A Study On Online Spam Review Detection Methods By Machine Learning Approach

Dr.Sudha Rajesh[1] , Dr. M. Mercy Theresa[2], Dr.J.Nithyashri[3], Dr. S. Jaanaa Rubavathy[4]

[1]Assistant Professor, BSAR Crescent Institute of Science and Technology
[2]Associate Professor, Prince shri Venkateswara Padmavati Engineering College,.
[3]Professor ,Karpaga Vinayaga College of Engineering & Technology,
[4]Associate Professor, Saveetha School of Engineering-Saveetha Institute of Medical and Technical Sciences,
[1]drsudharajesh84@gmail.com, [2]mercyjesudossa@gmail.com,[3] drjnithyashri@gmail.com,
[4]jaanaaruba@gmail.com.

Abstract: Online reviews for products, movies, shopping, etc. have become the major purchase decision of customers to make before buying the product. These reviews are kind of public opinion about the product which they bought and this may help another person who is willing to buy the product. The impact of reviews had made manufacturers, and retailers to be more concerned about making the product as a best one. Many retailers such as Reliance, Big Bazaar, More, etc are very concerned about online reviews and there are possibilities for these reviews to affect business either positively or negatively. Certain retailers are trying to create false reviews about the product through AI as a part of promoting the product. This process is termed the Opinion spam or opinion review spam here is where the reviewers manipulate a wrong review for selling the product for profit. Not all online reviews are trustworthy and truthful so in such cases, a model must be created for detecting online review spam. This research paper tries to highlight the online spam review detecting by machine learning approach using NLP (Natural Language Processing) where the extracted features from the text will be taken for further review. In this research main focus was shifted to the ML technique for detecting review spam, and classifying them. The main aim of the research is to provide a comprehensive ad strong ML approach for detecting review spam.

1. Introduction

(NAIK et al., 2019)Spam is known as the electronic message which includes various types of message services through the digital delivery system and through broadcast media. This system helps to send messages in larger numbers to the recipients on the internet.

(Tang et al., 2019) Nowadays spam on the internet is mainly through advertisements from the business sector or individuals business firm who wanted to promote their products send spam messages to the recipients on the internet through social media platforms like Twitter, Instagram, YouTube, Facebook, Websites, or Text messages. Spam messages are the ones where the sender's details are masked through which the original identity of the sender will not be known to the receivers. There is a difference between electronic spam and marketing spam so an advanced method is needed to identify and stop spam messages from an unauthorized person otherwise it will destroy the benefits of internet usage. Because anyone with internet can write a review about products which they mislead the consumers who are preferring to buy the products based on the review. Review for products on the internet has created a huge momentum for customers and it acts as a deciding factor for buying products.

(Barushka & Hajek, 2019) Online reviews for products are done for improving and enhancing the business sector of an individual or wholesale business dealers. Online reviews for a product can be true only if the consumer purchases the product and write down his/her own reviews for other consumers to buy the product. So there are possibilities for online review to create a blind trust about the product. It is sometimes dangerous for both buyers and sellers. In recent times before buying the product buyers look for an online review to decide whether to buy the product or not. The fake online review will completely break the trust of the individual who purchased the product after reading online reviews.

2. Literature Review

(Tang et al., 2019) proposed a framework for spam detection in which review spam where classified in 3 groups as
1) Untruthful Reviews,
2) reviews based on brand,
3) non –reviews with not-related content about the product.

Detecting of 1st type of review spam is considered the most difficult challenge and sometimes it seems to be impossible to identify the difference between fake and real reviews.

(Barushka & Hajek, 2019) proposed a framework by considering various online reviews as test samples to detect the real and fake reviews. Some online reviews considered in this research paper as follows

Review 1- The Big Hotel this is one of the great hotel to stay in with multiple facilities of the studio, gym, spa, swimming pool, etc. the studio wa too big, and the experience of watching a movie in this studio was highly different. The kitchen is having all types of utitensils such as microwave oven, dishwasher, fridge, freezer. The size of the bathroom was big with bath-tub and with proper toiletries. The parking slot was very secured at an affordable price. Every morning they provided morning breakfast which was the best part. The view of the hotel was good. An excellent experience ever experienced.

Review 2- During my latest business trip, both me and my wife recently stayed at the Omni Chicago Hotel in Chicago, Illinois, at one of their Deluxe suites. Unfortunately, and I think I speak for both of us, we were not fully satisfied with the hotel. The hotel advertises luxury level accommodations, and while the rooms resemble what one can see in the pictures, the service is certainly sub-par. When one plans a stay at such an establishment, they expect a service that goes beyond having fresh towels in the bathroom when they check-in. First of all, the air-conditioning in the room seemed to require a new filter and when it was first turned on, the air coming out seemed musty. Second of all, the fitness center was only open until 10:30 pm. For people who like to exercise after dinner, this can certainly be a problem. Especially considering that it does not take much to have the fitness center available around the clock or until midnight. For these, as well as other reasons, I would not recommend this hotel , if one is looking for luxury accommodations

There is no clear evidence from the 2 reviews which one is real and which one is fake for the reader who is looking for online reviews about a hotel to stay while making a trip. So it is difficult to judge which review is real or fake. Almost 18 M reviews have been created on the travel website Yelp and Trip Advisor. But in recent times it was found nearly 200 M reviews were found on the mentioned travel website. In such a case, the Big Data Analysis technique is needed to address this kind of large product review spam that takes place on the internet.

