Fake Users and Reviewers Detection System

Mitali Ghoshal[1] , Yuvnaswa Singh[2] , T.Balachander[3]

[1,2] Undergraduate student, Department of Computer Science and Engineering, School of Computing, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Chennai, India
[3]Assistant Professor, Department of Computer Science and Engineering, School of Computing, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Chennai, India
[1]mr3890@srmist.edu.in, [2]yb1848@srmist.edu.in , [3]balachat2@srmist.edu.in

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 10 May 2021

Abstract
In recent years, social media plays a major role in user decisions using reviews, feedbacks. This allows everyone to leave a review, creating a prime opportunity for spammers to write spam reviews about goods and services of various interests. Trending Research Area: Identifying these spammers and the spam material is a key subject of study, and while a significant number of studies have recently been conducted to this end, the methodologies put forward so far barely identify spam feedback, and none of them demonstrate the value of each derived function category. Every spam detector system fails to identify this kind of spammers or at least has some trouble to spot them. In our proposed system, we analyse two categories effectively named review-behavioral and user-behavioral to identify spam reviewers. Using the available web services, we are going to build this application and provide efficient accuracy in detecting spam reviewers in social media forums.

Keywords— fake, identity, spammer, detection, behavior, reviews, comments

1. Introduction

Cash and cheque use are declining in line with global trends, as consumers continue to prefer digital payments. Users are embracing the convenience of 'contactless' cards, online shopping is expanding, and mobile payments are gaining traction. The steady growth of online shopping exemplifies the shifting payment landscape. Online sales continue to account for a growing share of total retail globally. Forecasts predict that by 2020, online shopping will have more than doubled since 2015, accounting for 14.6 percent of all retail sales. This pattern can be seen in the fact that online retail spending is rising five times higher than conventional retail spending.

The transition to online shopping has coincided with an increase in online payment fraud, which is in line with global trends. In 2019, card-not-present fraud accounted for 78% of all transaction card fraud[1]. Despite the increase of online fraud, industry-led anti-fraud measures are being implemented in a number of countries. This is reflected in the CNP fraud growth rate, which has slowed in the last two years. Online social media portals have a significant impact on information dissemination and are a valuable resource for both producers and consumers when it comes to choosing products and services. People rely heavily on written reviews in their decision-making processes as a general behavioural trait, with positive and negative reviews encouraging or discouraging them[2]. Furthermore, written reviews aid service providers in improving the quality of their goods and services. As a result, reviews have gained prominence as a determinant of a company's performance. While positive reviews can benefit a company, negative reviews can harm its reputation and result in financial losses.

Because anyone with any identity will leave a comment as a review, spammers have a enticing opportunity to write fake reviews in order to deceive users [3]. These false reviews are then multiplied by social media's sharing function and spread across the internet. Spam is defined as reviews written with the intent of changing a user's perception of the quality of a product or service.

2. The State of the Art

A variety of spam review detection techniques have been investigated in order to combat the issue of spam reviews. The literature on spam review detection using the spammer behavioral features analysis technique is reviewed in this study. This study also aims to assess the contribution by comparing this new work to previous research.

A study developed by (Sreeram Gutha et al.,2020) detecting fake account on social media using machine learning algorithms. The system collects the dataset which are preprocessed by providing a framework of algorithms using which we can detect fake profiles in Facebook by comparing the accuracy of three machine learning algorithms and the algorithm with very high efficiency is found for the given dataset.

In another existing study (S. P. Maniraj et al,.2019) fake Account Detection using Machine Learning and Data Science. To reliably pick out fake accounts, they use a gradient boosting algorithm with a selection tree containing 3 attributes - spam commenting, artificial behaviour, and interplay rate.

Another related work (P. Srinivas Rao et al,.2018) fake Profiles Identification in Online Social Networks Using Machine Learning and NLP. In this machine learning algorithms and techniques are used to enhance the accuracy rate of the fake profiles detection. Here, they make use of Support Vector Machine (SVM) and Naïve Bayes algorithm.

