



IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 3, March 2022

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.165



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

Survey on Opinion Mining in Text Review

R Aishwrya, Prof. Prashanth Kumar K N, Prof. Suma L

M.Tech Student, Dept. of Computer Science & Engineering, Bangalore Institute of Technology, Bengaluru, India
Assistant Professor, Dept. of Computer Science & Engineering, Bangalore Institute of Technology, Bengaluru, India
Assistant Professor, Dept. of Computer Science & Engineering, Bangalore Institute of Technology, Bengaluru, India

ABSTRACT: Customer reviews on the internet have now become valuable criteria for consumers before purchasing anything. Authentic product reviews significantly effects the future sales of the product. In our paper, we thoroughly investigate the peculiarities of reviewers' behavior based on their uploaded reviews. The Fake review identification and removal from a dataset using various Natural Language Processing (NLP) algorithms, on the other hand, is essential in multiple ways. For determining whether a review is good or bad, two separate Machine Learning (ML) models are used to train the review dataset. The authentic critic is determined using the crowd sourcing concept's cumulative score. Product reviews are becoming gradually more important in the E-commerce business and also in many other platforms. Before buying a product, the company's items were trusted. Here, we should address the issue of f fake review so that significant e-commerce companies like Amazon, Flipkart and others can solve the issue and remove the fraudulent reviewers and spammers, preventing people from losing trust in online buying sites. The "amazon Yelp dataset" is utilized in training models in our study, and its limited dataset is being used for training on a limited scale and also be expanded to attain maximum flexibility and accuracy.

KEYWORDS: opinion mining, sentiment analysis, text mining

I. INTRODUCTION

People are purchasing nearly anything online through e-commerce sites and having it delivered to their homes and offices, and the sophistication with consumer reviews posted online has risen at a quicker pace. As a result, while buying online, individuals are not allowed to directly check the product, therefore they rely heavily on ratings from other purchasers, which must be kept as honest as possible so that the customer really isn't repeatedly misled by false reviewers or spammers. Although this online purchasing is easy but also time-consuming: wherein we have to go through each review and identify whether it's a misleading or false rating. That's why all this should be carried out in an organized manner to know what is the root cause of this issue of fake reviews. However this issue can get fixed via developing a machine learning model that handles the review area and meanwhile can identify if a review is legitimate or fraudulent. It's also worth noting that spammers who haven't used the product could be spotted in this manner.

The spammer reviews and the use of several customer ids could also be utilized in falsely filtering the ratings in order to achieve a high product rating. It could be refined via looking for terms including "awesome," "very good," and "great," among others. Some people prefer to exaggerate the goods or seek to imitate authentic reviews by repeating similar phrases over and over in attempt to leave an impression upon that purchaser. As a result, the problem caused by spam filtering necessitates training a vast amount of and also be efficient by having extra specific understanding like making sarcastic phrases being used among customers to express their dissatisfaction with the goods or services; for example, few times even though the item is alright and if the delivery or packaging is not good, will hamper review categorization. Rather than just of falsely classifying a bad rating as in sentiment analysis, we use an NLP method is being utilized in attempt of deleting unnecessary or old customer reviews.

II. LITERATURE SURVEY

In the current time, people are more exposed to online shopping which is more comfortable and easier for them. They don't have to spend time in commuting; they can just scroll down their phones on various e-commerce apps and buy items based on the customer ratings or reviews on that particular product. However not all the reviews will be genuine, there will be fake and real ones. In attempt to identify which review is fake one or not, it will take a lot of time. In paper [1], Carloine El Fiorenza et al., monitor the fake rating identification using Opinion mining algorithm and by creating a dictionary for sentiment analysis. Their aim is to provide genuine user satisfaction by monitoring the reviews based on the sentiment score and removing the spam ones by sentiment dictionary.