(Khurshid et al., 2018) proposed a framework by analyzing big data review spam using the Machine Learning method and found that the big data method was inefficient to deal with data of bigger size which seems to be difficult for using this type of algorithm for spam detection approach. A simplistic example for opinion mining is illustrated in fig 1 as follows. In opinion mining feature extraction text has been classified into 2 types as
1) behavioral feature set
2) Textural feature set
These 2 methods can be used for extracting features. Then the obtained data will be balanced by balancing them into various sub-sections by creating a unique balancing set. Then in the later stage, the obtained data from the balancing set will be sent for model training.
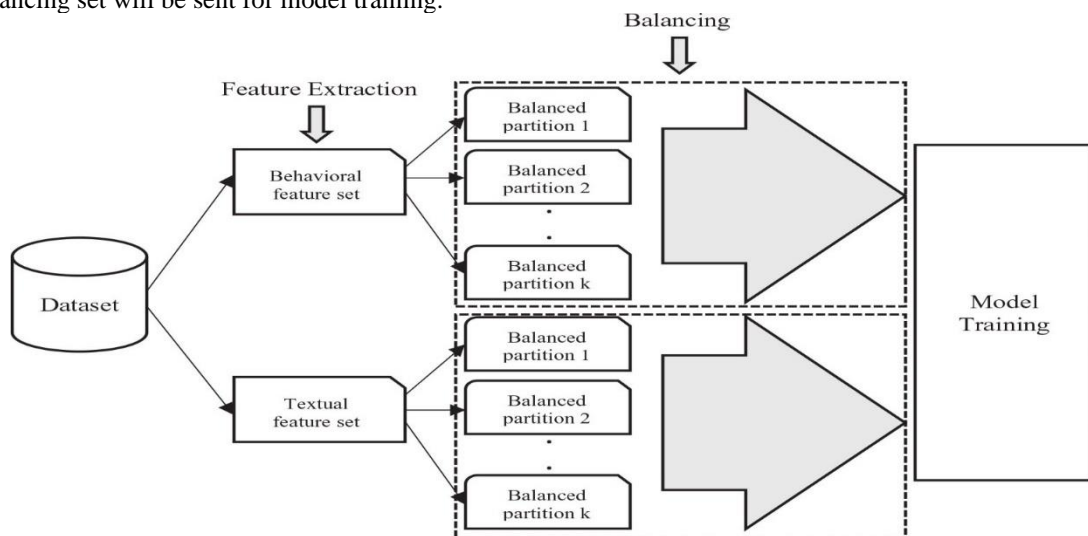

Fig.1 Opinion mining feature extraction for text

(Mohammadi & Mousavi, 2020) in opinion mining, the text will be considered based on sentiments as positive or negative polarity and the features of that textural passage will be analyzed. In the textual features extraction process a new set of the classifier will be developed for classifying the texts which have different opinions along

with mixed sentiment in it. In review spam detection opinion mining is one of the content mining processes which will not use features directly linked with the content. To describe more about the textural features of the review an NLP and text mining is needed.

(Shahariar et al., 2019) it was found that the ML-based review spam detection method is a little inefficient, they are more reliable than the manual detection process. One of the most known methods for text mining is usually done using a bag of words approach in which each and every individual word or small group of words can be used as features. But many previous studies indicated that this method is also inefficient for training the classifier with the capacity of detecting online review spam. So additionally feature engineering methods must be used for extracting features through which it is possible to easily detect online review spam. Many researchers suggest this feature engineering method for detecting review soam by using a variety of ML techniques.

(Wang et al., 2016)(Adike & Reddy, 2016) proposed a framework for review spam detection in which they utilized individual words for extracting features.
(Li et al., 2017) proposed a framework for spam detection using lexical features and syntactic features.
(Xue et al., 2019) proposed a framework utilizing review characteristics features along with unigram, and bigram frequencies.

(Duhan et al., 2017) the first step to be addressed in online review spam detection is the process of collecting data. Data is the foremost part of the ML approach because a wide range of reviews is available online. Collecting the appropriate data, labeling them according to its efficiencies are the most significant duty to be performed before sending the obtained data for training. But this process can be done by an alternative method using synthetic review spamming in which it only takes the true reviews by eliminating the fake one among a set of reviews.

(Etaiwi & Awajan, 2017) proposed a framework using a synthetic review spamming approach for creating a review spam dataset.
This research paper mainly focuses on the ML technique which will be used for detecting online review soam along with the concept of feature engineering and will also discuss the impact of the proposed features based on the performance and its spam detecting capacity. The advantages of supervised, semi-supervised and unsupervised have also been analyzed. The outcome of this research will be presented along with a comparative analysis study.

3.   Feature engineering for review spam detection

Feature engineering is one of the feature extraction method used for extracting features from data. This section mainly discusses the commonly used feature extracting method for review spam detection method. from the past study, it was found that there are many kinds of a feature extraction method for extracting a feature from a review. The most commonly known method was found to ve review text where a bag of wors method will be implemented in this method the feature will be considered based on individual words or a small group of words found in the review text.

(Mohammadi & Mousavi, 2020) only a few researchers have tried using the reviewer and product-centric features method by utilizing lexical, syntactical feature or features which describes reviewer's behavior. The features are usually categorized into 3 sections as
1)   Review centric features
2)   Reviewer centric features
3)   Product centric features

The first category is where the features are based on information found in a single review found on the internet. But the second and third categories are mostly reviews written by reviewers. The reviewers are either written by authors along with information about the one who wrote the review.

The features usually along with the mentioned category such as the bag of words method along with POS tags it or features et will be created based on features from review centric, product-centric, and reviewer centric.
(Etaiwi & Awajan, 2017; Jindal et al., 2020; Neisari, 2020) proposed a framework where most commonly used features such as the LIWC, and POS along with the bag of words method was incorporated to robust the bag of words approach rather than working as an individualistic approach.
(Barushka & Hajek, 2019) found that the abnormal behavioral features using reviewer centric feature performed very well when compared to the linguistic feature method using review centric features. The following subsections discuss and provide examples of some review centric and reviewer centric features.

