IJCSE International Journal of Computer Sciences and Engineering Open Access
Review Paper Vol.-7, Issue-5, May 2019 E-ISSN: 2347-2693

Improved Text Mining Techniques for Spam Review Detection

Akshat A. Uike[1*], Sumera W.Ahmad[2], Sunil R.Gupta[3]

[1,2,3]CSE, Prof Ram Meghe Institution of Technology & Research, Badnera, Maharashtra, India

Corresponding Author: akshat.uike@rediffmail.com,

DOI: https://doi.org/10.26438/ijcse/v7i5.147152 | Available online at: www.ijcseonline.org

Accepted: 09/May/2019, Published: 31/May/2019

Abstract— Text mining has played a important role in providing product recommendations to users. Online reviews have become an important factor when people make purchase and business decisions. Efficient recommendation systems help in improving business and also enhance customer satisfaction. The credibility of purchasing a product highly depends on the e-commerce online reviews. However most of people wrongly promote or demote a product by buying and selling fake reviews. Many websites have become source of such opinion spam. These fake/fraudulent reviews are deliberately written to trick potential customers in order to promote/hype them or defame their reputations. Our work is aimed at identifying whether a review is fake or truthful one. Naïve Bayes Classifier, Logistic regression and Support Vector Machines are the classifiers using in our work. This in turns leads to recommending undeserving products. This paper aims to classify online reviews into groups of positive or negative polarity by using machine learning algorithms. In this study, we find online reviews using SA methods in order to detect fake reviews. SA and text classification methods are applied to a dataset of movie reviews. More specifically, we compare five supervised machine learning algorithms: Naïve Bayes (NB), Support Vector Machine (SVM), K-Nearest Neighbours (KNN-IBK) for sentiment classification of reviews using two different datasets, including movie review dataset and movie reviews dataset. The measured results of our experiments show that the SVM algorithm outperforms other algorithms, and that it reaches the highest accuracy not only in text classification, but also in detecting fake reviews.

Keywords— Amazon E-Commerce dataset, Active Learning, Dataset acquisition, Data pre-processing, KNN Classifier, Rough Set Classifier, Support Vector Machine.

I. INTRODUCTION

The online has changed our lives since it was introduced. With rapidly expansion and usage of Internet people are now totally dependent on the web services, which also changed people's behavior of communicating and expressing their views. People gives their views in their respective discussion groups, forums, social media, and blogs and in e-commerce website for a product/service. These contents are user generated which are written in natural language. Opinion sharing on a product/service is based on their personal experience which is called as reviews.

Reading about product reviews before buying the product becomes a habit, especially for potential customers. If customer ready to buy a product, they usually read reviews of other customers about the current product. If the review is positive, there is a big chance to purchase the product, otherwise if review is negative, they tend to buy other product. While, for an industry/company, the positive reviews from customers can generate decisive financial benefits for companies, the positive review can be using as input for decisions related to product design and what services are provided to customers.

In order to solve this malignant problem, we propose an interactive semi-supervised model to identify fake reviews which is evaluated later on using real life data and compared with some sophisticated prior research work. In this paper we propose a novel approach of Active Learning to detect fake reviews. We have using original dataset of Amazon reviews. Make their decisions of whether to purchase the products or not by analyzing and reflecting the existing opinions on those products.

Opinion mining has played a significant role in providing product recommendations to users. Efficient recommendation systems help in improving business and also enhance customer satisfaction. The credibility of purchasing a product highly depends on the online reviews. However many people wrongly promote or demote a product by buying and selling fake reviews. Many websites have become source of such opinion spam. This in turns leads to recommending undeserving products. This literature survey is done to study the various fake review detection techniques in detail and to get ourselves familiar with the works done on this subject.

