ISSN : 2455-9679
SJIF Impact Factor : 6.008

IJRTSM

INTERNATIONAL JOURNAL OF RECENT TECHNOLOGY SCIENCE & MANAGEMENT

"A REVIEW ON SPAM DETECTION BASED ON MACHINE LEARNING TECHNIQUES"

Swati Toriya [1], Prof. Varsha Thakur [2]

[1] M. Tech Scholar, Department of Computer Science Engineering, Mittal Institute of Technology Bhopal, M.P., India
[2] Professor, Department of Computer Science Engineering, Mittal Institute of Technology Bhopal, M.P., India

ABSTRACT

Online reviews are usually the most important factor in a customer's decision to purchase a product or service and a valuable source of information that can be used to determine public opinions about those products or services. Because of their influence, manufacturers and retailers follow close to customer feedback and comments. Online reviews raises potential concerns that misconduct may create false reviews for artificially promoting or devaluing products and services. This practice is called opinion (spam) spam, where spammers manipulate and poison comments (i.e. make false, false or misleading comments) for profit. Since not all online reviews are authentic, it is important to develop techniques to detect spam. By using Natural Language Processing (NLP) to extract meaningful functions from text, various machine learning techniques can be used for spam detection. In addition, proofreading information in addition to the text itself can be used to aid this process. In this article, we examine the excellent machine learning techniques proposed to address the problem of comment spam detection, as well as the performance of various methods for classifying and detecting spam comments.

Key Words: *Review spam, Opinion mining, Web mining, Machine learning, Bigdata; Classification.*

I. INTRODUCTION

As the Internet continues to grow in size and importance, the number and influence of online comments continues to increase. Comments can affect people in a wide variety of industries, but they are especially important in e-commerce. In e-commerce, reviews and comments on products and services are usually the most convenient way for buyers to decide whether to buy, if not the only way or not to buy. Online comments can be generated for various reasons. To improve and enhance their business, online retailers and service providers can generally ask their customers to provide feedback on the experience of the products or services they have purchased and whether they are satisfied. If customers have a very good or bad experience with a product or service, they may also have a tendency to comment on the product or service. While online reviews can be helpful, blind trust in these reviews can be dangerous for both sellers and buyers. Many people check online reviews before placing online orders; However, reviews can be poisoned or thrown for profit, so any decision based on online reviews must be made with caution. In addition, business owners can reward those who write good reviews of their products, or they can pay someone to write negative reviews of their competitors' products or services [1]. These fake comments are considered spam and because of the importance of the comments, they can have a huge impact on the online market. Due to the loss of consumer confidence, spam comments will also have a negative impact on businesses. This problem is serious enough to attract the attention of the mainstream media and the government. Eg. The BBC and the New York Times reported that "fake reviews are becoming a common problem on the Internet, and a photography company has recently suffered fewer than hundreds of defamatory consumer reviews." In 2014, the Canadian government issued a warning, "urging consumers to be wary of fake online adaptations, giving the impression that ordinary consumers do," and estimates that one-third of all online

ISSN : 2455-9679
SJIF Impact Factor : 6.008

reviews are fake1. Since comment spam is a common and destructive problem, it is an important but challenging problem to develop methods to help businesses and consumers distinguish between real reviews and fake reviews. [2]

Feature Engineering Using the Linguistic Approach

Linguistic methods are the main methods of extracting or constructing functions from commentary datasets. In addition, this method uses only the comment text and performs various steps, such as data preprocessing, tokenization, conversion, and function selection .In this study, the researchers analyzed and discussed common functional engineering techniques.

Pre-Processing

Removing Stop Words or Punctuation Generally: The comment text contains unnecessary words such as "is", "the", "and" and "a". These words are not useful for detecting spam comments, so it is best to delete them before selecting them to avoid noise and unnecessary marking. Comment e.g. "This is a very good car." After removing stop words and punctuation, the comment is displayed as "good car".

Part of speech tagging: This involves basic labeling of word functions with part of the speech based on the context of the review text n addition; the relationship to adjacent and related words is also highlighted in the review text. A simplified form of tag-of-speech tagging is recognizing words such as nouns, verbs, adjectives, adverbs, and so on.

Stemming Word-The voice algorithm converts different types of words into a single recognizable form. Consider e.g. Words like "work", "work" and "work" as examples of the word work. Before tagging the comment text, voting must be applied to it.