3.1 Review centric features

the review centric features will be divided into various subsections as
1) Bag of the word method
2) Bag of words along with tern frequency features
3) The LIWC- (Linguistic Inquiry and Word Count)
4) The POS- (parts of speech)
5) Stylometric and Syntactic features

Then the review characters of review-centric features will be discussed for text with information about the review which has not been extracted.

3.1.1 Bag of words
In this method, either individual words or a small group of words from the text will be used as features. The feature type in this method is known to us as "n-gram". This feature is selected by taking n number of contiguous words from the given text such as either n can be 1 or 2 or 3 words selected from the given text.
1) n= 1 is called as unigram
2) n= 2 is called as bigram
3) n= 3 is called a trigram

(Dewang & Singh, 2018) (Hussain et al., 2019) proposed a framework utilizing the above-mentioned methods. (Crawford et al., 2015) proposed a framework using only n-gram features but this method was found to be inefficient for the supervised learning method and the training for the trainers was based on a synthetic false review. The unigram text feature extracted from sample reviews is tabulated in 1 and the sample review is also mentioned below as follows. The occurrence of words in the sample review will be denoted as 1 if the word exists in the reviews or else it will be denoted as 0 for otherwise.

1. Review-1: The hotel rooms were so great
2. Review-2: We had a great time at this hotel great stay
3. Review-3: The rooms service is bad

Table 1. Example of text features dataset structure, for reviews 1, 2 and 3

Review the hotel rooms were so great we had a time at this service is bad stay																
Review1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
Review2	0	1	0	0	0	1	1	1	1	1	1	1	0	0	0	1
Review3	1	0	1	0	0	0	0	0	0	0	0	0	1	1	1	0

3.1.2 Term frequency
Term frequency is almost similar to bag of the words feature. (Hussain et al., 2019) (Xie et al., 2012) proposed a framework using the term frequency feature. The term frequency feature method is shown in table form in tabulation 2. This term frequency is more about being concerned about the number of the frequency the word occurred in the sample review through which the number of times the word occurred in review can be identified as term frequency.
4. Review-4: The hotel rooms were so great,were very comfortable
5. Review -5:We had a great time at this hotel great stay
6.Review-6:The rooms service is bad so bad

Table 2. Example of text features frequencies dataset structure, for reviews 4, 5 and 6

Review the hotel rooms were so great comfort we had a time at this service is bad stay															
Review4	1	1	1	2	2	1	1	0	0	0	0	0	0	0	0
Review5	0	1	0	0	0	1	0	1	1	1	1	1	1	0	0
Review6	1	0	1	0	0	0	0	0	0	0	0	0	0	1	1

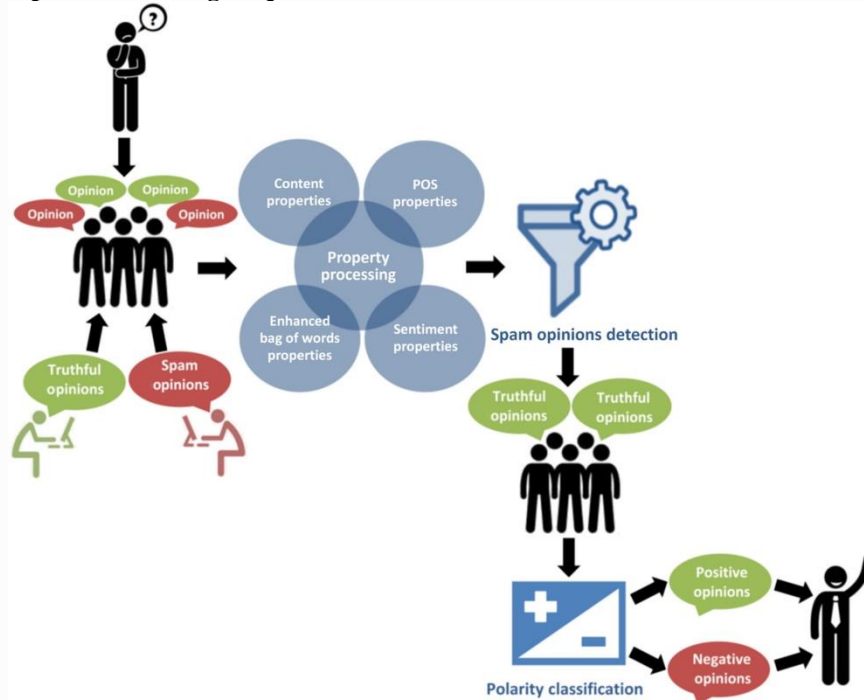**3.1.3 LIWC output and POS tag frequencies**

Fig 2. POS frequency Analysis approach

This is textural analysis method based on software in which the user they themselves can create their dictionary for analyzing the dimensions of languages based on their interest.

(Ghai et al., 2019) POS is based on tagging the word feature along with the part of the review which is based on content and definition from the review text found in the sentence.

(You et al., 2020) proposed a framework using this method and the outcome was better when compared to the bag of words method.

Table 3 is the outcome of the LIWC method applied on review 7. Personal texts in the reviews are the one which is about work, leisure activity, and home it is basically about the personal concerns of the review writer.

Formal texts are the ones about the psychological and linguistic process.

Table 4 is about POS tags for each and every word. Table 5 discuss the frequencies of words tagged from the sample review.

7. Review 7: *I like the hotel so much, the hotel rooms were so great, the room service was prompt, I will go* back to this hotel next year. I love it so much. I recommend this hotel to all of my friends.