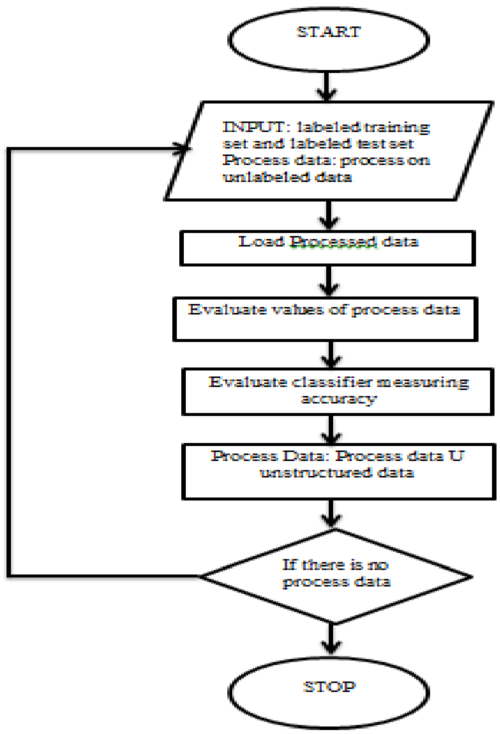

Fig1: Flowchart of data preprocessing

Data Acquisition: Data extracted from dataset is used as the unlabeled data, labeled dataset created by Ottertail is used for both training and testing purpose in this method. As the dataset created already preprocessed so we only do some preprocessing on the unlabeled dataset. We use reviews extracted from as the unlabeled data in this learning method.

Data Preprocessing: Unstructured data in MS Excel format acquired from the source is converted into structured data i.e. in My SQL Database format. Preprocessing procedures includes- tokenization & lowercasing letters, removing stop words, removing punctuations, stemming etc.

Active Learning: Active learning is a special case of semi supervised machine learning which can interactively request the user to determine the class of some unknown data points to achieve the desired results.

Labeling the whole dataset manually is extremely time consuming and labor intensive. So, the algorithm actively queries the user for labeling the new, confusing data points.
In this type of learning, learner itself chooses the data point examples that's why it needs a much lower number of examples to learn a concept than it is required in typical supervised learning

The algorithm trains the model based on a training dataset and evaluates using a test dataset.
After each evaluation, the algorithm selects a certain number of samples from the unlabeled dataset in order to be labeled by a user and is then added to the existing train dataset. The model will start training again with the new better training set.
The selection of unlabeled samples is based on a decision function which is the distance of the samples X to the separating hyper plane. Although the distance is between [-1, 1], we use absolute values because we need the confidence levels.

Rough Set Classifier:
It is the machine learning technique which has the concept of set theory to make decision. We discuss the impact of inductive reasoning on the rough set to concept approximation. We show how among formulas used for classifier construction from decision rules one can search for new patterns relevant for the incremental concept approximation.

We present applications of rough set methods for feature selection in pattern recognition. In the overview of methods for feature selection we discuss feature selection criteria, including the rough set based methods. Our algorithm for feature selection is based on an application of a rough set method to the result of principal components analysis (PCA) used for feature projection and reduction. Finally, the paper presents numerical results of face and mammogram recognition experiments using neural network, with feature selection based on proposed PCA and rough set methods.

SVM (SUPPORT VECTOR MACHINE) :
Support Vector Machine" (SVM) is a supervised machine learning algorithm which can be used for both classification and regression challenges. Support Vector Machine, abbreviated as SVM can be used for both regression and classification tasks. But, it is widely used in classification objectives.

KNN (K-NEAREST NEIGHBORS):
In this article, we will talk about another widely used classification technique called K-nearest neighbors (KNN). Our focus will be primarily on how does the algorithm work and how does the input parameter effect the output/prediction.
KNN algorithm fairs across all parameters of considerations. It is commonly used for its easy of interpretation and low calculation time.
This Result Paper contributes its own approach to the field of opinion spam detection by considering spamming clues linked to both reviews and reviewers. The examined indicators of spam activity include review content, rating and timestamp but the real focus is on review author behaviour,

as an established spammer is directly linked to highly suspicious reviews. Review author behaviour analysis is twofold, including both activity on a particular product but also past reviewing activity, as a reviewer's history of contributions is an indication of their overall reputation as an author. An additional employed technique is responsible for detecting bursts in a product's reviewing activity. And while literature focus has been shifted from review content categorization to primarily spammer behaviour analysis, another contribution by this Review Paper involves an examination of review context by utilizing a deep learning classifier and approach for the first time, with the intention of ascertaining its usefulness in detecting opinion spam.

The propose methodology is evaluated on a dataset of real reviews from E-commerce website in accordance with the rest of the literature. Despite the absence of ground truth annotated data, which is a major issue in evaluating opinion spam detection methods, an alternative evaluation method is employed and the reported results display positive detection accuracy, attesting to the chosen spam indicators' reliability. The ultimate goal is to present an effective spam filtering system, ready to be put to practical use thanks to the straightforward approach of the propose methodology. Because it is extremely important to combat manipulative and deceptive practices in favour of a safer, ethical and unbiased opinion-sharing online space.