Tokenization -In this method, a single word or a group of words is used as a function. This language technique is called unigram when you choose a word, bi-gram when you choose two words, tri-gram when you choose three words, etc. This technique is commonly called n-gram For example, consider a comment "good car" and apply different n-gram techniques to it. Unigram: ["good", "car"], Bi-gram: ["good car"], Uni + Bi-gram: ["car", "good", "good car"]. This work uses different n-gram combinations on the review data.

Transformation Document-Term matrix is used to represent tokens generated by the n-gram model in the form of a sparse matrix. A sparse matrix defines the frequency of terms or tokens in the collection of the reviews. It was observed by the literature review that most of the researchers use the following two techniques for transformation.

Reviewer centric and review spam detection - We mentioned earlier that identifying reviewers who wrote fake reviews is very important in detecting spam. For spam detection, it may be better to use a combination of comment-centered features and comment-centered features than just comment-centered methods. In addition, it is easier to collect behavioral evidence for spammers than to identify spam comments [3]. Emphasizes functional technology and the influence of these functions on the performance of spam detectors. In addition, the benefits of supervised, non-supervised, and semi-supervised learning methods are analyzed, and current research results and comparative analysis using each method are introduced. Finally, we provide recommendations for detecting spam comments that require further investigation, as well as best practices for future research. To the best of our knowledge, this article contains information about all datasets that have been used in the review literature or generated for use. [4] The structure of this article is as follows. The functional engineering part of comment-spam detection outlines functional techniques in this field, including comment-centered spam detection and comment-centered spam detection. The section "Comment-centered comment spam detection" discusses and analyzes the current research in the use of supervised, non-monitored and semi-monitored machine learning for comment-centered spam detection. The section on detecting spam detection outlines research using proof-focused features. The section on comparative analysis and recommendations contains discussions and comparisons between the different proposed methods. The conclusion summarizes our findings and reviews the importance of past and future work. [5]

Data Source: User perception is a remarkable rule for changing the nature of the management presented and improving deliveries. Websites, audit areas, information and small online journals can provide a good understanding of the

ISSN : 2455-9679
SJIF Impact Factor : 6.008

project's collection level and management.

A. Blogs As the use of the Internet continues to expand, blogs and blog pages are getting faster and faster. The blog site has become the most well-known intention to express personal feelings.

B. Crawl Pages Online sites are the most popular source of crawl data. Users post their opinions on these sites, and these sites analyze their opinions for future use.

C-Weibo the Weibo website is the most important and popular source. Twitter is a Weibo site where people can post comments and share their views in the form of "tweets".

II. LITERATURE REVIEW

Rutuja Katpatal et.al. (2018) Today, Twitter spam has become a major problem. Subsequent work focused on detecting spam on Twitter through a machine learning system, which uses the actual elements of the tweet. In our data collection of spam tweets, we observed that the measurable characteristics of spam tweets varied over a period of time. With these guidelines, the effectiveness of the machine -based learning category decreases. This problem is called Twitter's spam drive. To address this problem with a specific end goal, a technique called the Lfun scheme is used, which can detect tweets that have changed from unmarked tweets and incorporate them into the classification training process. The new training data is used to train other data containing unmarked tweets, which will lead to the detection of spam tweets. Our proposed solution will correct the training data after a certain period of time, such as discarding old samples, deleting unnecessary information and storing space. [6]

Sayali Kamble et.al. (2018) With the growing popularity of Twitter and more Twitter users tweeting around the world, real-time search technology has emerged, allowing everyone to follow monitor the impact of news and events on Twitter. These actions allow the user to disseminate information and allow the user to talk about that action and then publish their status. These services have opened the door to new types of spam. At all times, the most common behavior on Twitter is considered to be an opportunity to generate traffic and revenue. Spammers will post irrelevant tweets and malicious links, and will post frequent tweets on the same topic in an attempt to attract the attention of tweeters. The longer the spam tweet, the more it will be exposed to the suspects. So it is very important to look for spam tweets as soon as possible. Real -time finding is necessary to minimize the losses caused by spam. In order to monitor spam, there are many machine learning techniques that can analyze the statistical properties of tweets. The proposed system uses a variety of APIs to verify and analyze URLs to see if those URLs are corrupting, which helps improve the detection of spam activities in a timely manner. [7]

Kshitiz Badola et al. (2021) with the growing demand for social life in today's world, the popular Twitter platform plays an important role in each citizen's social interaction, or tweeting for others or exploring different parts of the world. But this area is now infected with some viruses. To increase traffic to spam sites, they link their URLs to informational tweets, and there is no connection between the content of the URL and the tweet message. These are called spam tweets. This article provides a unique way to determine if a user's tweets are spam or not spam, and uses filtering technology to use a vector converter on tweets and associated URLs to predict the differences between them. these.[8].