8. Review 8: I_PRP like _VBP the_ DT hotel_ NN so_RB much_RB,_, The_DT htel_NN rooms_NNS were_VBD so_RB great_JJ,_, the _DT room_NN service_NN was_VBD prompt_JJ,_,I_PRP will_MD go_VB back_RB for _IN this_DT hotel_NN next_JJ year_NN._. I_PRP love_VBP it_PRP s0_RB much_RB._.I_PRP recommend_VBP this _DT hotel_ NN for_IN all_DT of_IN my_PRP$ friends_NNS._.

Table 3 LIWC results when applying Review7 text

LIWC Dimension	Your data	Personal texts	Formal texts
Self-references(I, me, my)	12.50	11.4	4.2
Social words(Mate, talk, they, child)	2.50	9.5	8.0
Positive emotions(Love, nice, sweet)	5.00	2.7	2.6
Negative emotions(Hurt, ugly, nasty)	0.00	2.6	1.6

Overall cognitive words(cause, know, ought)	0.00	7.8	5.4
Articles(a, an, the)	7.50	5.0	7.2
Big words(>6 letters)	7.50	13.1	19.6

Table 4 POS tags abbreviation descriptions

Tag	Description	Tag	Description
CC	Coordinating conjunction	PRPS	Possessive pronoun
CD	Cardinal number	RB	Adverb
DT	Determiner	RBR	Adverb, comparative
EX	Existential there	RBS	Adverb, superlative
FW	Foreign word	RP	Particle
IN	Preposition or subordinating conjunction	SYM	Symbol
JJ	Adjective	TO	To
JJR	Adjective, comparative	UH	Interjection
JJS	Adjective, superlative	VB	Verb, base form
LS	List item marker	VBD	Verb, past tense
MD	Modal	VBG	Verb, gerund or present participle
NN	Noun, singular or mass	VBN	Verb, past participle
NNS	Noun, plural	VBP	Verb, non-3rd person singular present
NNP	Proper noun, singular	VBZ	Verb, 3rd person singular present
NNPS	Proper noun, plural	WDT	Wh-determiner
PDT	Predeterminer	WP	Wh-pronoun
POS	Possessive ending	WPS	Possessive wh-pronoun
PRP	Personal pronoun	WRB	Wh-adverb

Table 5 POS tagging frequencies for Review 7

POS Tag		DT	IN	JJ	MD	NN	NNS	PRP	RB	VB	VBD	VBP
Review	Review 7	6	3	3	1	7	2	6	6	1	2	3

3.1.4 Stylometric

(Wang et al., 2016) proposed a framework using Stylometric feature or either the text of lexical or syntactic feature character. Lexical features mostly indicate the character and type of words the writer is preferred to use while writing the review and it also includes features of upper case characters and the average word length used in the review.

The syntactic feature is about the writing of the reviewer and it also considers some features like the punctuations or function words which include "a", "of", and "the".

3.1.5 Semantic

This feature is about identifying the underlying meaning of the reviews and the underlying meaning of each word used in the review. (NAIK et al., 2019) proposed a framework creating a sematic language model for the process of detecting false reviews found online. The rationale is that changing a word like "love" to "like" in a review should not affect the similarity of the reviews since they have similar meanings.

3.1.6 Review characteristic of Review centric features

(Jindal et al., 2020)(Adike & Reddy, 2016) the features contain information about the review known as the metadata. The characters of the reviews can be the length of the review, data, rating, the time period of the review uploaded online. Reviewer, review, and store id, etc. the illustration of review characters of review-centric feature is tabulated in 6.

Table 6 Reviews characteristics dataset structure

Review	Review ID	Product ID	Reviewer ID	Rating	Helpfulness	Review char	Review words	Date	Time

Revi ew4	152	0123 45	226	1	1	38	9	8/5/2 013	09 :24
Revi ew5	153	0123 45	789	5	0	35	10	9/1/2 015	12 :06
Revi ew6	154	0123 45	789	5	0	25	7	9/1/2 015	12 :07

3.2 Reviewer centric features and Product-Centric Features

By identifying the spammers we can detect fake reviews because n number of reviews are found online about the product. Various combinations of features, characteristics of the features, and its behavioral pattern of the features have been discussed earlier by (Arabameri et al., 2020; Chen et al., 2015; Hussain et al., 2020; Pandey & Rajpoot, 2019; Zeng et al., 2019). The illustration of the reviewer-centric feature is tabulated in 7 and more details about the features of the reviewer-centric feature will be discussed below as follows

Table 7 Reviewers characteristics dataset structure

R eview	Pr oduct ID	Re view ID	Rev iewer name	E mail addre ss	# of Revie ws	Fi rst revie w	L ast revie w	M ax# revie ws per day	A verag e ratin g	D ate	T im e
R eview 1	12 345	15 2	JO	jo @gm ail	20 00	09 /01/1 3	09 /30/1 4	3 0	5	09 /30/1 4	1 2:0 5
R eview 2	12 345	15 3	LI	jo @gm ail	23 00	09 /01/1 3	09 /30/1 4	3 1	5	09 /30/1 4	1 2:0 6
R eview 3	12 345	15 4	SA	sa @gm ail	3	05 /02/1 1	06 /05/1 4	1	4	06 /05/1 4	1 2:0 0

3.2.1 Maximum number of reviews

It was found that about 73 percent of spammers will write more than 6 reviews per day. By identifying the number of reviews written by a reviewer per day will help us in detecting the spammers because original reviewers will not write more than one or two reviews in a single day.

3.2.2 Percentage of positive reviews

It was found that about 84 percent of spammers will consider their own review as a positive review. So it will be useful to detect untrustworthy reviewers based on positive reviews written by reviewers about themself.

3.2.3 Review length

The review length is very important for identifying the spammers because the spammer's review length was not above 130 words. In a study, it was found that reliable reviewers will have more than 250 words in their reviews.