A review can be fake positive or fake negative, in order to identifying we use opinion mining or sentiment analysis where in the text analysis is done based on the emotions of the opinion. In paper [2], Elshrif Elmurngi et al., compares five different Supervised Machine Learning (ML) algorithms in attempt to identify fake movie reviews that are both fake positive and fake negative using two different datasets and later conclude that Support Vector Machine algorithm gives more accuracy compare to other four algorithms such as Decision Tree, K-Star, K-Nearest Neighbour and Naive Bayes. Here sentiment analysis is done on the document level, where in the document is classified as opinion document which can be a negative opinion or a positive one.

Identification of fake reviews is done using Supervised and Unsupervised Machine Learning techniques. In paper [3], Aishwarya M. Kashid et al., identify fake or spam reviews based on the analysis of the review-centric features and the main review. Review spam identification methods are is done using data-driven methodologies that take into account a variety of factors connected with reviews, customers that are reviewers, and the social media structure that is utilized in classifying the reviews based on their authenticity. They conclude that the Unsupervised ML is less efficient but it doesn't need labeled datasets whereas the Supervised ML is more efficient in detecting the fake reviews based on small dataset.

The spam or fake review system is becoming more and more nowadays which hampers the sales of the product and also the business of the company. In paper [4], Anusha Sinha et al., offer an approach for identifying these spam ratings using Sentiment analysis or Opinion Mining. Wherein with the help of the opinion of the user, we try to make decide using decision tree algorithm based on the text analysis and also by using spam dictionary where we can filter the spam words. The software helps in blocking the IP address of the spammer manually helping in removal of spam reviews. The main issue of polarity is sentiment classification is achieved.

In paper [5], Ahlam Alrehili et al., recommended an Ensemble ML technique where the analysis is done based on the text and identify whether it is negative or positive. They used the Voting ensemble method, along with different algorithm such as SVM, Bagging and Boosting, Random forest, and Naive Bayes are combined classifiers. The Weka tool is against these classifiers in various scenarios like unigram, bigram, trigram without and with the stop of removal word. They conclude that the random forest gave the highest yield in scenario the unigram but the voting ensemble method outperformed in all the scenarios.

In paper [6], Rodrigo Barbado et al., develop a Fake Framework Feature called F3 used to characterize and arrange the features for spam review identification. The main goal is o identify the spam reviews in the online consumer electronic space conducted among four cities in USA. The F3 collects information that both user centric and review centric. They conclude that the Ada Boost classifiers and Radom forest gives the high F-score in spam review identification among electronic product reviews and the user centric F3 method gives better outcomes compare to review centric model.

In paper [7], P. Kaladevi et al., developed an Integrated CNN (Convolutional Neural Network) and LSTM (Long Term Short Memory) Recurrent NN based like Deep NN Opinion mining. The ICNN-LSTM-DNN method is utilized in the huge social media data for sentiment analysis. It uses the twitter posts tweets that were posted during the 2019 Election time in India as the data for sentiment analysis or opinion mining. The recommended method automatically separates the user facts from the opinion in real time tweets. This method outperformed the existing techniques in terms of F-score, accuracy, recall and precision. However the disadvantage of this is on identifying the difference between the levels of collision among every metric used for emphasizing an emotion of interest.

In paper [8], Yu Wang et al., try to analyze whether the culture differences has to do something the online customer reviews. They mined the product feature and their related opinions of online customer reviews for same product from Amazon China and USA websites. They conclude that the Chinese are more into the product aesthetics where in the Americans are more into product features and its usability. Also the Americans post more negative reviews on the product features compare to Chinese. Hence this shows how different cultural background helps to analyse the online customer reviews.

In paper [9], Piyush Jain et al., suggest an Opinion Mining method for identifying spam reviews that uses tracking of IP Address, Location, Ontology, and negative words dictionary using the Naive Bayes Supervised ML algorithm, and identifies if its brand based review or product based, and also detects if reviews are given by same account. Firstly the spam reviews are identified in the existing dataset, and then a new dataset is formed that contains spam filtered reviews. The spam filtered data is used for giving more accuracy and precision in identifying the fake reviews.

