabled for that user and no more reviews can be submitted by the user even in the Temp table.

Table 1. Tokenization Mechanism

Action	Result
Input Data	This mobile is very expensive!! but the quality is very good.
Tokenization	['This', 'mobile', 'is', 'very', 'expensive', '!', '!', 'but', 'the', 'quality', 'is', 'very', 'good', '!', '.']
Lowercasing Letters	this mobile is very expensive!! But the quality is very good
Punctuation removal	this mobile is very expensive but the quality is very good
Stop Words Removal	['t', 'h', 'i', 's', ' ', 'm', 'o', 'b', 'i', 'l', 'e', ' ', 'i', 's', ' ', 'v', 'e', 'r', 'y', ' ', 'e', 'x', 'p', 'e', 'n', 's', 'i', 'v', 'e', ' ', '!', '!', ' ', 'b', 'u', 't', ' ', 't', 'h', 'e', ' ', 'q', 'u', 'a', 'l', 'i', 't', 'y', ' ', 'i', 's', ' ', 'v', 'e', 'r', 'y', ' ', 'g', 'o', 'o', 'd', ' ', '.']
Stemming	['mobile', 'very', 'expensive', 'quality', 'very', 'good']

4.3 Tokenization

If the latter case is true then Initialization of spam words is done and the review is converted in the form of string using comma concatenation. After that we split the review according to the commas and then perform tokenization. During tokenization review is stored in an array and each index is compared with spam dictionary. If word on any index position is present more than 2 times in the spam dictionary then it is defined as spam. If 2 or less spam words are present in the review then Singular Value Decomposition is used. We use concept of decision tree while performing the above procedure.

Decision trees uses tree structure to build the classification models. It divides a dataset into smaller subsets. Leaf node represents a decision. Based on feature values of instances, the decision trees classify the instances. Each node represents a feature in an instance in a decision tree which is to be classified, and each branch represents a value. Classification of Instances starts from the root node and sorted based on their feature values. Categorical and numerical data can be handled by decision trees.

4.4 Positive and Negative Dictionaries

If on tokenization, review is split into “m” words, then “m” matrices are formed out of those. The matrices are then compared to already defined Positive and negative dictionaries using the concept of Cosine Similarity. If number of matches to positive dictionary are greater than those to negative dictionary, then it is considered as positive review and added to positive review table, else negative review and

added to negative review table.

In the case where number of matches are same the review is discarded and not posted. If no of words in review are greater than those in pre-defined dictionaries then normalization is performed and biased entries are added.

Here we have also used C5 Classifier to form small clusters. The C5 classifier is tested first to classify unseen data and for this purpose resulting decision tree is used. C4.5 algorithm follows the rules of ID3 algorithm, similarly C5 algorithm follows the rules of algorithm of C4.5. C5 algorithm has many features like: The large decision tree can be viewing as a set of rules which is easy to understand. C5 algorithm gives the knowledge on noise and missing data. Problem of over fitting and error pruning is solved by the C5 algorithm. In classification technique the C5 classifier can anticipate which attributes are relevant and which are not relevant in classification. C5 is faster than C4.5. Memory usage is more efficient in C5 than C4.5. C5 gets smaller decision trees in comparison with C4.5. The C5 rule sets have lower error rates on unseen cases. So comparing with C4.5 the accuracy of result is good with C5 algorithm. C5 automatically allows removing unhelpful attributes.

4.5 Feature Extraction

After the reviews are placed in the corresponding positive or negative review table we apply feature extraction to the review.