3.2.4 Reviewer deviation

By identifying the review rating will helps us in detecting dishonest reviewers. Review rating is important to know about spam reviewers and reliable reviewers.

3.2.5 Maximum content similarity

A similar type of reviews for a different type of product with a slight change in text is the important point for identifying spammers.

4. Proposed Algorithm for Review and reviewer based features

There are several spam detection techniques preferred. But this paper deals with 3 important Machine Learning Approach for spam detection. The proposed spam detection method is shown in fig 3. As
1.   The Supervised Learning
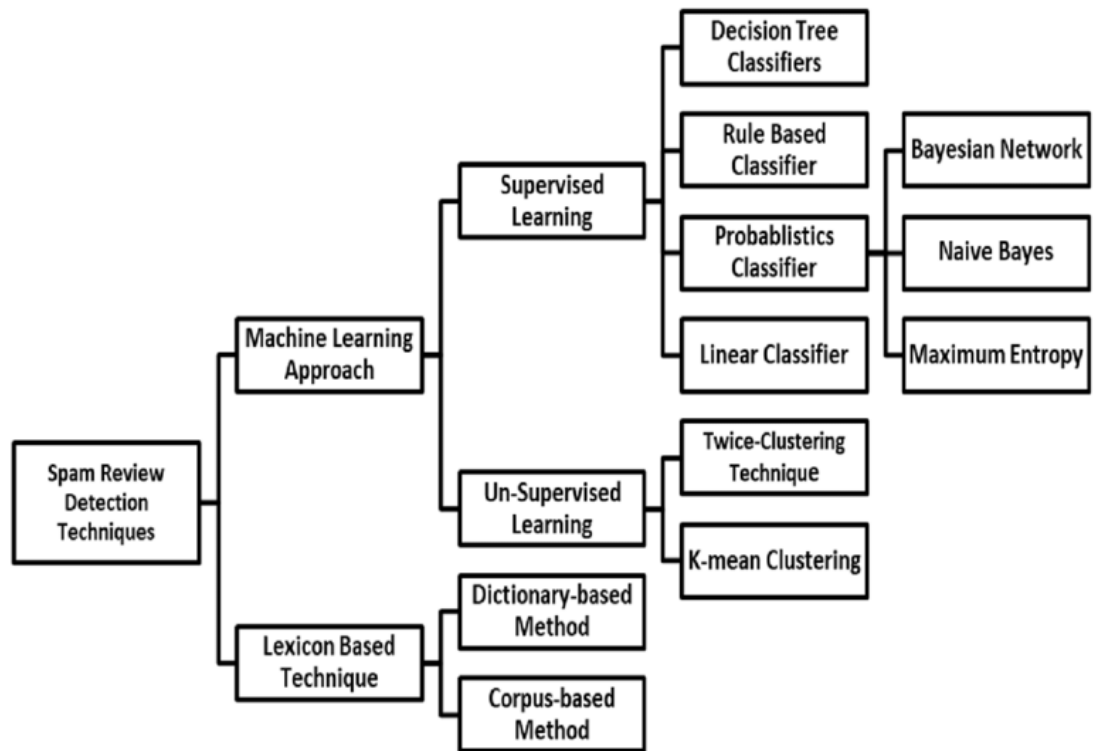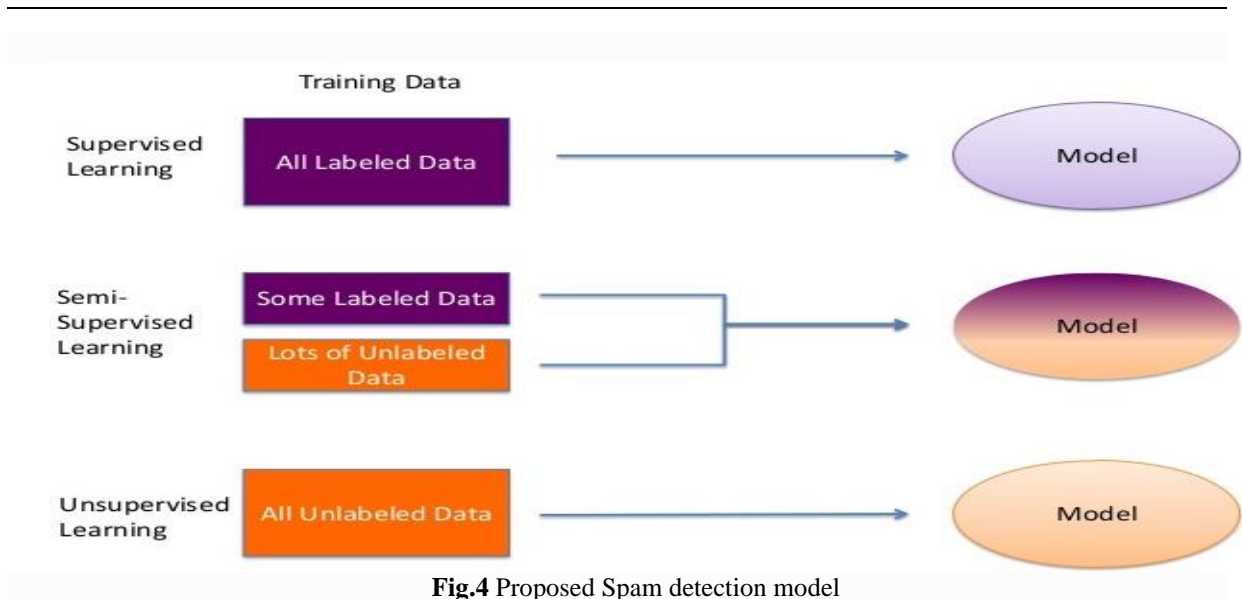2.   The Semi-supervised Learning
3.   The Unsupervised Learning