In paper [10], Amit Sawant et al., recommend a system to monitor and delete the deceptive reviews on items to provide more genuine reviews. Firstly, it understands the variation in the ratings of the product then identifies the IP Address to see if the customer is genuine or a spammer. Lastly uses a sentiment analysis to identify if the review is positive or negative. This analysis provides overall understanding of the product and also maintains the sales of the products with very little spam reviews. This system is easy to use and is effective and also simple to implement.

In paper [11], Abhishek Punde et al., develop a system that examine and deletes the spam reviews using ML techniques. The Sentiment Analysis is done on the proper authentic reviews in order to categorize if it's positive or negative. Also bag of words is used to label the words based on the opinion. The sentiment analysis uses two algorithms namely: Native Bayes and Decision Tree. They conclude that the Opinion mining or sentimental analysis is key aspect in deciding the sales of the product. It emphasizes on text mining and also data recovery, and feature weighting. The Decision tree classifier is slightly more accurate than the Native Bayes classifier.

In Paper [12], G. M. Shahariar et al., suggest a system that deals with both unlabelled and labelled data and also the deep learning methods Multi-Layer Perceptron (MLP), CNN and a variant of RNN which is LSTM. Also employ conventional ML methods like Naive Bayes (NB), SVM and K-nearest. They conclude that deep learning methods provide solution for classification of a problem completely. The conventional ML methods require less training datasets compared to Deep learning methods for fake review identification. The conventional SVM and NB methods give less accuracy along with the Word2Vec method. The proposed system MLP, CNN and LSTM gives better accuracy than the existing techniques.

In paper [13], Abdul Rafay et al., they centre around reviews for a wide range of hotel business and to have recommended a sentiment analysis and opinion mining framework to take out the classifying on organization reviews. to perform robust outcomes both parallel and multiclass classifying are utilized by utilizing a huge and rich text reviews dataset given by yelp set of data task round. The CLSTM text classification method has got better scores in both binary and multi-level classification compared to MNB text classification method.

In paper [14], Aagam Shah et al., customer survey investigation gains us with the specific comprehension of the client's responses towards to the item. Throughout this paper, we have now recommended totally unique methodology by performing unique level sentimental examination of client survey by n-gram classification and pos tagging. This classification would then be utilized as entropy for ML algorithm. This paper influences upon the suggested method with promising outcomes and further developed accuracy by assessing the information with the assistance of two algorithms, MaxEnt model and NB classifier, in the wake of breaking down couple of algorithms including SVM and random forest.

In paper [15], Yuanyuan Wu et al., developed a precursor result intercession calculated system to foster an underlying exploration plan for examining fake audits. In light of an audit of the surviving literary works on this problem, we distinguish 20 future exploration research questions and recommend 18 recommendations. Eminently, research on falsify audits is regularly restricted by absence of excellent datasets. To reduce this issue, we exhaustively order and sum up the current fake audits related public datasets. We finish up by introducing the conceptual framework and functional implications of the present examination.

In paper [16], Zhi-Ping Fan et al., develop an information fusion in attempt to rank the items based on the online customer ratings and audit the current investigations on cycles and techniques for data mixture for every stage. Moreover, we momentarily study the current examination on data combination in view of web audits in several fields. At long last, we sum up the primary finishes of this document and indicate the long run exploration heading. The procedure of the online customer reviews and ratings involves: item feature extraction, opinion mining and ranking the items.

In paper [17], Mansi Pakhale et al., suggested a framework which assists the clients with discovering the malicious audits or information. A system is developed for identifying the false ratings, which uses Opinion mining, various ML methods and Deep learning methods and also uses spam dictionary. Utilization of calculations such as NB, decision tree and SVM makes it feasible to arrange the information and use it for the spam recognition with a specific precision. Likewise, deep learning has been a significant player as its calculations makes it conceivable to get more noteworthy exactness and its neural network algorithms and methods helps in accuracy. Their software has an extra element for identifying negative comments on the YouTube videos.