The study (Gayathri A et al,.2018) detecting Fake Accounts in Media Application Using Machine Learning. This paper proposed many classification algorithms namely like Support Vector Machine, Random Forest, Spam Filtering, Supervised and Unsupervised Machine Learning, Filtering and deep neural network to implement a multi-layer perceptron model.

Another related study (Samala Durga Prasad Reddy, 2019) fake Profile Identification using Machine Learning. This paper presents a technique for the automatic and powerful identification of fake accounts. This technique makes use of strategies along with the Random Forest Classification to segregate the profiles into fake or genuine groups. As it is an automated detection tool, it is straightforward to do for hundreds of thousands of accounts whose profiles can't be checked manually through the usage of social community sites.

The current study proposed a spam review detection model using spammer behavior analysis based on the gaps identified in the reviewed literature. It adds a number of spammer behavioral features to detect spammers, including a maximum number of reviews, an activity window, a review count, and the ratio of the first review. The proposed model is based on a weighted algorithm that detects spam reviews.

The CNP fraud rate was 68 percent in 2011 and has since risen to 78 percent in 2016. Spammers account for nearly 20% of Yelp users (a multibillion-dollar restaurant booking app Up to now, the methodologies presented have only identified a small percentage of spam reviews and neither has shown the importance of each functional form extracted[4][5]. Since anybody with any identity can leave a review, spammers have an easy way to trick users by writing fake reviews[11][13]. These fake reviews are then replicated and distributed around the internet thanks to social media's sharing feature [6][7].

As a result, there is no effective existing application tool to predict fake users spamming the app and attempting to undermine the app's user trust. Spam features are used as heterogeneous data networks in the current framework, for modelling analysis data sets, and spam detection in such networks is mapped into a classification issue[8][9]. With regard to various measures evaluated from the Yelp and Amazon Web pages, our performance can be improved with the use of spam value[10][12].

3. Objective of the proposed work

We now live in a time when social networks have become essential. From the last 20 years, online social sites have attracted millions of users. Teenagers nowadays prioritize having accounts on social media sites. Users of social media platforms should make new friends from all over the world. It also

aids in the transfer of video, audio, and messages, as well as the development of skills and the ability to learn new things.

Social media sites are also used by business and government organizers to improve the quality of their services. These platforms make life so much easier and more intelligent. Online social media platforms can be used for a variety of purposes. They do, however, have a dark side, as many malicious activities are carried out by fake accounts. As a result, an effective social media application that is regularly updated to detect fake accounts and prevent them from spamming the app has a higher chance of gaining user trust.

4. Proposed Methodology

In our proposed system, we design and develop an e-commerce application. In this we are going to detect and prevent ***spam reviewers*** *efficiently.*

Detection of spam reviewers is as follows:

1. Analyze reviewer behaviors and reviewer history (R-PPP-R).
2. Percentage of capital words used in the review.
3. Usually, spammers hide their identity for security reasons.
4. Review written on the same released date of the product /item / service (R-ETF-R).
5. Review with same number of exclamations (R-RES-R).

Detection of spam users is as follows:

1. Automated validation of user e-mail address.
2. Automated validation of user billing address.
3. Automated validation of user phone number.
4. Automatic obtaining and validating user IP address.
5. Automatic obtaining and validating user device ID.
6. Credit card number validation.
7. Manually checking user integrity by checking user social media profiles (i.e., user timelines, photos, friends & families).
8. Validating user business website.

Our application would provide a score analyzing all the above parameters. If the score is below the threshold, ***phone call verification*** would be performed to validate the user. This automation would reduce the human efforts in validating the users and avoid spamming to build a trust-based application online.

5. Software Requirements

JAVA:

The need for a platform-independent language to produce applications that could be incorporated in different consumer electronic devices was the primary inspiration for developing this language.