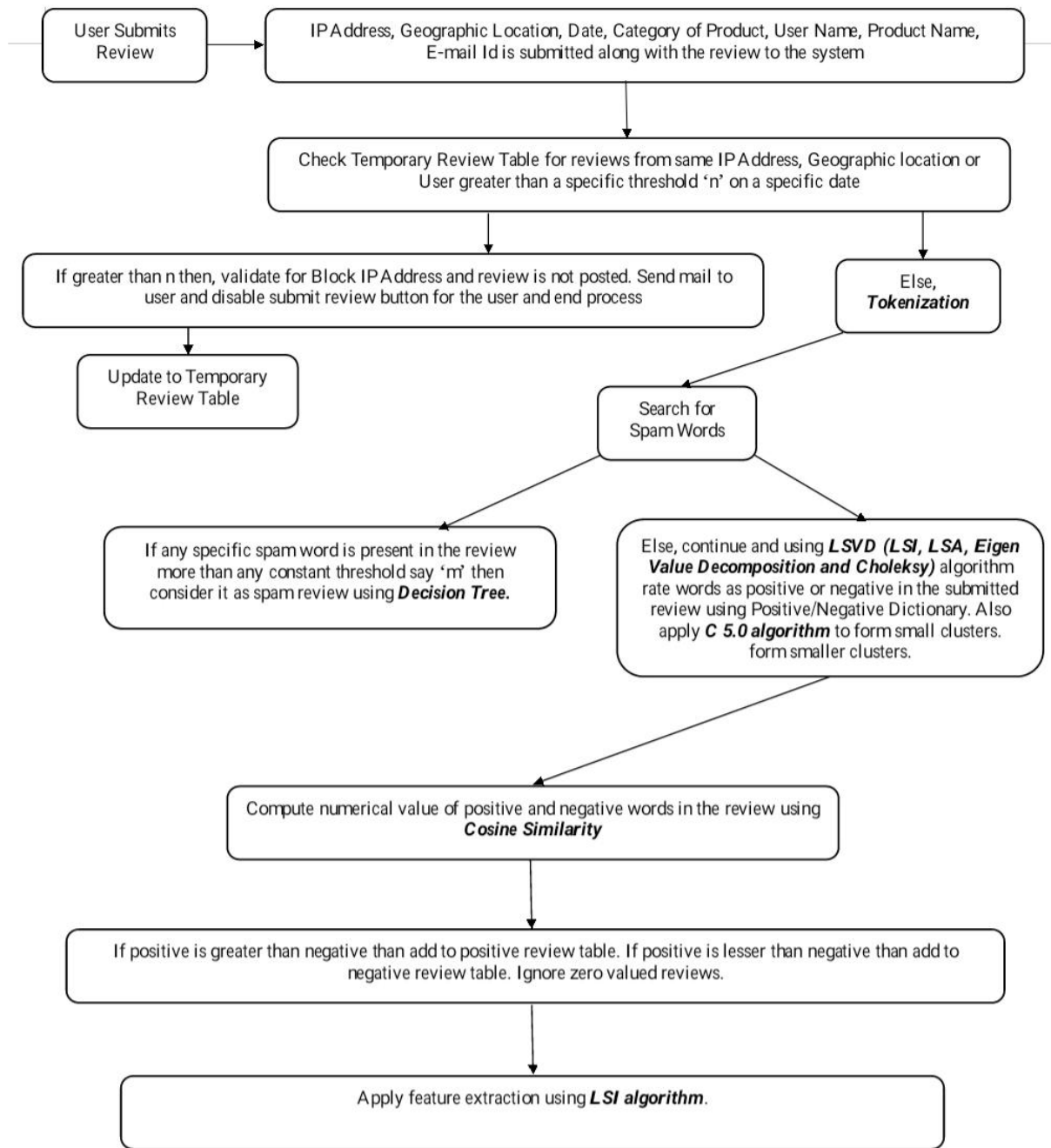


Fig 3 : Flowchart of Proposed Work

Later Semantic Indexing is used to extract the features of the product based on the review posted. LSI performs a low-rank approximation of document-term matrix. It design a mapping such that the low-dimensional space reflects semantic associations (latent semantic space). It computes document similarity based on the inner product in this latent semantic space.