Shivangi Gheewala et.al. (2018) an online social network (OSN) is an information platform for people to establish and manage social interactions. At OSN, billions of users regularly participate in social media, content and ideas dissemination, networking, recommendations, monitoring, alerts and social events. OSN's popularity has opened up new perspectives and challenges in social network research, and has attracted the attention of many sectors. A social network is a place where social events, business events, entertainment and information are exchanged. It creates a harmonious environment where people can share their interests and what they do, or be interested in the interests of others. While social networks have brought great benefits to people, they have also harmed all types of people. This sector has caused huge economic losses in our society and even threatened national security. All social networks Facebook, Twitter, LinkedIn, etc. are highly vulnerable to malware activity. Twitter is one of the largest Weibo

ISSN : 2455-9679
SJIF Impact Factor : 6.008

networking platforms. On average, millions of users send more than 5 billion tweets to Twitter every day. Twitter is so crowded and crowded that it's easy to get involved in bad work. Malicious activities include malware infiltration, distribution and social attacks. Spam uses social engineering attack strategies to send spam tweets, spam URLs, etc. That's why Twitter is a potential place to spread unusual spam accounts. This influence prompted researchers to develop a model to analyze, identify, and heal from Twitter scams. The Twitter network is flooded with tens of thousands of spam accounts, which can compromise the security and privacy of ordinary users. Improving real -time user security and identifying spam profiles has become an important part of research. [9]

Niddal Imam et.al. (2019) found that image-based spam is a perennial problem on online social networks (OSNs) such as Facebook and Twitter. Spam is common in all forms of online communication, such as email and the Internet. However, due to the increasing number of spammers and the potential negative impact on users, searchers and researchers are turning to spam on the OSN. Various types of spam are found in the OSN. Spam images are images with entrenched dirty text and one of the most difficult types of spam to deal with. Image processing will overwhelm the classifier and affect the detection efficiency. Therefore, spammers use this problem to attack more complex attacks, such as avoidance attacks. After reviewing popular Arabic tags and topics on Twitter, I found a lot of image -based spam. Therefore, this article proposes a way to detect image -based spam with Arabic posts on Twitter using deep learning (DL) technology. This article uses an effective and accurate Stage Text Detector (EAST) and Repetitive Neural Network (CRNN) model for text identification and text recognition. After the text extraction process, blacklists and white lists are used to classify the text as spam or non -spam. The proposed article classification technology is flexible and robust against certain article classification attacks. [10]

K.Ushasree Santoshi et.al. (2021) Twitter is a popular social media platform with over 300 million monthly users and 500 million daily tweets. This is the main reason why spammers use Twitter for obscene behaviors, such as spreading malware that steals user information and tweets that contain incorrect or incorrect URLs, trustingly or unfollowing users, and sending fake tweets to attract the attention of natives. Prohibited advertising. According to reports, Twitter has been collecting data on active users and analyzing their behavior in recent years. Clear reports show that more than 32 million users interact with servers every day to get entertainment information. Therefore, in today's social world, it is very important to identify and filter out harmful tweets or trends that are harmful or disliked by users. This article talks about how to analyze tweets based on the words included in the tweets and classify them as spam and ham. Although there are various machine learning and in-depth analysis methods for classifying and filtering spam tweets, such as SVM, collection methods and binary spy models use simple Bayesian categorization. More recently, twitter users have been experiencing at least some data theft software by visiting or accessing unwanted spam or tweets. Because many people lose money or personal information, they need to be carefully considered. Aside from software stealing data, the tendency to fraud is also alarming. It needs to be tested. Because of its automatic operation, spam can interact with many people. [11]