JVM:

The Java Virtual Machine is a computer abstraction that is the base of the Java platform. It is the security aspect responsible for the technology's hardware and operating system independence, limited compiled code size, and ability to shield users from malicious programs. It, like a true computer machine, has an instruction set and manipulates different memory areas at run time. It is fairly common to use a virtual machine to execute a programming language; perhaps the most well-known virtual machine is the UCSD Pascal P-Code machine. For security purposes, the JVM imposes strict syntactic and structural restrictions on code in a class file.

CSS:

Like XHTML is used to define the content of a web page, Cascading Style Sheets (CSS) is used to specify how the content should be displayed. Methods for tracking how records are viewed in formats other than a browser on a computer, such as print and mobile devices, are included. It also includes instructions for describing a document's non-visual appearance, such as how it sounds when read aloud by a screen reader. Style sheets are also an excellent tool for automating development because they help one to modify all of a site's pages by editing a single style sheet paper.

SERVLETS:

Java Servlets are Web or Application Server-running programmes that act as a bridge between Web browser or other HTTP client queries and HTTP databases and applications on the HTTP server. They are often used for the same purposes as programs that use the Popular Gateway Interface (CGI). Servlets, on the other hand, have a number of advantages over CGI. Performance has improved dramatically. Servlets can be used for user input by means of web page formats. Database or other source documents can be presented and web pages can be dynamically designed.

XML:

XML (Extensible Markup Language) tags are used to describe data and store and organize it. XML has 3 essential traits that make it beneficial in a extensive variety of structures and solutions:

- XML is extensible.
- XML incorporates the data, does not present it.
- XML is a public standard.

NETBEANS:

The official Java 8 IDE is Net Beans IDE. With the Net Beans IDE, the programs can be upgraded conveniently and efficiently to use modern Java 8 language constructions such as lambdas, operations and process references using our editors, code analyzers, and converters. Batch analyzers and converters are available to scan several applications at once for patterns that can be converted to the new Java 8 language constructs.

MY SQL:

MySQL is the world's second most common open-sourced relational database management system (RDBMS). MySQL is a usual web application database and a core component of the popular open-source LAMP Web Application Stack (and other 'AMP' stacks). LAMP is an acronym for "Linux, Apache, MySQL, and Perl/PHP/Python."

JSP:

The most important of many good reasons is that JSPs make creating complex websites extremely easy. Anyone who is able to write HTML will easily create websites that are rich, interactive and adaptive, allowing users to enjoy their online time. JSPs also allowed large groups or people, via a system known as JavaBean, to function in complicated tasks such that each part is easy and manageable, without losing control. Also, JSPs provide a great deal of flexibility in HTML generation because they are able to produce custom-like HTML tags.

SQL YOG:

SQLyog is a GUI tool for the RDBMS MySQL.