RESEARCH OBJECTIVES:-

Individuals and organizations increasingly use reviews from the social media for:

1. For making decisions relating to product purchases
2. For product designing and marketing
3. To make election choices
4. 31% of consumers read online reviews before actually making a purchase (rising)
5. By the end of 2014, 15% of all social media reviews will consist of company paid fake reviews

Our main objectives are:

1. Working on Amazon reviews data set.
2. In review data set we need to identify that review given by user was genuine or fake.
3. NLP based sentiment analyzer and text mining algorithms we will be using to classify, predict (positive, negative and neutral) reviews.
4. Any occurrence of unrelated (non referrals) words will leads to fake review.
5. Fake reviews- Unauthorized, Non trustworthy, Contents of unrelated words.

II. RELATED WORK

A number of literature review have been conducted which focusing on spam detection in e-mail and on the web, however, only recently have any studies been conducted on opinion spam.

In the paper [1] Jindal and Liu (2008) have worked on "Opinion Spam and Analysis" and have found that opinion spam is widespread and different in nature from either e-mail or Web spam. They have classified spam reviews into 3 types: Type 1, Type 2 and Type 3. Here Type 1 spam reviews are untruthful opinions that try to mislead readers or opinion mining systems by giving untruthful reviews to some target objects for their own gains. Type 2 spam reviews are brand only reviews, those that comment only on the brand and not on the products. Type 3 spam reviews are not actually reviews, they are mainly either advertisements or irrelevant reviews which do not contain any opinions about the target object or brand. Although humans detect this kind of opinion spam they need to be filtered, as it is a nuisance for the end user. Their investigation was based on 5.8 million reviews and 2.14 million reviewers (members who wrote at least one review) crawled from amazon.com and they have discovered that spam activities are widespread. Thus they had to use duplicate spam reviews as positive training examples and other reviews as negative examples to build a model.

In the paper [2] Ott, et al. 2011, they have given focus to the deceptive opinion spam i.e. the fictitious opinions which are deliberately written to sound authentic so as to deceive the user. The user cannot easily identify this kind of opinion spam. They have mined all 5-star truthful reviews for 20 most famous hotels in Chicago area from trip advisor and deceptive opinions were gathered for the same hotels using amazon mechanical trunk (AMT). They first asked human judges to evaluate the review and then they have automated the task for the same set of reviews, and they found that automated classifiers outperform humans for each metric. The task was viewed as standard text categorization task, psycholinguistic deceptive detection and genre identification. The performance from each approach was compared and they found that the psycholinguistic deceptive detection and genre identification approach was outperformed by n-gram based text categorization, but a combined classifier of n-gram and psychological deception features achieved nearly 90% cross-validated accuracy. Finally they came into a conclusion that detecting deceptive opinions is well beyond the capabilities of humans. Since then, various dimensions have been explored: detecting individual [3](Lim et al., 2010) and group spammers [4] and [5](Mukherjee et al., 2012), time-series (Xie et al., 2012) and distributional analysis (Feng et al., 2012a).

In [6] Mukherjee, et al., 2013, authors have briefly analysed "What yelp filter might be doing?" by working with different combination of linguistic features like unigram, bigram, distribution of parts of speech tags and yielding detection accuracy. Authors have found that a combination of linguistic and behavioral features comparatively yielded more accuracy.

In paper [9] Yoo and Gretzel (2009) gather 40 truthful and 42 deceptive hotel reviews and, using a standard statistical test, they have manually compared the psychologically relevant linguistic differences between them.

In paper [12] Prof. M. A..Pund et al presents an innovative and effective pattern discovery technique which includes the processes of pattern deploying and pattern evolving, to improve the effectiveness of using and updating discovered patterns for finding relevant and interesting information. In proposed system they can take sufficient .txt file as inputs & they apply various algorithms & generate expected results.

In paper [13] Hu, Bose, Koh & Liu, (2012) Internet users can easily and openly express their opinion about a product or brand by using social media or online product reviews and reach up to millions of potential buyers. With the assistance of opinion mining tools, businesses can retrieve valuable information with regard to product, service and marketing improvements from this kind of user-generated content (Heydari, Tavakoli, Salim & Heydari, 2016).