Surendra Sedhai et.al. (2018) most spam detection technologies on Twitter are designed to identify and block users from sending spam tweets. In this article, we provide a proposal with semi-supervisor spam detection (S3D) for tweet-level spam detection. The proposed system consists of two main modules: a spam detection model in real mode and a model update model in batch mode. The spam detection model consists of four simple detectors: 1) Domain blacklist detector, used to detect tweets with blacklisted URLs; 2) Average duplicate detector used to reliably mark and alert warnings through double-digit ad text; 3) a secure ham detector for noting tweets posted by secure users that are free of spam, and 4) a multi -category spy to note the remaining tweets. The information required for the spyware will be updated in series based on the tweets marked in the previous window. Experiments with multiple data sets show that the system learns to balance new spam campaign patterns and maintains good spam filtering reliability in the stream of tweets. [12]

Dani Gunawan et.al. (2018) more recently, the use of text messages or text messages has become the promotion of products or services or even fraud. Indonesian mobile phone users are also facing a similar situation. An easy way to solve this problem is to create a blacklist of phone numbers or specific keywords and phrases. However, this method does not work because the spammer may change the phone number or change the content of the text message. Meanwhile, another way is to use text classification, such as Naive Bayes, close proximity (kNN), and vector support

ISSN : 2455-9679
SJIF Impact Factor : 6.008

machine (SVM) to determine the pattern of text messages. This research proposes Twitter's LDA algorithm to detect spam in Indonesian. There are a total of 985 short messages, which is divided into 774 short messages in the set training data and 211 short messages in the test data. This data includes 860 spam emails and 125 ham text messages. Before using the training and testing process, all text messages must be submitted in advance. There were 5 experiments performed in this study, with an mean f-94.26% and a mean frequency of 96.49%. According to this result, Twitter's LDA algorithm shows good success in identifying spam in Indonesia. [13]

Wafaa Daffa et.al. (2018) with the rapid growth of Weibo websites like Twitter, various ads are also growing at a rapid pace. Spam detection is a major problem on social networking sites. This article discusses the detection of spam URLs on Twitter by providing malicious behavior categories, detection strategies, detection capabilities, detection techniques, and limitations (if available). We further examined the effectiveness of machine learning classification based on various published works. So, we used four classifiers in the Twitter account data, which included 10,713 users, marked with 5358 benign and 5355 spam, with 17 internal and user-based effective features. Our results show that of the four categories, the non-mixed forest category yields the highest score with a degree of 96.4%, whereas the accuracy of the J48 category. is slightly different, which is 94.5%. [14]

E. Elakkiya et.al. (2020) Facebook, Twitter, YouTube and other social networking sites are popular among Internet users for sharing information and communication. This proliferation has also attracted online criminals to use social networks to spread their criminal activities, including spam. Spam messages can contain unexpected information or malicious links, which can harm Twitter users and make them feel insecure. Therefore, there is an urgent need to find effective spam, which requires a clear definition of the characteristics of spam behavior. This is done through information theory-based functional selection methods, since these methods select information functions that retain their original meaning. Most information theory-based methods focus on retaining features that have more information and eliminating features that lack more information and Features to improve performance. This could result in new losses. In addition, work that is not particularly important may be needed when combined with other work. Therefore, the Anti-Spam Search Engine (CIFAS) has been developed by the service community to combat combinatorial search, which uses fuzzy access as a form of exercise to provide optimal results. In this work, the fuzzy cross entropy is used to construct the information recorded in the selected activity as the information recorded in the full sequence. The test results show that the CIFAS proposal provides the same spam protection as the original feature, and is better than the existing method in fact, the error rate and F and measurement. of standard Twitter data [15].

III. PROBLEM STATEMENT

The fierce competition between filtering methods and spammers is going on every day because spammers are starting to use difficult methods to overcome spam filters, e.g. by using random sender addresses or at the beginning or end of the subject line in the message Random characters are added. Lack of machine learning focuses on developing models that can predict activity. Spam is a waste of time for users because they have to classify unwanted spam and consume storage space and communication bandwidth. Other existing rules need to be continuously updated and maintained, which places greater burdens on some users and makes it difficult to manually compare the accuracy of the classified data.