6. Architecture Diagram

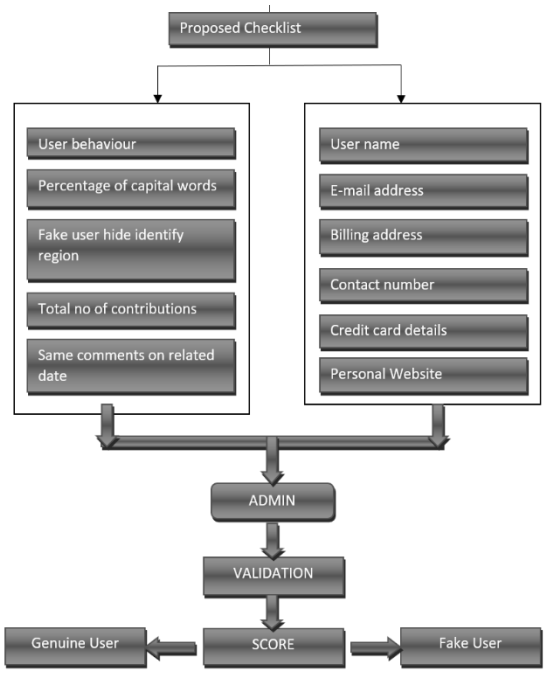


Figure 1. Fake News and Reviews Detection System

7. Modules

7.1 Checking the User Behavior

The user's previous activities are examined. The user's name (first and last name, as well as an optional middle name) is checked by comparing it to Facebook, LinkedIn, and other Google forums to see if the user is involved (checking images, contacts, timelines, and so on). If it's the user's first time leaving a comment, they'll be scrutinized more closely to see if they're a real user or a spammer. The term "review-early time frame-review" applies to comments that have the same release date as the object. The date is evaluated and kept track of if a product and review are introduced. Fake users often use '!' as much as possible in their sentences to create a favorable impact on users and to draw attention to their reviews among others as shown in Figure 1. To get the spotlight, many Fake users use capital letters.

7.2 IP and MAC Address Verification

The Internet Protocol (IP) address is a four-part string that identifies any device connected to the Internet. An IP address is made up of four numbers separated by intervals (each between 0 and 255). The MAC address is the manufacturer's unique identification identifier assigned to most network adapters or NICs. Both of these are tracked and verified before moving further in deciding whether the user is valid or not.

7.3 Personal Details Verification:

Validating phone numbers allows you to contact and verify your clients. Our powerful tool uses web services to provide reliable, fast results based on data from over 200 telecommunications providers. We can also clean up telephone numbers before entering our database to ensure they are correct. Our telephone validation tool ensures the exact number is collected in order to keep this precious channel open. The location and line type are also detected in order to guarantee user legitimacy. Additionally, the customer's e-mail address is checked and validated. Email addresses are often requested as user identification on websites for data validation purposes. Although some organizations provide services to verify an email address at the time of admission, usually with the use of an Application programming interface, there is no guarantee that the results will be correct. The API endpoint is used to verify email addresses. Fast algorithms that predict the validity of email addresses are required by large websites, bulk mailers, and spammers. Heuristic algorithms and mathematical models are heavily used in such approaches.

In order to ensure that it is a confirmed person, the billing address is also an essential information that must be verified. A system for verifying an address for a person claiming to be the owner of a credit card is the AVS (Address Verification System). The system compares the billing address of the customer's credit card against the file address of the credit card company. AVS ensures that the credit or debit card billing address corresponds to the customer's address. In our project we verified billing address using library in the Google Maps JavaScript API.

8. System Implementation

We have designed an e-commerce application which is going to be maintained by an admin, which can be a single entity, a bot, or a sales management team of an organization. The admin can login using the admin username and password and then upload new products (be it for his own organization, or partners who want to advertise their product on the online forum), write the product description. The admin has the authority to upload new products, and then specify the product type, the rate, a basic description and the product image. Once uploaded, the admin can see the list of uploaded products in addition to the list of registered users of the application. Any new customer needs to first register their names with the application. If they have already registered, they can directly login with a username and password and can see all their details on the profile tab. Once they are logged in, they can browse through the various products that are available.

$$\%Accuracy = \frac{All\ correctly\ identified\ accounts}{Total\ number\ of\ accounts} \times 100$$

In case a customer views a project and intends to buy it, they can add it to the cart, which then takes them further to a new booking page, wherein the customer can type in the payment details like email address, credit card number and the associated bank, billing address and so on. All the orders that have been placed can be viewed in the booking details tab, along with the status of shipment. For any queries or issues, the customer can contact customer support thorough email or the helpline number.

Process Flow

1. Select the profile to be tested.
2. Extract the features required for instance,
    a) user behaviour
    b) region
    c) number of contributions
    d) all-capital words percentage
    e) suspicious comments
    f) Username
    g) Email-address
    h) Contact details
    i) Credit card details
    j) Personal website, if exists
    k) Pass through the trained classifier.
    l) Determine whether the profile is genuine or not.
    m) IP address
    n) Count of suspicious activities
    o) Number of friends and followers
3. Pass through the trained classifier.
4. Determine whether the profile is genuine or not.
5. Depending on the evaluation, report if spam or not.

