International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8, Issue-1, May 2019

Predilection Decoded: Web Based Spam Detection And Review Analysis For Online Portals

Anam Jawaid, Saima Dev, Radhika Sharma, Veena G.S

Abstract: ***A vast majority of people depend on pre-existing information available on social media to aid them in their decisions. The most common being: Reviews on various products available in the market. With internet services being provided to any and every human being, there are certain drawbacks with such as leaving negative or disingenuous reviews about various products and services offered on internet platforms varying in interests. The classification and determination of such spammers along with the spam content is quite growing topic for analysis and more deep research. A substantial quantity of researches have been carried out regarding this topic, however, the methodologies that have been presented are of high complexity and do not have an easy to use interface for the same. In this research paper, we put forth a simple yet highly effective framework that uses basic algorithms of cosine similarity and sentiment analysis, to implement a web-based model for spam and fake review detection. We segregate the comments as fake, meta-fake and genuine reviews. Sentiment Analysis, Negative Ratio Checking and Cosine Similarity are used for detection of fake reviews and spam content along with other examinations. Incorporating changes based on customer feedback is one of the most important activities carried out by product designers. Spam detection and fake review identification can help an organization analyze, improve and enhance their product based on the suggestions in the real classified reviews given by the customers. If this information is made public by the organization, people can decide whether to buy the product or not based on the real reviews that have been identified by the system.***

Keywords: Spam detection, Dataflow Diagrams, Datasets, Cosine Similarity, Negative Ratio Text, Bar graph, Pie Chart, Meta Fake review table, Database, Review Analysis, Bias Detection, Spam Detection, Java Server Page, Web interface, User interface

I. INTRODUCTION

In these modern technological times, various social media portals act as an imperative part for conveying various information, which is vital for various producers in carrying out advertising campaigns for their products and services. The content available on various social media platforms comes in handy and plays as a deciding factor in the manufacturing processes of various such products. This in turn paves an important path of Globalization. Ever since the advancement of social media and various online portals, people have been relying mostly on such available reviews for selecting their required products, During the deciding phase, the individual looks for positive and negative reviews which further help them to be encouraged or discouraged to show further interest in the product or service. In addition, written reviews also help service providers to enhance the quality of their products and services. These reviews also hold the responsibility of securing individuals' visibility online in various organic search rankings. Their importance increases day by day which causes a growing number of competitors to start engaging more in order to persuade more customer reviews which would further increase their reputation in the online community. Online reviews, thus, have become a deciding circumstance in the boom of various business chains. As we've studied that positive reviews can reap in fruitful benefits for a business, the same way, unfavorable reviews can face a direct hit at the image of a company and cause unwanted losses. Keeping in mind the harsh reality that individuals can post reviews anonymously, provides an enticing freedom for spammers to post their fake reviews which are designed to fool the users mind set. These disingenuous reviews are then multiplied in incalculable digits and spread over the world wide web with the help of the extensive power of various social media outlets. In this proposed method, the product detects reviews by five major steps. This five layered modelled has been designed to identify spam in reviews given by people on social media or online platforms. We put forth a novel framework that exploits the spam contents for aided review of the modelling of datasets in their heterogeneous form which further utilizes this information to map a spam detection methodology into a classification issue. We segregate various comments as fake and genuine reviews. Sentiment Analysis is performed on the genuine reviews using n gram technique and cosine similarity. Spam identification and sentiment analysis on the reviews can help an organization analyze, improve and enhance their product based on the suggestions in the reviews given by the customers. People can decide whether to buy the product or not based on the reviews given by the customers who have bought the product.

Revised Manuscript Received on 30 May 2019.
***** Correspondence Author
Anam Jawaid*, Department of Computer Science and Engineering, Ramaiah Institute of Technology, Bangalore, India
Saima Dev, Department of Computer Science and Engineering, Ramaiah Institute of Technology, Bangalore, India
Radhika Sharma, Department of Computer Science and Engineering, Ramaiah Institute of Technology, Bangalore, India
Dr. Veena G.S, Department of Computer Science and Engineering, Ramaiah Institute of Technology, Bangalore, India

Our proposed system provides clients with a user-friendly interface which can be easily understood and comprehended. The methodologies used by the back system are simple yet effective.

II. LITERATURE SURVEY

Amir and Bin propose a model that predicts whether a review is a deceptive positive review using classifier that will use two sets of features. The first feature Developed-LIWC (Linguistic Inquiry and Word Count) [1] is utilized to examine around eighty characteristics of wordings yielding output as a feature. The second feature is the Unigram feature [1], in the sense it will evaluate the frequency of each word in a review. The paper also uses SVM to classify reviews based on extracted features and Chi-Square to select a subsection of features for classification which effectively reduces number of features of the dataset.