IV. SCOPE OF STUDY

These intervals help focus on the project. The scope is: i. Modified existing machine learning algorithms. ii. Use and classify datasets, including preparation, classification and visualization of data. Data scores to determine the accuracy of spam detection. Use specific algorithms to learn classification rules from these messages. These algorithms are used to classify objects into different categories. The algorithm provides input and output data and has a self-learning program for solving a given task. Searching for the best algorithm and model can be very time consuming. Two types of classifiers are best used to classify the type of email, whether it is spam or non-spam. This algorithm is used to predict the probability and classification of date results.

ISSN : 2455-9679
SJIF Impact Factor : 6.008

Machine learning method

Machine learning is one of the most important and prominent methods for detecting spam comments, and it is usually divided into supervised and unsupervised learning low, the researchers discussed various machine learning methods that were proposed for registering spam comments.

A-Supervised learning -Monitored learning can be used to detect spam comments as a classification problem that divides comments into two categories: spam comments and non-spam comments. As far as we know, the first researchers to use supervised learning to study misleading opinion spam were Jindal et al. [twenty-one]. They discussed the development of opinion mining and focused on using natural language processing (NLP) to extract or summarize opinions from the text. Prior to their contribution, the textual content features that may indicate unusual activity (such as creating comment spam) have not been resolved. Efforts to investigate spam in comments and design spam detection technology.

B-Unsupervised learning- Since it is difficult to generate accurately labeled data sets for commenting on spam, the use of supervised learning is not always useful. Unsupervised learning provides a solution to this because it does not require labeled data. Raymond et al. Developed a new model without supervision of text mining and integrated it into a semantic language model to detect untrue comments. [1] And compare with the supervised learning method. Their model creates an approximation method using a semantic language model (SLM) to estimate the overlap of semantic content between reviews and calculate the degree of untruthfulness of reviews based on repeated recognition results. In addition to performing unsupervised spam monitoring, they also developed high-level concepts for association mining to extract context-dependent knowledge about concept association. Their model follows the hypothetical logic that if the semantic content of one comment is close to the semantic content of another comment, it is likely that the two comments are duplicates and therefore are examples of spam comments. In their experiment, they constructed a dataset based on real reviews collected from Amazon.

C-Semi-supervised learning - In other fields, it has been shown that the use of unlabeled data in combination with a small amount of labeled data can significantly improve the accuracy of the student compared to fully monitored methods [3]. In a study by Li et al. [9] using the framework of a collaborative training algorithm to exploit a large number of available unlabeled comments, a semi-monitored double-view method for detecting comment spam was created. The co-training algorithm, developed by Blum and Mitchell [4], is a guided method that uses a set of labeled data to incrementally apply labels to unlabeled data. It trains 2 classifiers on 2 different function sets and adds the most reliably labeled instance of each classifier to the training set. This makes it possible to efficiently generate large data sets and use them for classification, reducing the need to manually generate tagged training examples. A modified version of the co-training algorithm is also proposed, which only adds instances assigned to the same label by two classifiers. Their data sets were generated with the help of students. They manually marked 6000 comments collected from Epinions.com, of which 1394 were marked as spam. Four groups of comment-centered features were created: content, mood, product, and metadata. Two other groups of commenter-centered features were also created: profile and behavior.

V. Open Issues and Future Directions

This study showed that there are still some differences and unresolved issues in the research on the detection of spam comments. The main research gaps are described as follows:

Unavailability of labeled datasets

The lack of tagged datasets is an unsolved problem and a challenge in detecting spam comments. A tagged dataset of hotel reviews is available, but its number of attributes is limited. Researchers need access to standardized datasets to train classifiers to identify spam or non-spam comments.

The growing rate of review datasets

There are already millions of reviews on review-based sites, such as Amazon.com, and the number of reviews and reviewers is growing rapidly. Such a large data set involves high computing power for experiments [6], and the

ISSN : 2455-9679
SJIF Impact Factor : 6.008

realization of semantic algorithms is one of the biggest challenges in this field. The semantic analysis of words relies on WordNet and WordNet, both of which have a huge dictionary that can be used for sentiment analysis of comments. So far, no semantic based model has been proposed for registering spam comments.

Limited data attributes

Currently publicly available comment sets have limited attributes. This restriction makes it difficult for researchers to accurately detect spam comments. The biggest challenge here is the inaccessibility of the cube. Many researchers rely on datasets collected through crawling; However, such datasets also have limited attributes. To improve the accuracy of the algorithm, several attributes are needed, such as the spammer's IP address, the registered email address on the review page, and the place where the reviewer logs in to write the review.