9. Results and Discussion

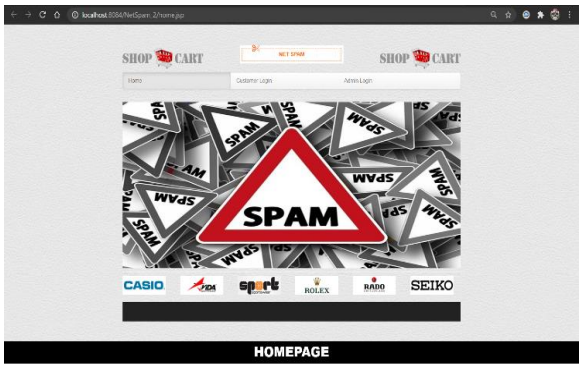


Figure 2. Experimental Results

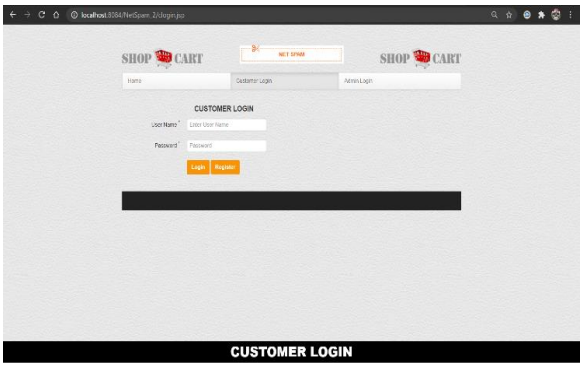
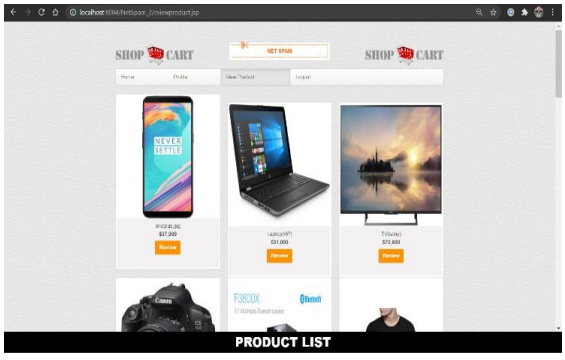