5. RESULTS

The hybrid approach of the above explained methodology brings about the high yield of spam detection and classification of review into positive and negative ones. Since the number of algorithms are higher thus the overall traversal time and algorithm switch along with its run time brings in a degradation in time utilization. But the accuracy and hit

ratio's are higher. Thus on a compromise on the time brings about a better performance and output. We calculated the positive and negative reviews output. (Refer Fig 4) Previous works depicted an average of 89 percent but we could improve it to 91 percent. For that we went through the comments manually and compared with programmatic results, so the results obtained were showing an average percentage of matching, and the value was around 91 percent as shown in the below figure . The graph shows that, the line in orange color for manual analysis, which is supposed to be 100 percent accurate and on the other hand, the line in blue color with Programmatic analysis which was around 91 percent accurate.

We also studied the spam words table a more in detail. This

comprised of the occurrence and dominance of alphabets in a range of spam words. (Refer Fig 5) This study could help future researchers on focusing more certain alphabets for a better yield.

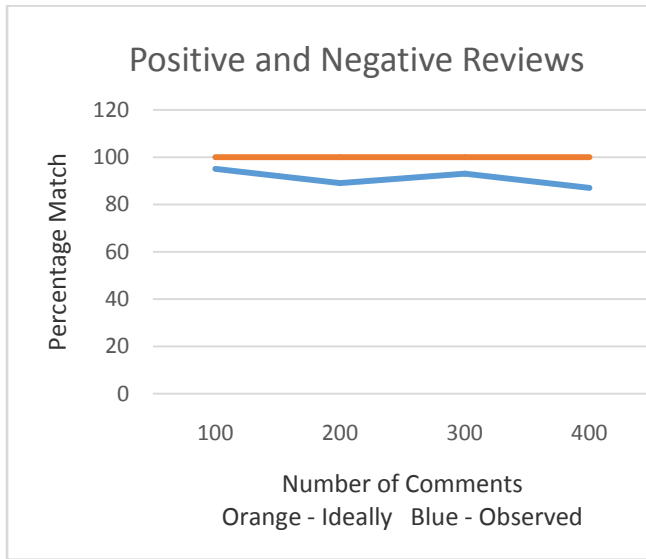


Fig 4 : Positive and Negative Reviews

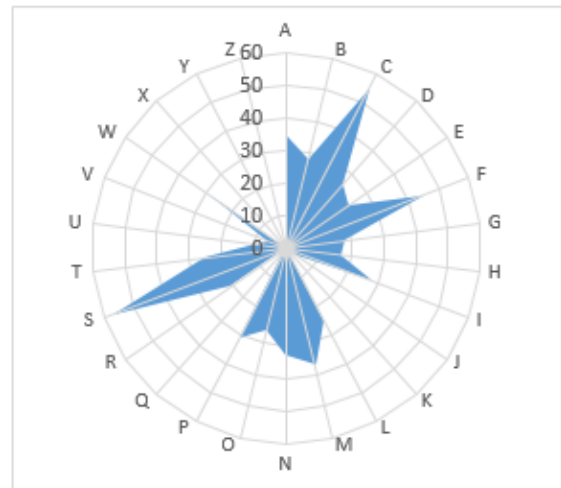


Fig 5 : Occurrences of Spam words alphabetically

Fig 6 : Temp Review Table

Id	Emailid	Review	Ip	Latitude	Longitude	Name	Date
6043	avishikansal1996gmailcom	Galaxy J2 is very bad but has a good screen	192.168.1.7	28.52434202	77.20531037	a	22/04/2018
6044	avishikansal1996gmailcom	IPHONE5 is very expensive	192.168.1.7	28.5244218	77.20521321	a	22/04/2018
6047	avishikansal1996gmailcom	Galaxy J2 is very bad	192.168.0.101	28.62944039	77.0818303	hari	23/04/2018
6048	avishikansal1996gmailcom	IPHONE5 is very bad	192.168.0.101	28.62944039	77.08180627	hari	23/04/2018
6049	avishikansal1996gmailcom	IPHONE5 is bomb	192.168.0.101	28.62935176	77.0818303	hari	23/04/2018