Despite the highlights of extraction cycle of the surveys. In paper [18], Ahmed M, Elmogy et al., applied a few elements, designing to extricate, different practices of the commentators. The study evaluates the presentation of a few trials accomplished on a genuine yelp dataset of eateries surveys either with or without features extricated from client’s practices. In the two cases, we think about the exhibition of a few classifiers; KNN, NB, SVM, logistic regression and random backwoods. Likewise, unique language models of n-gram specifically bi-gram and tri-gram are taken into contemplations during the assessments.

In paper [19], Joni Salminen et al., recommended system that uses language models such as ULMFiT and GTP-2 to create spam reviews based on the Amazon dataset. GTP-2 being the better one was used to generate a dataset for classification of identifying spam reviews. They found fake survey data is an issue for specialists, web based business locales, and firms occupied with online company. Our outcomes demonstrate that current text age strategies yield fake surveys that show up so practical that it is trying for a human to distinguish them. Luckily, ML classifiers improve in such manner, with practically ideal precision in distinguishing audits created by different machines.

In paper [20], W.M. Wang et al., suggest a Kansei text mining methodology which consolidates text mining and Kansei designing ways to deal with consequently extricate and sum up item includes and their relating emotional reactions in view of online item portrayals and shopper surveys. Clients can effectively and ideal survey the emotional parts of the items. To assess the viability of the proposed approach, tests have been led based on open information from amazon.com. The outcomes showed that the proposed approach can really recognize the emotional data as far as element full of feeling conclusions. Furthermore, we have fostered a model framework that envisions item includes, emotional qualities, full of feeling watchwords, and their connections. The proposed approach helps customers settling on buy choices, yet in addition helps makers understanding their items and rivals' items, which could give bits of knowledge into their item improvement.

III. LITERATURE SURVEY DETAILS

Pap er no.	Short notes	Advantages	Disadvantages
[1]	There might be some difference in the procedure that is taken while buying the product so the reviews are directly related to the sales of the product and thus it necessary for the online websites to spot fake reviews as it’s their own reputation that comes into consideration as well, so a Fake Review Detection is used to spot any fraudulent going on because it’s not possible for them to verify every product and sale manually so a program comes into the picture that tries to detect any pattern in the reviews given by the customers	The project will detect the fake reviews by deploying opinion mining algorithms and creating a word dictionary.	To implement the system and check performance by applying proposed approach to various benchmark data sets.
[2]	This paper aims to classify movie reviews into groups of positive or negative polarity by using 5 machine learning algorithms. In this study, we analyze online movie reviews using SA methods to detect fake reviews. SA and text classification methods are applied to a dataset of movie reviews which are V1.0 and V2.0. The results show that the SVM algorithm outperforms other algorithms by reaching highest accuracy not only in	Using the accuracy analysis for these five techniques, SVM algorithm is the most accurate for correctly classifying the reviews in movie reviews datasets, i.e., V2.0 and V1.0.	Extending the study to use other datasets such as Amazon dataset or eBay dataset and use different feature selection methods. Also, sentiment classification algorithms are applied to detect fake reviews using various tools.



	text classification, but also in detecting fake reviews.		
[3]	To spread of misinformation, to promote or to damage certain business, various opinion spam's are done either to mislead the human readers or the sentiment analysis or opinion mining systems which are automated. The analysis of the main review and the reviewer-centric features are proposed to detect the fake reviews by using supervised machine learning approaches.	Supervised classifiers are in general more effective, and usually employ review and reviewer-centric features.	Unsupervised solutions are in general less effective to employ review and reviewer-centric features.
[4]	In this paper, we propose a framework to detect fake product reviews or spam reviews by using Opinion Mining. In sentiment analysis, we try to figure out the opinion of a customer through a piece of text. With the help of the decision tree check if the review is related to particular product. We use Spam dictionary to identify the spam words in the reviews.	The admin has to manually block the IP of the spammer account by identifying its pattern.	There is no automatic blocking of IP of spammer.
[5]	On a set of customer reviews collected from Amazon the sentimental analysis approach is conducted. Classifying each review into one of these classes: positive review or negative review by using ensemble machine learning method. The ensemble machine learning method used Voting which combined five classifiers: Naive Bayes, Support Vector Machines (SVMs), Random forest, Bagging and Boosting. We test six different scenarios to evaluate our proposed model against the five classifiers. The scenarios are using unigram (with/without) stop of word removal, bigram (with/without) stop of word removal and trigram (with/without) stop of word removal. The result shows the random forest technique give the highest accuracy which equals to 89.87% in the case of using unigram (with) a stop of word .	With stop word removal the accuracy increased in all the three scenarios unigram, bigram, and trigram. Moreover, the unigram provides the best result among Ngram scenarios.	If the stop words has not been removed this lead to the decreased accuracy rate in all three scenarios except trigram.