Fig 3. Overall Spam review detection Techniques

4.1 List of the algorithm used for the study

Review-centric spam detection is one of the most commoly used spam detection methods for online spam review. In this method, the ML-based techniques are used for developing the model considering the content and metadata of the online review. Among the 3 ML method, the supervised learning method is the most preferred method for detecting online review spam. The supervised learning method requires labeled data, while the unsupervised learning method prefers non-labeled data, the semi-supervised method includes both labeled and non-labeled data. These 3 methods are illustrated in table 8 as follows

Table 8 Types of machine learning techniques

Method	Attributes
Supervised Learning	Learning from a set of labeled data
	Requires labeled training data
	Most common form of learning
Unsupervised Learning	Learning from a set of unlabeled data
	Finds unseen relationships in the data independent of class label
	Most common form is clustering
Semi-supervised Learning	Learning from labeled and unlabeled data
	Only requires a relatively small set of labeled data which is supplied with a large amount of unlabeled data
	Ideal for cases such as review spam where vast amount of unlabeled data exist

Fig.4 Proposed Spam detection model

4.1.1 Supervised learning

This method is used for detecting review spam by mainly considering the classification issue for separating the reviews into 2 groups as

1)   Spam reviews and
2)   Non-spam reviews

(Gong et al., 2020) proposed a framework using a supervised learning method for deceptive opinion spam. This method mainly focused on extracting the textural features using Natural Language Processing (NLP).

Different types of methods were proposed under the supervised learning method they are

1)   NLP- Natural Language Processing
2)   AUC- (Area Under the receiver operating characteristic Curve) this is used for finding duplicate spams which will yield results of 51%
3)   The Naïve Bayes and Support vector machine was used for classification but they did not perform well.

4.1.2 Unsupervised learning

It is difficult to always provide a labeled dataset for review spam detecting utilizing supervised learning method so the above-mentioned method may not be suitable always. The unsupervised learning method is the best answer because it does not require a labeled dataset for detecting review spam. [1] proposed a framework using an unsupervised learning method through which a text mining model was developed and implemented in semantic language for detecting fake reviews found online.

Different types of methods were proposed under the unsupervised learning method they are

1)   For calculating the degree of untruthfulness for reviews based on the duplicate identification results by estimating the overlap of semantic contents among reviews using a Semantic Language Model (SLM). The data model was trained using SVM. The SLM model in unsupervised learning achieved a high detection rating for duplicate spam detection.

2)   Random Undersampling (RUS) and Random Oversampling (ROS) were used for collecting the degree of untruthfulness and the collected model was trained using Naïve Bayes and SVM but Naïve Bayes overpowered SVM in performance.

4.1.3 Semi-supervised learning

(Yao et al., 2019) With unlabelled dataset along with a labeled dataset will improve the performance of the model when compared to that of the supervised learning method.

Different types of methods were proposed under the semi-supervised learning method they are

1)   PU-learning is one of the commonly known semi-supervised learning methods implemented by [35] which helps to learn from positive reviews and also from the unlabelled dataset.
2)   [34] proposed a framework using the PU-learning method for detecting Deceptive opinion spam.
3)   [3] the evaluation process for PU-learning is done using F-Measures.
4)   The classifiers are trained using SVM and Naïve Bayes model

4.2 Performance yield Reviewer centric review spam detection using an algorithm

(Lin et al., 2014) the combination of both review centric feature and reviewer centric features yields a better result for detecting online spam reviews. Collecting behavioral evidence of spammers will help in identifying the review spam.

(Saumya & Singh, 2018) proposed a framework by completely analyzing the supervised learning approach while utilized in deceptive review detection. The researcher tested the proposed model using AMT (Amazon Mechanical Turk) on a real-world fake reviews dataset procured from Yelp.

But it was found that with the n-gram feature this study performed well on the Amazon Mechanical Turk dataset but the same while using in Yelp dataset the performance was slightly worse. It was also found that the behavioral feature yields better results than the linguistic feature on the Yelp dataset.

3 different set of features were used in this experiment
1) The LIWC
2) The POS tag
3) Bigrams

4.2.1 Ratio of Amazon Verified Purchase (RAVP)

RAVP is based on the amazon verified purchase which is done by separating the product based on reviews written about the product by consumers or users. it is the fact that verified reviews will be trustworthy than non-verified reviews. The reviewer with higher nu,ber of RAVP will be considered as a trustworthy reviewer.

4.2.2 Rating Deviation (RD)

By identifying the review rating will helps us in detecting dishonest reviewers. Review rating is important to know about spam reviewers and reliable reviewers.

4.2.3 Burst Review Ratio (BRR)

BRR is based on the Computed ratio of the review wriiten by the reviewer. The burst occurs to the number of reviews written by the reviewer.

4.2.4 Review Content Similarity (RCS)

The RCS is based on a similar type of reviews for a different type of product with a slight change in text is the important point for identifying spammers.

4.2.5 Reviewer Burstiness (RB)

The RB is based on the number of times the product and reviewers get burst. The more that this occurs, the more likely the reviewer is a spammer.

5. Comparative analysis, Decision of Proposed method, and further suggestions

It is always important to consider the techniques and approaches used by researchers in the past to detect online review spam. In the previous section lot of discussion took place based on the methods needed to be considered for extracting features and an overview of the ML approach was also discussed.

It was found that feature engineering had great significance in impacting the performance of the classifier. (Khurshid et al., 2018). In previous studies, different sets of learning methods, datasets, performance metrics have been used by achieving different results based on the preferred method for the study.