Online opinions thus can have great impact on brand and product reputation as well as related sales and management decisions. This gives an incentive to businesses to create, for example, positive fake reviews on their own products and negative fake reviews on their competitors' products (Akoglu & Rayana, 2015). There is a variety of ways to spam the internet with fake content. For instance, by hiring professional firms which are specialized in writing spam reviews, by using crowdsourcing platforms to employ review spammers or by using robots to create synthetic reviews. Reviews produced by someone who has not personally experienced the subjects of the reviews are called spam reviews (Heydari et al, 2015). The person who creates the spam review is called an individual review spammer. Individual review spammers working together with other review spammers are group spammers (Mukherjee, Liu & Glance, 2012).

Due to the amount of reviews posted online and the proficiency of review spammers, it is often very hard to detect spam reviews and separate them from trustworthy ones. However, it is essential not only for businesses but also for customers that review spam can be identified and removed in a reliable way. Researchers have suggested a variety of methods and tools to identify spam reviews, review spammers and spammer groups (e.g. Jindal & Liu, 2008; Mukherjee et al, 2012; Xie et al, 2012). One of these tools is reviewskeptic.com1 developed by Ott, Choi, Cardie, & Hancock (2011). The authors claim that review sceptic is able to detect spam reviews on hotels based on psychological and linguistic criteria with 90% accuracy. However, hotel reviews are only a fraction of the opinions posted on the Internet. Many reviews are related to individual products, services, brands or stores. Review sceptic claims to be a well-working yet very specialized tool for spam review detection. The aim of this research is to assess review sceptic's performance on non-hotel reviews and based on the existing literature to give recommendations on how the tool could be enhanced to detect also non-hotel review spam effectively.

The identification of spam reviews will be a relevant research topic as long as opinions will be expressed on the internet. Not only the tools for detection are improving but also the ways of producing review spam are getting more advanced. For example, sellers on Amazon 2 now have the opportunity to provide their products for free or at a discount in exchange for a review. Thereby, the review is still marked as a verified purchase and thus seems more trustworthy to potential buyers and to conventional review spam detection methods. However, the honesty of the reviews obtained in this way is highly questionable (Bishop, 2015). This example shows the importance of developing, testing and improving new methods for spam review detection which can keep up with the novel ways of producing spam reviews constantly.

In paper S P. Rajamohana et al [17] 2017 focused light on deceptive reviews that are available in the internet which increasingly affects businesses and customers. Hence it is important to detect and eliminate such fake reviews from online websites. This paper reveals several approaches used for review spam detection and performance measures were identified.

III. METHODOLOGY

Propose system will automatically classify user opinions into fake, genuine and neutral. This automatic system can be useful to business organization as well as to customers. Business organization can monitor their product selling by analyzing and understand what the customers are saying about products. Customers can make decision whether he/she should purchase or not purchase the products. This can helpful to people to buy valuable product and spend their money on quality products.

The posted reviews are useful only if reviews posted without any incorrect intention. Online survey shows that around 90% of customers are satisfied after reading 10 on 10 rating and good review. Many customer decide to buy product or not only by following reviews, But when intention of a person is not good, behind giving review such opinion. There is needed to detect such activities to make sure that opinions/reviews on the web are trustworthy source of information. Therefore, it is very important to know if the reviews are genuine because fake reviews lead to false reputation of the product and mislead the user.

In order to solve this malignant problem, we propose an interactive semi-supervised model to identify fake reviews which is evaluated later on using real life data and compared with some sophisticated prior research work. In this paper we propose a novel approach of Active Learning to detect fake reviews. We have using original dataset of Amazon reviews. Make their decisions of whether to purchase the products or not by analysing and reflecting the existing opinions on those products.

This paper aims to classify reviews into groups of positive or negative polarity by using machine learning algorithms. In this study, we analyses online movie reviews using SA methods in order to detect fake reviews. SA and text classification methods are applied to a dataset of movie reviews. More specifically, we compare five supervised machine learning algorithms: Naïve Bayes (NB), Support Vector Machine (SVM), K-Nearest Neighbours (KNN-IBK) for sentiment classification of reviews using two different datasets, including movie review dataset and movie reviews dataset.

DESIGN AND IMPLEMENTATION

The effectiveness of the proposed approach is performed and evaluated on artificial data sets generated by a traffic road simulator.