Multilingual review spam detection

Comments are user-generated content, and users can post comments in any language of their choice. So far, few researchers have used languages other than English to process datasets, such as Arabic, Chinese, or Malaysian [4]. A thorough examination of spam detection in multilingual comments is necessary. 5. Identify spammers by analyzing feedback from other users on their written comments. To detect spam, researchers have made some progress by analyzing the content of the comments and the behavior of the comments. So far, the reviewer's personal profile information has not been used by any work. Normally, other users will have follow-up comments or comments on a given comment. For example, many sites will ask questions such as "Do you think this review is helpful?" So far, such feedback or comments on a given comment have not been used as a function to detect spam comments.

VI. CONCLUSION

This research conducted a systematic literature review in the field of detection of spam comments and emphasized recent research contributions in various functional engineering methods, methods for the detection of spam comments and various measurements for performance evaluation. To extract accurate and practical evidence, this work plans review methods that focus on search strings, ask research questions, select papers from well-known publishers, and apply formal evaluation criteria for inclusion and exclusion of research.

REFERENCES

1. Lau RY, Liao SY, Kwok RCW, Xu K, Xia Y, Li Y (2011) Text mining and probabilistic language modeling for online review spam detecting. ACM Trans Manage Inf Syst 2(4):1–30
2. Dixit S, Agrawal AJ (2013) Survey on review spam detection. Int J Comput Commun Technol ISSN (PRINT) 4:0975–7449
3. Ott M, Choi Y, Cardie C, Hancock JT (2011) 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-Volume 1 (pp. 309–319). Association for Computational Linguistics
4. López V, del Río S, Benítez JM, Herrera F (2015) Cost-sensitive linguistic fuzzy rule based classification systems under the MapReduce framework for imbalanced big data. Fuzzy Sets Syst 258:5–38
5. Bandakkanavar RV, Ramesh M, Geeta H (2014) A survey on detection of reviews using sentiment classification of methods. IJRITCC 2(2):310–314
6. Rutuja Katpatal;Aparna Junnarkar An Efficient Approach of Spam Detection in Twitter 2018 International onference on Inventive Research in Computing Applications (ICIRCA) Year: 2018 DOI: 10.1109/ IEEE Coimbatore, India
7. Sayali Kamble;S.M. Sangve Real Time Detection of Drifted Twitter Spam Based on Statistical Features 2018 International Conference on Information , Communication, Engineering and Technology (ICICET) Year: 2018 DOI: 10.1109/ IEEE Pune, India

ISSN : 2455-9679
SJIF Impact Factor : 6.008

8. Kshitiz Badola;Mridul Gupta Twitter Spam Detection Using Natural Language Processing by Encoder Decoder Model 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS) Year: 2021
9. Shivangi Gheewala;Rakesh Patel Machine Learning Based Twitter Spam Account Detection: A Review 2018 Second International Conference on Computing Methodologies and Communication (ICCMC) Year: 2018
10. Niddal Imam;Vassilios Vassilakis Detecting Spam Images with Embedded Arabic Text in Twitter 2019 International Conference on Document Analysis and Recognition Workshops (ICDARW) Year: 2019
11. K.Ushasree Santoshi;S.Sree Bhavya;Y.Bhavya Sri;B. Venkateswarlu Twitter Spam Detection Using Naïve Bayes Classifier 2021 6th International Conference on Inventive Computation Technologies (ICICT) Year: 2021
12. Surendra Sedhai;Aixin Sun Semi-Supervised Spam Detection in Twitter Stream IEEE Transactions on Computational Social Systems Year: 2018
13. Dani Gunawan;Romi Fadillah Rahmat;Arsandi Putra;Muhammad Fermi Pasha Filtering Spam Text Messages by Using Twitter-LDA Algorithm 2018 IEEE International Conference on Communication, Networks and Satellite (Comnetsat) Year: 2018
14. Wafaa Daffa;Omaimah Bamasag;Amal AlMansour A Survey On Spam URLs Detection In Twitter 2018 1st International Conference on Computer Applications & Information Security (ICCAIS) Year: 2018
15. E. Elakkiya;S. Selvakumar;R. Leela Velusamy CIFAS: Community Inspired Firefly Algorithm with fuzzy cross-entropy for feature selection in Twitter Spam detection 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT) Year: 2020