Somayeh and Azreen propose a framework [2] which uses two sets of questions one each for the reviewer and the review to detect the bias in review. It uses an annotation system [2] and an annotator to answer clues for reviews and reviewers. The features are selected based on its existing work. If the threshold for the review and reviewer are greater than what has been predefined then the review is categorized as fake. Abdullah and Resul provide a detailed survey [3] and useful insights on various spam detection methods that are used by the most popular social media site i.e. Twitter. They have been categorized into account based, tweet based, graph based and hybrid-based method and features [3]. The analysis of these methods shows that most of the old methods are likely to be exploited by attackers and hence a system to detect fake reviews and biases in reviews should apply methods which are different and simple yet effective. Some of methods [3] and features explained include Account based features and method which are User name, biography, creation date, likes etc. are controlled by the user or not and approaches such as honeypot, Tweet based features and methods such as sender, hashtags, links etc. are user controlled or not and approaches such as URL Filtering. Graph based features and methods such as distance and connectivity between two users is user controlled or not using data structures and extraction and Hybrid methods which are combination of the methods. An alternate strategy is adopted, by Geli and Arjun, which uses the burstiness idea [4] of surveys to recognize audit spammers. Blasts of surveys can be either because of sudden notoriety of items or spam assaults. Commentators and surveys showing up in a burst are regularly related as in spammers will in general work with different spammers and veritable analysts will in general seem together with other authentic commentators. This prepares for authors to manufacture a system of analysts showing up in various blasts. At that point it is shown commentators and their co-occurrence in blasts as a Markov Random Field (MRF), and utilize the Loopy Belief Propagation (LBP) technique [4] to gather whether an analyst is a spammer or not in the diagram. Additionally, a few highlights are proposed and utilized which include instigated message going in the LBP system for system deduction. Further a novel assessment technique is proposed to assess the identified spammers consequently utilizing directed grouping of their surveys. Furthermore, area specialists are utilized to play out a human assessment of the recognized spammers and non-spammers. Both the characterization result and human assessment result demonstrate that the proposed strategy outflanks solid baselines, which exhibit the viability of the technique. Stamper, a methodology [5] is proposed for distinguishing altered group calculations that essentially increases current standards for avoidance by versatile aggressors. Stamper configuration depends on two key experiences: First, Sybil assault discovery picks up quality in numbers: we propose factual examination strategies that can decide whether a huge group calculation has been altered by Sybils, not with-standing when it is in a general sense hard to deduce which of the partaking personalities are Sybil. Second, Sybil characters can't fashion the timestamps of their exercises as they are recorded by framework administrators; Stamper breaks down these unforgeable timestamps to thwart versatile aggressors. A connection was established to Stamper to identify altered calculations in Yelp and Twitter. Not only recently known altered calculations were distinguished with high precision, yet additionally a huge number of beforehand obscure altered calculations were revealed in these frameworks. Congrui and Quienchang propose a method to improve the quality of a blog website. They have built their own heuristic model [6] where they have considered features like, length of the comment, similarity between comments, difference between texts using KL Divergence and finding popular words and propaganda. They have employed a decision method to build a tree-model identical to decision tree. After obtaining information of every comment, they use a statistical method to help in classification of the comment. The most fit feature is selected as the tree's root. They study the skewness coefficient of the circulation on every single feature to choose the best feature. Thus, the higher the skewness coefficient is, the best fit the feature is. The YouTube comments section has a feature [7] where the comments are marked as 'hasSpamHint'. They noticed that this feature correctly detected some spam comments and ignored the others. This might have happened due to infeasibility of manually analyzing voluminous comment data and also classifying a comment using 'hasSpamHint' is performed at comment level and not user level [7]. They proposed an approach where they extract and retrieve the comments already marked with hasSpamHint for a given video. They extract the user-id of these spam comments. YouTube API allows you to retrieve recent commenting activity. They have extracted some metadata such as timestamp, video-id, text of comment and so on. They have defined 4 indicators to identify spam and the value of each of the indicators can identify spam users. The 4 indicators [7] are: ATDC(Average Time Difference between comments) which should be low to be detected as spam, PCHF(Percentage of Comments that have hasSpamHint flag-if many users have high PCHF value they're classified as Spam users,

CRAV(Comment Repeatability Across Videos)-If the user posts the same comment multiple number of times in the same video or across different videos then they are spam users, CRR(Comment Repetition and Redundancy)- If the user posts the same comment multiple number of times in the same video or across different videos in small time intervals or well-spaced out intervals, even then they are classified as spam users.