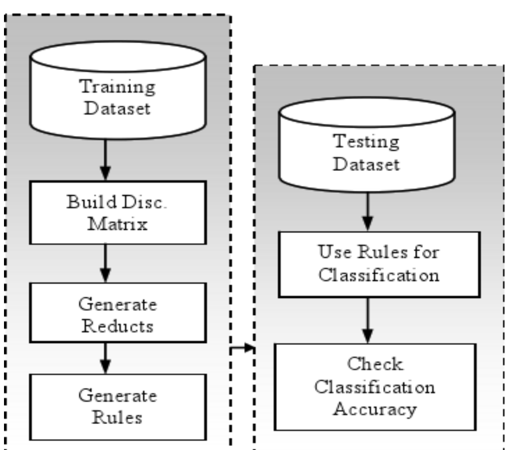

Fig 2: Work Flow Diagram of Rough set Classifier Algorithm

In this article, we will talk about another widely used classification technique called K-nearest neighbors (KNN). Our focus will be primarily on how does the algorithm work and how does the input parameter effect the output/prediction.

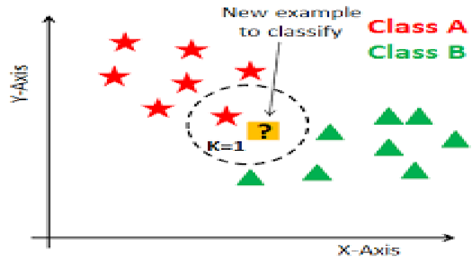

Fig 3: positive or negative polarity by using machine learning algorithm

Algorithm:

Let may be the number of training data samples. Let p be an unknown point.

Store the training samples in an array of data points arr[]. This means each element of this array represents a tuple (x, y).

for i=0 to m

Calculate Euclidean distance d(arr[i], p).

Make set S of K smallest distances obtained. Each of these distances correspond to an already classified data point.

Return the majority label among S.

Support Vector Machine (SVM): We looked at the machine learning algorithm, Support Vector Machine in detail. we discussed its concept of working, process of implementation in python, the tricks to make the model efficient by tuning its parameters, Pros and Cons, and finally a problem to solve. I would suggest you to use SVM and analyses the power of this model by tuning the parameters.

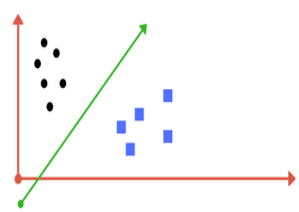

Fig 4 : SVM and analyses

Finally, a different take on the topic of opinion spam detection was proposed by, in which fuzzy logic techniques were utilized in order to quantify the spam city of a product review, as a module of a larger system aimed at analysing product reviews and determining a score for the quality of a product. The authors considered five types of spam reviews, namely Duplicate Reviews, No-Reviews, Brand Reviews, Intentional Reviews and Semantically Similar Reviews, a definition referring to written reviews influenced by other reviews.

IV. Results and Discussion

We compare our proposed model with some recent and sophisticated methodology to justify the competency of it.

The following mentioned approaches are based on review-content centric features. The comparative analysis based on count of fake and genuine review and Execution time of algorithm

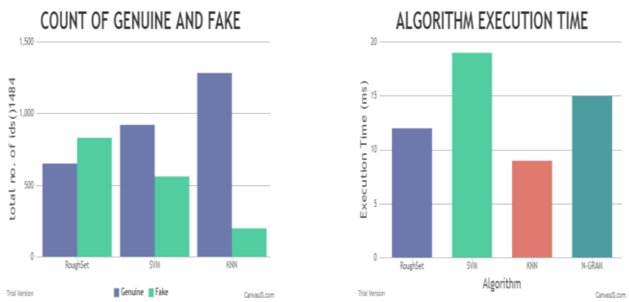

Fig 5: comparative analysis based on count of fake and genuine review and Execution time of algorithm

Since the detection accuracy percentage varies with different sets of test reviews, we have used 5-fold cross validation technique by considering folds of trained dataset and test dataset in the ratio of 80:20. Test frequency accuracy obtained for unigram presence, unigram frequency, bigram presence, bigram frequency and review lengths.

This paper is dedicated to evaluating the proposed methodology introduced previously by undertaking an experimental analysis. First, an explanation of the experimentation process is provided as well as any other details necessary for the subsequent sections and then, the procured experimentation results on the spam detection techniques are presented. Finally, an overview and conclusions section discusses the results and their significance.