6042	avishikansal1996gmailcom	Don't buy these phones.	192.168.1.7	28.52438491	77.20516426	Avishi	22/04/2018
6045	avishikansal1996gmailcom	IPhone 5 is verygood	192.168.0.101	28.62942098	77.08183273	Avishi	23/04/2018
6046	avishikansal1996gmailcom	Galaxy J2 is verygood	192.168.0.101	28.62938582	77.08181191	somil	23/04/2018

Fig 7 : Spam Review Table

Id	Emailid	Review	Ip	Latitude	Longitude
5002	somillohanigmailcom	Iphone2 is not so good	192.168.56.1	28.62937928	77.08184761
5003	somillohanigmailcom	Iphone2 is not so good	192.168.56.1	28.62937928	77.08184761
5004	somillohanigmailcom	Iphone2 is not so good	192.168.56.1	28.62937928	77.08184761
5005	somillohanigmailcom	Iphone2 is not so good	192.168.56.1	28.62937928	77.08184761
5006	somillohanigmailcom	iphone2 is not so good	192.168.56.1	28.62930204	77.08187654
5007	somillohanigmailcom	Galaxy J2 is not so good	192.168.56.1	28.62927677	77.08187931
6002	somillohanigmailcom	Iphone2 is good phone	192.168.56.1	28.62928849	77.08186716
6003	somillohanigmailcom	Galaxy J2 is good phone	192.168.56.1	28.62929745	77.08185843
6004	somillohanigmailcom	Galaxy J2 is good phone	192.168.56.1	28.62929745	77.08185843

Fig 8 : Product rating

Id	Productname	PositiveRating	NegativeRating	TotalRating
1	Galaxy J2	5	7	12
2	IPHONE4	4	2	6
3	IPHONE5	0	2	2

We considered a few sets of input data to check how they respond to the algorithm and how efficiently it can put things in correct correlations and tables.

We submitted number of possible reviews for two categories of products to check feasibility and working of the algorithm.

The Temp Review had the reviews posted before being allotted to the spam or genuine reviews. The spam review table had reviews which were have one or more spam words. Review table has all the genuine reviews which would be displayed on the web page or review site. Product Rating is an

outcome of review table, each positive and negative word is counted and using Cosine Similarity the words become either positive or negative.

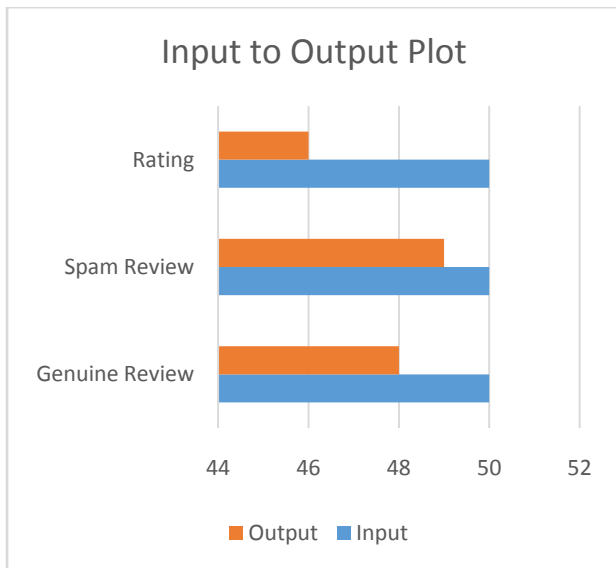
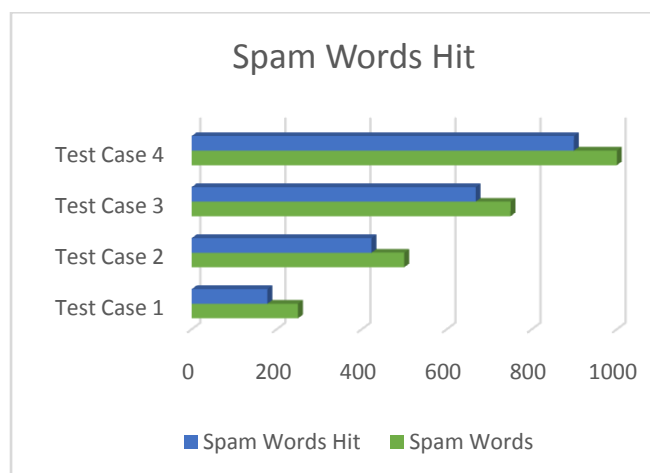