[6]	This article proposes a feature framework for detecting fake reviews that has been evaluated in the consumer electronics domain. Construction of a dataset for classifying fake reviews in the consumer electronics domain in four different cities based on scraping techniques and Feature framework for fake review detection.	Reached an 82% F-Score on the classification task and the Ada Boost classifier has been proven to be the best one by statistical means according to the Friedman test.	This framework can also be used in other tasks, such as toxic user activity detection or fake news detection, which could involve complex semantic analysis.
[7]	An Integrated Convolutional ssNeural Network and Long Short Term Memory (LSTM) Recurrent Neural Network-based Deep Neural Networks-based Sentiment Analysis Methodology (ICNN-LSTM-DNN) was proposed over the big social data for opinion mining. This proposed ICNN-LSTM-DNN-based sentiment analysis approach is an adaptable sentimental analysis mechanism which is capable of investigating social media posts and extracts user’s facts and opinion in real-time.	This proposed ICNN-LSTM-DNN-based sentiment analysis scheme was presented as an important adaptive technique used for the objective of extracting opinions of the people about a specific subject that purely depends on the social media contents.	This proposed model has inability in differentiating the degree of impact between each metric used for accentuating a feeling of interest.
[8]	In this study, we investigate whether consumers with different cultures concentrate on different product features in online consumer product reviews and show different opinions toward individual product features of the same products. End, we extract product features and their associated opinions from online consumer reviews of the same products available at Amazon websites for U.S. and Chinese consumers.	Product manufacturers should understand unique preferences and needs of certain products and their features desired by consumers with different cultures so that they can better customize or improve their products to suit the needs of different markets.	Data collected from two OCR platforms in two countries which are very different in terms of Hofstede’s cultural dimensions.
[9]	Opinion Spam detection is an exhausting and hard problem as there are many faux or fake reviews that have been created by organizations or by the people for various purposes. They write fake reviews to mislead readers or automated detection system by promoting or demoting target products to promote them or to degrade their reputations. The proposed technique includes Ontology, Geo location and IP address tracking, Spam words Dictionary using Naïve Bayes, Brand only review detection and tracking account used.	New algorithm is proposed that detects spam reviews more precisely and performs opinion mining using spam filtered data.	Manually the admin has to find out the fake reviews based on IP address of the user.