Figure 3. End User Login



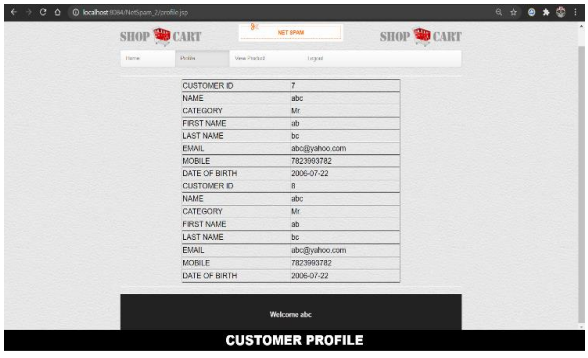


Figure 4. Product Listing

Figure 5. End user profile details

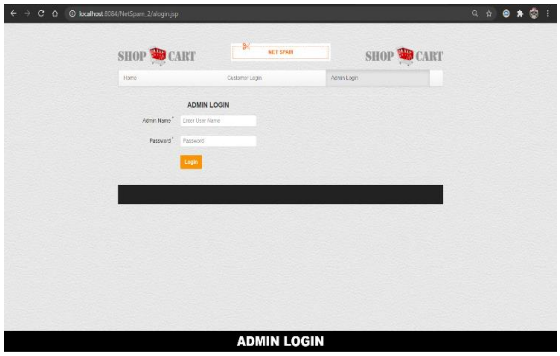


Figure 6. Admin login to validate end users

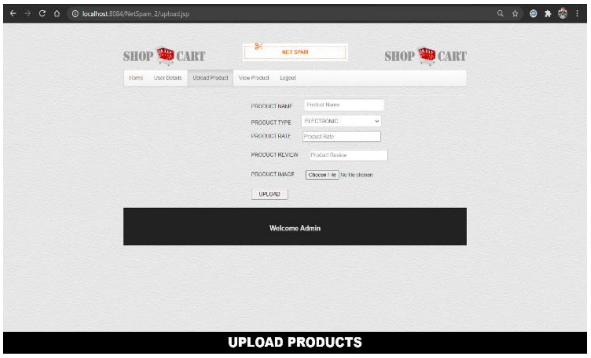


Figure 7. Upload products

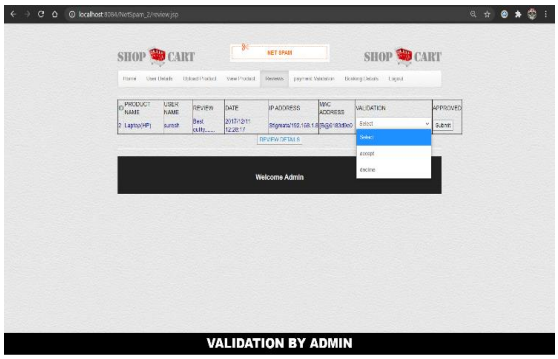


Figure 8. Validating the end users based on the checkpoints

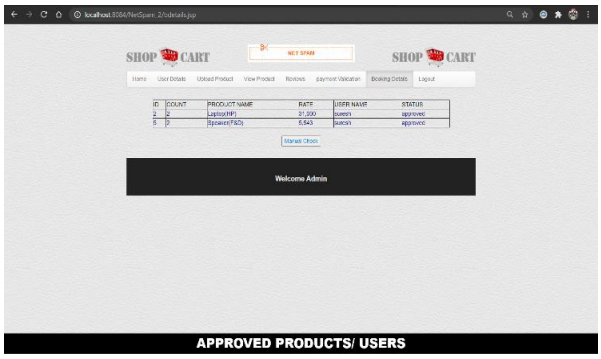


Figure 9. Result of validation

Our proposed system is very successful in detecting fake user reviews, which leads to improved results and is shown in the Figure from 2 to 9. In addition, we are really good at detecting fake users. As a result, our proposed framework offers an effective and precise architecture for detecting and preventing fake users.

Experimental Parameters

Table 1 shows the number of genuine and fake users and Table 2 represents the parameters accuracy, false negative and false positive.

Total Users	Genuine Users	Fake Users
30	23	7

Table 1. Performance of our proposed system

Accuracy	False Positive	False Negative
0.914	0.111	0.001

Table 2. Accuracy of our proposed system

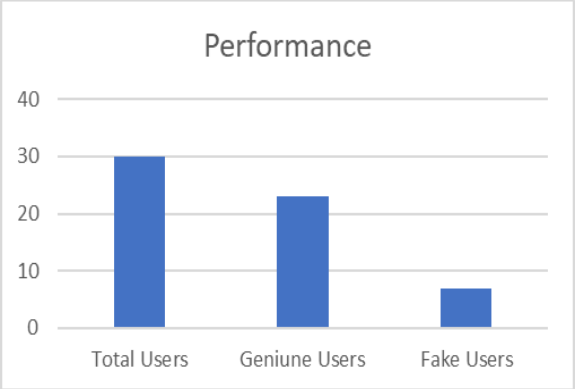


Figure 10. Performance of our proposed system

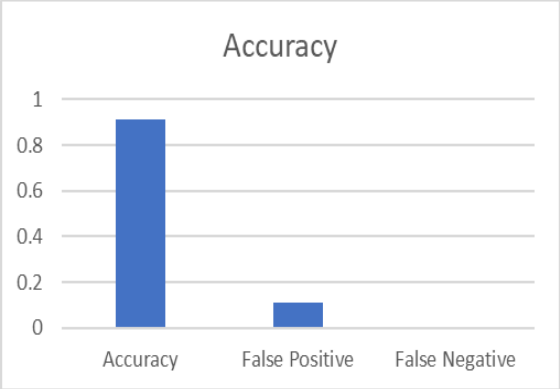


Figure 11. Accuracy of our proposed system

The performance and accuracy of the proposed system are shown in the Figure 10 and 11.