Table 8 gives a complete comparative study based on the method as discussed earlier with the different approaches used for achieving the best result for the study. From the previous study, it was found that the combination of multiple methods of features will increase the performance.

(Duhan et al., 2017) proposed a framework by including both review-centric and reviewer-centric features to increase the performance by adding it to the textural features.

(Barushka & Hajek, 2019) there was a slight increase in performance while combining bigram along with LIWC. Following this many experiments preferred the same type of dataset for improving the performance. It is evident that bigram along with LIWC will achieve better performance. Other studies utilized different datasets in such cases it is difficult to make a comparative analysis.

Almost many studies used the ML-based approach. This study used a supervised learning approach in all parts except in 3 areas. They are in
1) LR- Logistic Regression
2) NB- Naïve Bayes
3) SVM- Support Vector Machine

The support vector machine will always better performance than LR and NB. But in some, it will be vice-versa. The SVM was not compared with other learners in such cases it is difficult to decide SVM as the best learning method. further research on this topic will mainly focus on testing the performance with multiple learners along with multiple datasets by utilizing different features

Table 9 Comparison of previous works and results for review spam detection along with the relative complexity of the approach (including feature extraction and learning methodology)

Dataset	Features used	Learner	Performance metric	Score	Method complexity
5.8 million reviews written by 2.14 reviewers crawled from amazon website	Review and reviewer features	LR	AUC	78%	Low
5.8 million reviews written by 2.14 reviewers crawled from amazon website	Features of the review, reviewer and product characteristics	LR	AUC	78%	Medium
5.8 million reviews written by 2.14 reviewers crawled from amazon website	Text features	LR	AUC	63%	Low
6000 reviews from Epinions	Review and reviewer features	NB with Co-traning	F-score	0.631	High
Hotels through Amazon Mechanical Turk(AMT)by Ott et al	Bigrams	SVM	Accuracy	89.6%	Low
Hotels through Amazon Mechanical Turk(AMT)by Ott et al	LIWC+Bigrams	SVM	Accuracy	89.8%	Medium
Hotels through Amazon Mechanical Turk(AMT)by Ott et al+ gathered 400 descriptive hotel	LIWC+POS+Unigrams	SAGE	Accuracy	65%	High
Yelps real life data	Behavioral features combined with Bigrams	SVM	Accuracy	86.1%	Medium
Hotels through Amazon Mechanical Turk(AMT)by Ott et al	Stylometic features	SVM	F-measure	84%	Low
Hotels through Amazon Mechanical Turk(AMT)by Ott et al	n-grams features	SVM	Accuracy	86%	Low
Data set collected from amazon.com	Syntactical ,lexical and stylistic features	SLM	AUC	.9986	High
Their own crawled Arabic	Review and reviewer features	NB	F-measure	.9959	Low

reviews from tripadvisor.com					

6. Conclusion

This research mainly focused about various machine learning approaches and other techniques utilizing for detecting online review spam. From the experimentation carried out in the study, it was found that the supervised machine learning approach yield better performance when compared to other leaning approaches such as unsupervised and semi-supervised approach. But the supervised approach is considered a restricted one because using a labeled dataset for training and identifying the reviews manually showed lesser accuracy. By which many experiments used small datasets for study. The feature extraction from the textural features was mainly done through a bag of words method and through POS tags for training the spam. Many experiments incorparted multiple types of features to obtain better results on the basis of the performance of the classifier. From this study, it was found online review soam detection has huge significance among researchers as well as owners of a business firm. This is because the fake review can impact the consumers and affect the behavioral pattern and also the purchasing decision to be made by them while purchasing a product.

References

A. NAIK, U. U., RAMAKRISHNA, V., & SCHOLAR, P. G. (2019). A network-based spam detection framework for reviews in online social media. *Complexity International*, *23*(03).

B. Tang, X., Qian, T., & You, Z. (2019). Generating behavior features for cold-start spam review detection. *International Conference on Database Systems for Advanced Applications*, 324–328.

C. Barushka, A., & Hajek, P. (2019). Review spam detection using word embeddings and deep neural networks. *IFIP International Conference on Artificial Intelligence Applications and Innovations*, 340–350.

D. Khurshid, F., Zhu, Y., Xu, Z., Ahmad, M., & Ahmad, M. (2018). Enactment of ensemble learning for review spam detection on selected features. *International Journal of Computational Intelligence Systems*, *12*(1), 387–394.

E. Hussain, N., Turab Mirza, H., Rasool, G., Hussain, I., & Kaleem, M. (2019). Spam review detection techniques: A systematic literature review. *Applied Sciences*, *9*(5), 987.

F. Mohammadi, S., & Mousavi, M. R. (2020). Investigating the Impact of Ensemble Machine Learning Methods on Spam Review Detection Based on Behavioral Features. *Journal of Soft Computing and Information Technology*, *9*(3), 132–147.

G. Shahariar, G. M., Biswas, S., Omar, F., Shah, F. M., & Hassan, S. B. (2019). Spam Review Detection Using Deep Learning. *2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, 27–33.

H. Wang, X., Liu, K., He, S., & Zhao, J. (2016). Learning to represent review with tensor decomposition for spam detection. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 866–875.

I. Adike, M. R., & Reddy, V. (2016). Detection of fake review and brand spam using data mining. *Int J Recent Trends Eng Res*, *2*(7), 251–256.

J. Dewang, R. K., & Singh, A. K. (2018). State-of-art approaches for review spammer detection: a survey. *Journal of Intelligent Information Systems*, *50*(2), 231–264.

K. Li, H., Fei, G., Wang, S., Liu, B., Shao, W., Mukherjee, A., & Shao, J. (2017). Bimodal distribution and co-bursting in review spam detection. *Proceedings of the 26th International Conference on World Wide Web*, 1063–1072.