[10]	In recent times it is seen that the companies tend to provide reviews to its own product in order to increase its sales, and on other hand there are some users who tend to degrade the product by giving fake negative reviews. Hence, a system which can detect such types of fake reviews be it positive or negative is needed. These reviews are also important in terms of the future progress of the ecommerce website.	This system is its efficiency and the ease to implement and use it. The admin can easily track the analyzing process of reviews in the systems.	This is a general model apart from this there is a lot of techniques which can be used.
[11]	In some review sites some great audits are included by the item organization individuals itself, so as to make so as to deliver false positive item reviews. To eliminate this type of fake product review, we will create a system that finds out the fake reviews and eliminates all the fake reviews by using machine learning. We also remove the reviews that are flood by a marketing agency in order to boost up the ratings of a particular product. Finally Sentiment analysis is done and Bag-of-words to label individual words according to their sentiment.	Sentiment Analysis not only encompasses concepts of text mining but also the concepts of information retrieval. Major challenges in Sentiment Analysis includes feature weighting which plays a crucial role for good classification.	The work can be further extended to emerging areas like Mobile learning and investigation with soft computing techniques like neural network.
[12]	To detect any deceptive text reviews. To achieve the working with both labeled and unlabeled data and proposed deep learning methods for spam review detection which includes Convolutional Neural Network, Multi-Layer Perceptron, a variant of Recurrent Neural Network (RNN) that is Long Short-Term Memory. We have also applied some traditional machine learning classifiers to detect spam reviews and performance comparison for both traditional and deep learning classifiers.	Deep learning algorithms such as Word2Vec is also used in order to obtain better vector representations for words and improve the accuracy of classifiers trained with traditional machine learning algorithms.	Traditional machine learning algorithms such as SVM and NB reach a certain threshold where adding more training data does not improve their accuracy. Drawback of this problem using deep learning is on small dataset which leads to over fitting problem.
[13]	We focus on reviews for all kinds of restaurants business and have proposed a sentiment analysis and opinion mining model to perform the classification on business reviews. In order to achieve robust results both binary and multilabel classification are used by using a large and rich text reviews dataset provided by Yelp Dataset. Compare the results of a machine learning based algorithm "Multinomial Naive Bayes" and deep learning algorithm "convolution Long Short Term Memory" (CLSTM) with word2vec and Global Vector (Glove). After analyzing the performance of	Best classifier C-LSTM has obtained overall good scores for both binary and multi-class classification with a word to vectors encoding technique.	Although the accuracy for MNB in binary classification seems to be good but it has a biased result as it is learning a single class (positive) and neglecting the other (negative) class.



	each model with different metrics, it has been observed that the best model for classifying the review ratings is CLSTM.		
[14]	Proposed a unique approach by performing abstract-level sentimental analysis of user review by n-gram classification and POS tagging. This classification is then used as entropy for machine learning algorithm. This paper leverages upon the proposed methodology with promising outcomes and improved accuracy by evaluating the data with the help of two algorithms, MaxEnt model and Naïve Bayes classifier, after analyzing few algorithms including SVM and random forest.	It is clearly observed that the combined methodology of performing MaxEnt and Naïve Bayes together with the help of n-gram classification and POS tagging provides estimation of ratings with highest accuracy.	Due to the imbalance in the nature of user reviews and the correspondent ratings, there is a conflict between prediction and actual ratings.
[15]	This study proposes an antecedent–consequence–intervention conceptual framework to develop an initial research agenda for investigating fake reviews. Based on a review of the extant literature on this issue, we identify 20 future research questions and suggest 18 propositions. We conclude by presenting the theoretical and practical implications of the current research.	The ACI framework is used as an initial research agenda and serves as a roadmap to guide further study of fake reviews.	Apart from contributing to the overall understanding of the extant research and the corresponding limitations, we analyze the data sources and research methods within this field of research and, accordingly, highlight promising research directions
[16]	We review the existing studies on processes and methods of information fusion for each stage. Furthermore, we briefly review the existing research on information fusion based on online reviews in other fields. Finally, we summarize the main conclusions of this paper and point out the future research direction.	The studies on methods and techniques for information fusion based on online reviews should not only focus on the issues of ranking products based on online reviews, but also need to conduct in-depth studies on the other fields of information fusion, s	Most of the existing studies assume that content of online reviews is real, but there may be some false information and useless information in numerous online review data.
[17]	We propose a framework to detect fake product or content reviews or spam reviews by using Opinion Mining and different Machine Learning and Deep learning algorithms. The Opinion mining is also known as Sentiment Analysis. Here we use Spam dictionary to identify the spam words in the reviews. In Text Mining we apply several algorithms from Machine learning and detect the reviews as the spam and the real ones.	It monitors the spam review made on any product/video. And user can be sure about the products availability on that application and reviews too.	The admin has to remove the spam reviews manually from the system, in the future work it could be automated. • Only the verified purchaser can be allowed to make a review a certain product.