Fig 9 : Input to Output Plot

The hit ratio of spams for a set of 50 reviews was 49. The hit ratio of genuine reviews from a set of 50 was 48 and rating percentage was 92 percent where 8 percent were the sarcastic reviews as the algorithm does not account for sarcastic reviews and words (Figure 9).

We also tested the scalability of the algorithm by uniformly selecting a set of each alphabets of spam words to account for a definite size of spam word dictionary. In a set of 250, 178 spam words were caught which is almost 71.2 percent. When we took 500 spam words and with the same input data 423 spams were caught which is 84.6 percent. For 1000 it became 89.9 percent.



6. ACKNOWLEDGMENTS

Our deep thanks and gratitude to the people who have contributed to this paper and without whom this would not have been possible. Special thanks to almighty god, Mr Kamlesh Kumar Aggarwal, Mr Kavind Kr Kansal, our project coordinator Mr Manish Kumar (Assistant Professor) and Dr. Vishnu Sharma (Professor Head of Department).

7. CONCLUSION AND FUTURE SCOPE

Spam Reviews are predominant in today's digital world. From buying cosmetics to grocery all are supported by reviews. It creates a wider trustworthy opinion about choice of product for a buyer. It is definitely fulfilled by one's posting it but it is

moderated only by avoiding spam one's, which are nothing more than a marketing technique to mislead people. Thus Review detection is a dire need with comprehensive algorithms to not only discard spam reviews but also block spammers and provide an automatic opinion deduced by posted reviews. Thus to apply the above described method is a stepping stone in the improvement of e-commerce and online shopping.

The above method solves the problem of accuracy in review spam detection, moreover the previous algorithms lacked a holistic and overall spam detection with rating and mail shooting to those who post frivolous reviews. The blocking of IP address and limiting the number of reviews and posts on a particular day per user acts as a safety measure which increases the number of pure and genuine reviews.

The biggest limitation of the proposed work is that it cannot account for sarcastic words.

Since, it is highly applicable to the fields of E-Commerce, with higher order uses both the algorithms complexity as well as performance can be improved to a very high order. Some of the future scope are mentioned below-

1. Tracing of M.A.C address of spammer.
2. Application on Hadoop and Big Data.
3. Incorporation of sarcastic words in dictionary. The above work does not incorporate sarcastic words commonly used by reviewers. Sarcastic words change the tone of the review completely and might cause problem to algorithm to decide on which is a positive review or not. Thus effective machine learning is required for the same.