V. CONCLUSION AND FUTURE SCOPE

In this way we proposing system will automatically classify user opinions into fake, genuine and neutral. This automatic system can be useful to business organization as well as to customers. Business organization can monitor their product selling by analyzing and understand what the customers are saying about products. Customers can make decision whether he/she should purchase or not purchase the products. This can helpful to people to buy valuable product and spend their money on quality products.

For future work, we would like to extend this study to use other datasets such as movie dataset or eBay dataset and Fake New Detection and use different feature selection methods. Also, we can use better sentiment classification algorithms to identify fake and genuine reviews using numerous tools such as Statistical Analysis System (SAS), Python and R studio.

REFERENCES

[1] Jindal, N., Liu, B.: "*Opinion Spam and Analysis*" in Proceedings of the International Conference on Web Search and Web Data Mining (pp. 219–230), 2008.

[2] Ott, M., Choi, Y., Cardie, C., Hancock J.T*," Finding Deceptive Opinion Spam by any Stretch of the Imagination*" in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Portland, Oregon, USA, pp. 309–319 (2011), 2011.

[3]Jindal, N., Liu, B. and Lim, E.-P. "*Finding Unusual Review Patterns Using Unexpected Rules*" CIKM (2010)., 2010 92.

[4]Mukherjee, A., Liu, B., Wang, J., Glance, N. and Jindal, N. 2011. Detecting Group Review Spam 2011

[5] Mukherjee, A., Liu, B., & Glance, N. "*Spotting fake reviewer groups in consumer reviews*" in Proceedings of the ACM international conference on world wide web (pp. 191–200). ACM., 2012

[6] Mukherjee, A., Venkataraman, V., Liu, B., & Glance, N." *Fake Review Detection: Classification and Analysis of Real and Pseudo Reviews*", UIC-CS-03- 2013. Technical Report.

[7] B. Liu, "*Sentiment analysis and opinion mining,"* Synthesis lectures onhuman language technologies, vol. 5, no. 1, 2012, pp. 1–167.

[8]RAYMOND Y. K. LAU, S. Y. LIAO, RON CHI WAI KWOK, KAIQUAN XU, YUNQING XIA, YUEFENG LI," *TextMining and Probabilistic Language Modeling for Online Review Spam Detection* "ACM Trans. Manag. Inform. Syst. 2, 4, Article 25 (December 2011)

[9] Yoo and Gretzel "Comparison of deceptive and truthful reviews" (2009)

[10]Xie, S.,Wang, G., Lin, S., & Yu, P. S" *Review spam detection via temporal pattern discovery",* in Proceedings of the ACM international conference on knowledge discovery and data mining (pp. 823–831). ACM., 2012.

[11] Fei, G., Mukherjee, A., Liu, B., Hsu, M., Castellanos, M., & Ghosh, R." *Exploiting burstiness in reviews for review spammer detection",i*n Proceedings of the ICWSM. Citeseer., 2013

[12] Miss Dipti S.Charjan , Prof. Mukesh A.Pund " *Pattern Discovery For Text Mining Using Pattern Taxonomy"* (IJETT) Volume 4 Issue 10- October 2013.

[13]Lin, Y., Zhu, T., Wang, X., Zhang, J., & Zhou, A. "*Towards online antiopinion spam: Spotting fake reviews from the review sequence",i*n 2014 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM) (pp. 261–264). IEEE.

[14] Heydari, A., Tavakoli, M., Salim, N."*Detection of fake opinions using time series*" Expert Systems with Applications, 58, 83-92, 2016

[15] Ye, J., Kumar, S., Akoglu, L"*Temporal Opinion Spam Detection*" Multivariate Indicative Signals, 2016.

[16] Liu, Pan, et al. "*Identifying Indicators of Fake Reviews Based on Spammer's Behavior Features.*" Software Quality, Reliability and Security Companion (QRS-C), 2017 IEEE International Conference on. IEEE, 2017.

[17] SP.Rajamohana, Dr.K.Umamaheswari, M.Dharani, R.Vedackshya. "*Survey of review spam detection using machine learning techniques*." 2017/978-1-5090-5778-8/17.

[18] Mr.Akshat Uike, Mr Ram Deshmukh, Dr.S R.Gupta, Dr. S.W.Ahmad4 ,"*Improved text mining techniques for spam review detection*" (IJIIRD), Vol. 03 Issue 01 2019