[18]	The paper compares the performance of several experiments done on a real Yelp dataset of restaurants reviews with and without features extracted from user’s behaviors. In both cases, we compare the performance of several classifiers; KNN, Naive Bayes (NB), SVM, Logistic Regression and Random forest. Also, different language models of n-gram in particular bi-gram and tri-gram are taken into considerations during the evaluations.	Not all reviewers behavioral features have been taken into consideration in the current work	Consider including other behavioral features such as features that depend on the frequent times the reviewers do the reviews, the time reviewers take to complete reviews, and how frequent they are submitting positive or negative reviews.
[19]	we address the creation and detection of fake reviews. First, we experiment with two language models, ULMFiT and GPT-2, to generate fake product reviews based on an Amazon e-commerce dataset. Using the better model, GPT-2, we create a dataset for a classification task of fake review detection. We show that a machine classifier can accomplish this goal near-perfectly, whereas human raters exhibit significantly lower accuracy and agreement than the tested algorithms. The model was also effective on detected human generated fake reviews. The results imply that, while fake review detection is challenging for humans, “machines can fight machines” in the task of detecting fake reviews. Our findings have implications for consumer protection, defense of firms from unfair competition, and responsibility of review platforms.	Current text generation methods yield fake reviews that appear so realistic that it is challenging for a human to detect them. Fortunately, machine learning classifiers do much better in this regard, with almost perfect accuracy in detecting reviews generated by other machines.	Experimenting with more datasets and platforms is required.
[20]	In this paper, we propose a Kansei text mining approach which incorporates text mining and Kansei engineering approaches to automatically extract and summarize product features and their corresponding affective responses based on online product descriptions and consumer reviews. Users can efficiently and timely review the affective aspects of the products. In order to evaluate the effectiveness of the proposed approach, experiments have been conducted on the basis of public data from Amazon.com. The results showed that the proposed approach can effectively identify the affective information in terms of feature–affective opinions. In addition, we have developed a prototype system	This approach not only helps consumers making purchase decisions, but also helps manufacturers understanding their products and competitors’ products, which might provide insights into their product development	The proposed approach can only handle English text. It does not support other languages or formats, such as images, audio, and video. Some special characters, such as emoji, are useful to express consumer emotions. They are not included in this study.

	that visualizes product features, affective attributes, affective keywords, and their relationships.		
--	--	--	--

Table 1 I. Literature Survey Details:**IV. CONCLUSION**

In our paper, we examine on the outcomes by comparing the two models such as: Naive Bayes model and Random forest model which are designed to demonstrate performance for "Amazon's yelp" dataset plus their significance to implement both these models into real world applications explored. As a result, the Naive Bayes method was outperformed in a wide margin by the Random forest algorithm. The issue of fake review identification is satisfactorily resolved and provides a decent understanding of its legality and necessity; our goal is to use a technique to complete the work of identifying the false reviews and deletion of it. In the upcoming work, we can test the hybrid models and novel models for the identification of false reviews. The work can be fastened up by the execution process via employing NVIDIA Graphics Processor and Google co-lab.