L. Xue, H., Wang, Q., Luo, B., Seo, H., & Li, F. (2019). Content-aware trust propagation toward online review spam detection. *Journal of Data and Information Quality (JDIQ)*, *11*(3), 1–31.

M. Duhan, N., Mittal, M., & others. (2017). Opinion mining using ontological spam detection. *2017 International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions)(ICTUS)*, 557–562.

N. Neisari, A. (2020). *Spam Review Detection Using Self-Organizing Maps and Convolutional Neural Networks*. University of Windsor (Canada).

O. Jindal, R., Seeja, K. R., & Jain, S. (2020). Construction of Domain Ontology utilizing Formal Concept Analysis and Social Media Analytics. *International Journal of Cognitive Computing in Engineering*.

P. Etaiwi, W., & Awajan, A. (2017). The effects of features selection methods on spam review detection performance. *2017 International Conference on New Trends in Computing Sciences (ICTCS)*, 116–120.

Q. Ghai, R., Kumar, S., & Pandey, A. C. (2019). Spam detection using rating and review processing method. In *Smart Innovations in Communication and Computational Sciences* (pp. 189–198). Springer.

R.   Xie, S., Wang, G., Lin, S., & Yu, P. S. (2012). Review spam detection via temporal pattern discovery. *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 823–831.

S.   Crawford, M., Khoshgoftaar, T. M., Prusa, J. D., Richter, A. N., & Al Najada, H. (2015). Survey of review spam detection using machine learning techniques. *Journal of Big Data*, *2*(1), 23.

T.   Rayana, S., & Akoglu, L. (2015). Collective opinion spam detection: Bridging review networks and metadata. *Proceedings of the 21th Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*, 985–994.

U.   Li, H., Fei, G., Wang, S., Liu, B., Shao, W., Mukherjee, A., & Shao, J. (2017). Bimodal distribution and co-bursting in review spam detection. *Proceedings of the 26th International Conference on World Wide Web*, 1063–1072.

V.   You, L., Peng, Q., Xiong, Z., He, D., Qiu, M., & Zhang, X. (2020). Integrating aspect analysis and local outlier factor for intelligent review spam detection. *Future Generation Computer Systems*, *102*, 163–172.

W.   Xue, H., Li, F., Seo, H., & Pluretti, R. (2015). Trust-aware review spam detection. *2015 IEEE Trustcom/BigDataSE/ISPA*, *1*, 726–733.

X.   Gong, M., Gao, Y., Xie, Y., & Qin, A. K. (2020). An attention-based unsupervised adversarial model for movie review spam detection. *IEEE Transactions on Multimedia*.

Y.   Hajek, P., Barushka, A., & Munk, M. (2020). Fake consumer review detection using deep neural networks integrating word embeddings and emotion mining. *Neural Computing and Applications*, 1–16.

Z.   Hussain, N., Mirza, H. T., Hussain, I., Iqbal, F., & Memon, I. (2020). Spam Review Detection Using the Linguistic and Spammer Behavioral Methods. *IEEE Access*, *8*, 53801–53816.

AA. Chen, C., Zhao, H., & Yang, Y. (2015). Deceptive opinion spam detection using deep level linguistic features. In *Natural Language Processing and Chinese Computing* (pp. 465–474). Springer.

BB. Pandey, A. C., & Rajpoot, D. S. (2019). Spam review detection using spiral cuckoo search clustering method. *Evolutionary Intelligence*, *12*(2), 147–164.

CC. Zeng, Z.-Y., Lin, J.-J., Chen, M.-S., Chen, M.-H., Lan, Y.-Q., & Liu, J.-L. (2019). A review structure based ensemble model for deceptive review spam. *Information*, *10*(7), 243.

DD. Arabameri, A., Tiefenbacher, J. P., Blaschke, T., Pradhan, B., & Tien Bui, D. (2020). Morphometric analysis for soil erosion susceptibility mapping using novel gis-based ensemble model. *Remote Sensing*, *12*(5), 874.

EE. Rout, J. K., Singh, S., Jena, S. K., & Bakshi, S. (2017). Deceptive review detection using labeled and unlabeled data. *Multimedia Tools and Applications*, *76*(3), 3187–3211.

FF. Saumya, S., & Singh, J. P. (2018). Detection of spam reviews: A sentiment analysis approach. *Csi Transactions on ICT*, *6*(2), 137–148.

GG. Kumar Tripathi, A., Sharma, K., & Bala, M. (2019). Fake review detection in big data using parallel bbo. *International Journal of Information Systems & Management Science*, *2*(2).

HH. Tang, X., Qian, T., & You, Z. (2020). Generating Behavior Features for Cold-Start Spam Review Detection with Adversarial Learning. *Information Sciences*.

II. Visani, C., Jadeja, N., & Modi, M. (2017). A study on different machine learning techniques for spam review detection. *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*, 676–679.

JJ. Othman, N. F., & Din, W. (2019). Youtube spam detection framework using na{\"\i}ve bayes and logistic regression. *Indonesian Journal of Electrical Engineering and Computer Science*, *14*(3), 1508–1517.

KK. Yao, C., Wang, J., & Kodama, E. (2019). A Spam Review Detection Method by Verifying Consistency among Multiple Review Sites. *2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; IEEE 5th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, 2825–2830.

LL. Lin, Y., Zhu, T., Wang, X., Zhang, J., & Zhou, A. (2014). Towards online review spam detection. *Proceedings of the 23rd International Conference on World Wide Web*, 341–342.

MM. Lin, Y., Zhu, T., Wu, H., Zhang, J., Wang, X., & Zhou, A. (2014). Towards online anti-opinion spam: Spotting fake reviews from the review sequence. *2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)*, 261–264.