10. Conclusion

We examine the behavioral qualities of reviewers in this paper. We also verify their email ids, phone numbers, addresses etc. to check whether or not they are bots, and if not, whether they are legal users. Spammers can hardly be held responsible for spamming because they are able to disappear (go offline) immediately after they have spammed the recipients. Our findings confirm that a large number of spammers only spam in a short period of time, and more alarmingly, some sophisticated spammers use short-term spam networks. We have found that in order to control spam effectively, we need to hold spammers liable, force them to stay online for longer periods while speculating and limit the flexibility of spammers to change their locations frequently and/or Internet service providers.

11. Future Work

The metapath principle can be extended to other problems in this field in the future. A platform like this can be used to find spammer communities, for example. Reviews can be linked together using group spammer features to find groups, and reviews with the highest correlation based on the metapath concept are referred to as communities.

References

1. [1] Nambouri Sravya, Chavana Sai praneetha, S. Saraswathi," Identify the Human or Bots Twitter Data using Machine Learning Algorithms", International Research Journal of Engineering and Technology (IRJET), Volume: 06 Issue: 03 | Mar 2019 www.irjet.net, e-ISSN: 2395-0056, p- ISSN: 2395-0072.
2. [2] M. Smruthi, N. Harini," A Hybrid Scheme for Detecting Fake Accounts in Facebook", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7, Issue-5S3, February 2019.
3. [3] Tehlan, Pooja, Rosy Madaan, and Komal Kumar Bhatia. "A Spam Detection Mechamism in Social Media using Soft Computing."
4. [4] Rao, P. S., J. Gyani, and G. Narsimha. "Fake profiles identification in online social networks using machine learning and NLP." Int. J. Appl. Eng. Res 13.6 (2018): 973-4562.
5. [5] Raturi, Rohit. "Machine learning implementation for identifying fake accounts in social network." International Journal of Pure and Applied Mathematics 118.20 (2018): 4785-4797. J. Wang, "Fundamentals of erbium-doped fibre amplifiers arrays (Periodical style—Submitted for publication)," IEEE J. Quantum Electron., submitted for publication.
6. [6] Van Der Walt, Estée, and Jan Eloff. "Using machine learning to detect fake identities: bots vs humans." IEEE Access 6 (2018): 6540-6549.

7. [7] Kulkarni, Sumit Milind, and Vidya Dhamdhere. "Automatic detection of fake profiles in online social networks." Open access international journal of science and engineering 3.1 (2018): 70-73. M. Young, The Techincal Writers Handbook. Mill Valley, CA: University Science, 1989.
8. [8] Ala'M, Al-Zoubi, Ja'far Alqatawna, and Hossam Faris. "Spam profile detection in social networks based on public features." 2017 8th International Conference on information and Communication Systems (ICICS). IEEE, 2017.
9. [9] Elovici, Yuval, and Gilad Katz. "Method for detecting spammers and fake profiles in social networks." U.S. Patent No. 9,659,185. 23 May 2017.
10. [10] Gurajala, Supraja, et al. "Profile characteristics of fake Twitter accounts." Big Data & Society 3.2 (2016): 2053951716674236
11. [11] T.Balachander, et al. " Machine Learning Pipeline for an Improved Medical Decision Support", International Journal of Advanced Science and Technology, Vol. 29, No. 6, (2020), pp. 2632 – 2640.
12. [12] T.Balachander, et al. " "Website vulnerability detector", International Journal of Advanced Science and Technology, Vol. 29, No. 6, (2020), pp. 2444 – 2450.
13. [13] Dr.M.B.Mukesh Krishnan, et al. " Agent Based Trust Estimation for Mobile Ad Hoc Network", Indian Journal of Science and Technology, Vol 8, I S9, B 223-E 227.