8. REFERENCES

- [1] Hari Shankar Aggarwal, Avishi Kansal, Aatif Jamshed, Noisy Information and Progressive Data-Mining giving rise to Privacy Preservation, ICACCA, Dehradun, 2017.
- [2] Nitin Jindal, Bing Liu, Review Spam Detection, ACM proceedings of 16th international conference on world wide web, pp-1189- 1190,2007.
- [3] Sihong Xie, Guan Wang, Shuyang Lin, Philip S. Yu, Review Spam Detection via Temporal Pattern Discovery, ACM 978-1- 4503-1462-6 /12/08.
- [4] Jiwei Li, Mylez Ott, Claire Cardie. Identifying Manipulated Offerings on Review Portals.
- [5] Fangtao Li, Minlie Huang, Yi Yang and Xiaoyan Zhu Learning to Identify Review Spam, Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence. M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.
- [6] Manali S.Patil, A.M.Bagade, Online Review Spam Detection using Language Model and Feature Selection, International Journal of Computer Applications 0975 – 8887 Volume 59– No.7, December 2012.
- [7] Long- Sheng Chen, Jui-Yu Lin, A study on Review Manipulation Classification using Decision Tree, IEEE conference publication, 2013.
- [8] N. Hu, L. Liu, V. Sambamurthy, Fraud detection in online consumer reviews, Decision Support Systems, vol. 50, pp. 614- 626, 2011b.
- [9] N. Hu, I. Bose, N. S. Koh, L. Liu, Manipulation of online reviews: An analysis of ratings, readability, and

- sentiments, *Decision Support Systems*, vol. 52, pp. 674-684, 2012.
- [10] Rajashree S. Jadhav et al, A New Approach for Identifying Manipulated Online Reviews using Decision Tree / *IJCSIT International Journal of Computer Science*
- [11] Mukherjee, A., Liu, B., Wang, J., Glance, N. and Jindal, N. 2011, Detecting Group Review Spam, In *Proceedings of the 20th WWW*, pp. 93-94.
- [12] Ntoulas, A., Najork, M., Manasse M. and Fetterly, D. 2006, Detecting Spam Web Pages through Content Analysis, In *Proceedings of the 15th WWW*, pp. 83-92.
- [13] Li, F., Huang, M., Yang, Y., & Zhu, X. 2011. Learning to Identify Review Spam. In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume Three* pp. 2488–2493. AAAI Press. doi:10.5591/978-1-57735-516-8/IJCAI11-414
- [14] Mukherjee, A., Liu, B. and Glance, N. 2012. Spotting Fake Reviewer Groups in Consumer Reviews, In *Proceedings of the 21st WWW*, pp. 191-200.
- [15] Xinkai Yang 2015. One Methodology for Spam Review Detection Based on Review Coherence Metrics, *International Conference on Intelligent Computing and Internet of Things ICIT*
- [16] M.N. Istiaq Ahsan*, Tamzid Nahian, Abdullah All Kafi£, Md. Ismail Hossain†, Faisal Muhammad Shahµ, 2016 Review Spam Detection using Active Learning
- [17] S.P.Rajamohana, K. Umamaheswari, S.Vasantha Keerthana, An Effective Hybrid Cuckoo Search with Harmony Search for Review Spam Detection, 3rd International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics AEEICB17
- [18] Shashank Kumar Chauhan, Anupam Goel, Prafull Goel, Avishkar Chauhan and Mahendra K Gurve, 2017. Research on Product Review Analysis and Spam Review Detection, 2017 4th International Conference on Signal Processing and Integrated Networks SPIN
- [19] Aatif Jamshed and Pawan Singh Mehra 2012. Modified Block Playfair Cipher using Random Shift Key Generation, *International Journal of Computer Applications*, Vol. 58, pp. 2012/1/1.
- [20] Aatif Jamshed ,Surbhi Chandhok and Romil Anand 2017, Analysis of Sequential Mining Algorithms, *International Journal of Computer Applications*, Vol. 165, pp. 12-2017/5.
- [21] Aatif Jamshed ,Surbhi Chandhok and Romil Anand 2017, An Analysis of Sentimental Data using Machine Learning Techniques, *International Journal of Computer Applications*, Vol. 166, pp. 3-2017.
- [22] Aatif Jamshed ,Garima Verma 2013. Mobile Devices integration with Grid by Using Efficient Scheduling for Local Resource, *Journal of Advanced Computing and Communication Technologies* ISSN: 2347 – 2804 Volume No. 1 Issue No.2, December 2013
- [23]