REFERENCES

1. Carloine El Fiorenza, Aditya Singh Kashyap, Kartikey Chauhan, Kishan Mokaria, Aashutosh Chandra , "Fake Product Review Monitoring and Removal for Genuine Online Reviews", Journal of Network Communications and Emerging Technologies (JNCET), Apr 2018.
2. Elshrif Elmurngi, Abdelouahed Gherbi, "Detecting Fake Reviews through Sentiment Analysis Using Machine Learning Techniques", DATA ANALYTICS 2017: The Sixth International Conference on Data Analytics, Jun 2018.
3. Aishwarya M. Kashid, Ankita K. Lalwani, Samiksha S. Gaikwad, Rajal A. Patil et al., "Fake Review Detection System Using Machine Learning", 2018 International Journal of Research in Engineering, Science and Management, 2019.
4. Anusha Sinha, Nishant Arora, Shipra Singh et al., "Fake Product Review Monitoring Using Opinion Mining", International Journal of Pure and Applied Mathematics, Jan 2018
5. Ahlam Alrehili, Kholood Albalawi, "Sentiment Analysis of Customer Reviews Using Ensemble Method", 2019 International Conference on Computer and Information Sciences (ICCIS), 2019, Sakaka, Saudi Arabia.
6. Rodrigo Barbado, Oscar Araque, Carlos A. Iglesias, "A framework for fake review detection in online consumer electronics retailers", 2019 Information Processing & Management, 1234-1244, 2019.
7. P. Kaladevi, K. Thyagarajah, "Integrated CNN- and LSTM-DNN-based sentiment analysis over big social data for opinion mining", Behaviour & Information Technology, 09 Dec 2019.
8. Yu Wang, Zhigang Wang, Dongsong Zhang, Ru Zhan, "Discovering Cultural Differences in Online Consumer Product Reviews", Journal of Electronic Commerce Research, 2019.
9. Piyush Jain, Karan Chheda, Mihir Jain, Prachiti Lade, "Fake Product Review Monitoring System", International Journal of Trend in Scientific Research and Development (IJTSRD), 105-109, Apr 2019.
10. Amit Sawant, Shruti Desai, Shraddha Bhang, Akash Pandey, "Fake Product Review Monitoring and Removal for Proper Ratings", 2019 International Journal of Engineering Research & Technology (IJERT) Volume 08, Issue 03, 10 Mar 2019.
11. Abhishek Punde, Sanchit Ramteke, Shailesh Shinde, Shilpa Kolte, "Fake Product Review Monitoring & Removal and Sentiment Analysis of Genuine Reviews", International Journal of Engineering and Management Research (IJEMR), Apr 2019.
12. G. M. Shahariar, Swapnil Biswas, Faiza Oma et al., "Spam Review Detection Using Deep Learning", 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Dec 2019, Vancouver, BC, Canada.
13. Abdul Rafay, M. Suleman, Affan Alim, "Robust Review Rating Prediction Model based on Machine and Deep Learning: Yelp Dataset", 2020 International Conference on Emerging Trends in Smart Technologies (ICETST), 30 Apr 2020, Karachi, Pakistan.
14. Aagam Shah, Komal Kothari, Umang Thakkar, Satvik Khara, "User Review Classification and Star Rating Prediction by Sentimental Analysis and Machine Learning Classifiers", Information and Communication Technology for Sustainable Development, Jun 2019.



15. Yuanyuan Wu, Eric W.T. Ngai, Pengkun Wu, Chong Wu, “Fake online reviews: Literature review, synthesis, and directions for future research”, Decision Support Systems, Volume 132, 2020.
16. Zhi-Ping Fan, Guang-ming Li, Yang Liu, “Processes and methods of information fusion for ranking products based on online reviews: An overview”, 2020 Information Fusion, 87-97, 01 Aug 2020.
17. Mansi Pakhale, Swati Pandey, Gauri Patil, “Spam Review Analysis and Detection System”, International Research Journal of Engineering and Technology (IRJET), Apr 2021.
18. Ahmed M. Elmogy¹, Usman Tariq, Ammar Mohammed et al., “Fake Reviews Detection using Supervised Machine Learning”, International Journal of Advanced Computer Science and Applications (IJACSA), Jan 2021.
19. Joni Salminen, Chandrashekar Kandpal, Ahmed Mohamed Kamel et al., Creating and detecting fake reviews of online products”, Journal of Retailing and Consumer Services, Jan 2022.
20. W.M. Wang, Z. Li, Z.G. Tian et al., “Extracting and summarizing affective features and responses from online product descriptions and reviews: A Kansei text mining approach”, Engineering Applications of Artificial Intelligence, 149-162, Aug 2018.



INNO  SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.165

 **doi**[®]
cross **ref**

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  ijircce@gmail.com



www.ijircce.com

Scan to save the contact details