Rohini and Sheikh propose a technique to identify and control spamming. The first step is preprocessing [8] which is used to filter the data and involves steps like lexical analysis, removal of punctuations, symbols and stop words. The second step is feature extraction which includes removal of repeated comments and slang words. Feature extraction [8] is nothing but a spam control task. They have used an Iterative algorithm to construct Domain Space Dictionary. They have acquired 15 most similar words and compared semantic similarity between them. Create a separate category where words begin with s_ as they belong to spam. Eliminate those words of candidate dictionary if they are present in vulgar, AD dictionary or Domain Space dictionary.

The proposed system provides clients with a user-friendly interface which can be easily understood and comprehended. The methodologies used by the back system are simple yet effective.

III. SYSTEM DESIGN

A. Product Overview

The product detects reviews by five major steps. First, it retrieves the data in the form of an excel file .the file content will be sent to the server via URL in the form of multipart, in the server-side servlet receives the file content and write the file content in the folder of the application. From that folder it reads the file content and store the file content in to the database. Followed by this is first phase of fake review detection. In this process, Data will be read from the database and checked whether the IP_Address and UserID is fake or not based on the metadata table and insert the fake reviews in to the fake review table.

Additional check performed is to see whether the number of reviews from the IP_Address are exceeding the threshold limit. Consequently, in the second phase, reviews will be read from the Real reviews table, considering each review with intense processing of over seven levels and calculations are made. If any user exceeds total percentage threshold, that user is considered fake and is inserted in to the meta fake user table. Once the analysis is complete report generation phase is carried out with generates five detailed reports.

These reports are finally mapped into graphs in the graph generation phase which clearly distinguish real reviews from those that are fake. Figure 3.1 shows the overall architecture of the system and figures 3.2, 3.3 and 3.4 show different levels of dataflow diagrams. These provide the details of each input and each output of every entity in the system. At increasing levels, more number of technical details and workflow have been depicted. It proves the simplicity yet the effectiveness of the architecture that is developed and hence less processing overhead with good performance.

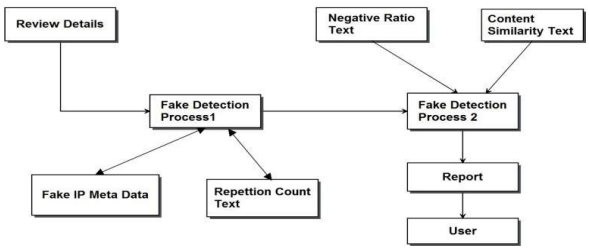

Fig 3.1: System Architecture

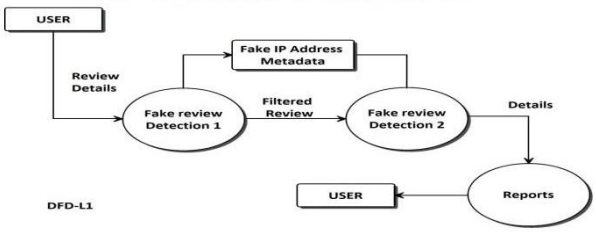

Fig 3.2: Data Flow Diagram Level 1

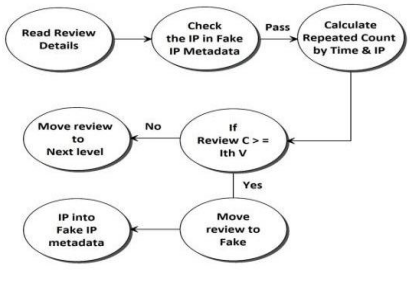

Fig 3.3: Data Flow Diagram Level 2

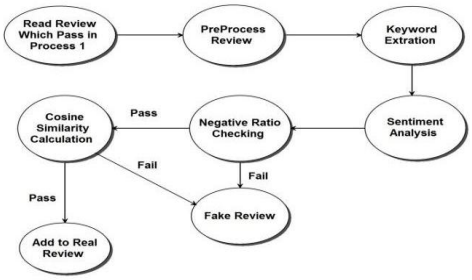

Fig 3.4: Data Flow Diagram Level 0

B. MODULAR DESCRIPTION

i. Data Retrieval

The system should be able to retrieve data in the form of an excel file successfully from its client. This forms the core of the system features as other system features work after successful data retrieval in the right format. Following are the sequence of actions required to be carried out by this system in this phase:

- User selects the file from the client machine.
- File content is sent to the server via URL
- Server-side servlet receives the file content and writes the file content in the folder of the application
- From that folder it reads the file content and store the file content in to the database.

ii. Fake Review Detection Phase I

Data will be read from the database and the system checks whether the IP_Address and UserID is fake or not based on the metadata table. Fake reviews are inserted in the fake review table. It also checks whether the number of reviews from the IP_Address are exceeding the threshold limit or are within the threshold time limit. Following are the sequence of actions required to be carried out by this system in this phase:

- Check the metadata table
- Check the IP_Address and UserID
- Insert the fake reviews in the fake review table

iii. Fake Review Detection Phase II

Reviews will be read from the Real reviews table, considering each review and a seven-level processing is applied. Steps 1 to 5 correspond to implementation of cosine similarity and 6 and 7 correspond to checking negative ratio. Following are the sequence of actions required to be carried out by this system in this phase:

- Unnecessary words and special characters are removed
- Categorize each word into noun or adjective
- Pair the noun and adjacent adjective
- Check whether the adjective which is paired with the noun is negative or positive
- Check whether the maximum number of pairs are positive or negative, based on the maximum count of positive or negative, assign the review value as positive or negative
- Calculate and insert the two gram and three-gram pairs in to the database
- Calculate the count percentage, positive percentage and n-gram percentage of each user and add all the percentages and get total percentage threshold, if any user exceeds total percentage threshold, consider that user fake and insert that user in to the meta fake user table.

iv. Report generation

Five distinct reports are generated as a result of the processing. Following are different types of reports generated by this system in this phase:

- Report 1 is generated based on the reviews given by the specific user for the specific product.
- Report 2 is generated based on the reviews given by all the users for the specific product.
- Report 3 is generated based on the reviews given by the specific user for all the products.
- Report 4 is generated based on the fake reviews given by all the users for all the products.
- Report 5 is generated based on the reviews given by the Meta Fake users and Meta Fake IP Address for all the products.

v. Graph Creation

Two graphs are generated as a result of the processing. The two graphs generated by the system in this phase are described below:

- Graph 1 (Pie Chart) is generated based on the total number of fake reviews, real reviews and meta fake reviews given by all the users for all the products
- Graph 2 (Bar Graph) is generated based on the number of reviews (fake, real and meta fake reviews) V/s Products.

IV. IMPLEMENTATION

Java Server Pages can be embedded with static HTML or XML to generate dynamic content. JSP can take inputs through a database and multiple other sources and is compatible with any browser, thus making the web application more dynamic and robust.

Initially, the history of results of the previous review analysis performed are cleared or refreshed from the database so that each detection process is started afresh and is not linked to the previous results.

The first stage of fake review detection analyses the name of the product, IP address, date and time for a review and checks whether the IP Address and User ID for the same are fake or not based on the data given in the metadata table in the dataset. It also examines if the number of reviews by a particular IP Address have surpassed the threshold limit in the given the threshold time limit. If any IP Address surpasses the threshold limit, the reviews given by that particular IP address are added to the fake review table and that IP Address is added to the Meta Fake IP Address table and the remaining reviews are added to the genuine reviews table. In this manner, the meta fake IP address table is constantly updated with new IP addresses that are possibly spam addresses.:

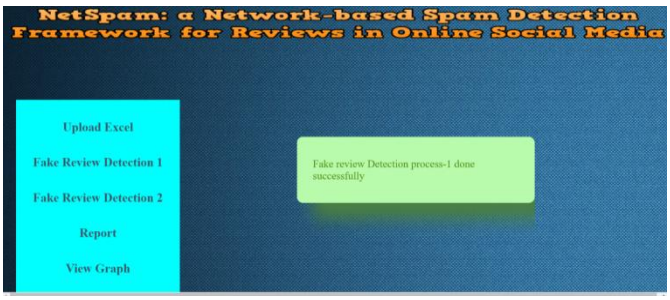

Fig 4.1: Execution of Module I

In the second stage of fake review detection, a detailed analysis of the content of the review is performed. This stage is added because spam cannot be detected only on the basis of IP addresses and the spammers can find new filters to escape being detected. A detailed sentiment analysis is performed for each review that has been inserted into the genuine review table. This process comprises of seven levels. The first level eliminates inessential words and characters.

The second level classifies every word as a noun or adjective. The third level groups a noun and adjoining adjective. The fourth level checks whether the noun-adjective pairs formed in the previous level are negative or positive. The fifth level checks the maximum number of pairs that are positive or negative. If the maximum number of pairs are positive, the review is positive else, negative. In the sixth level the two gram and three-gram pairs are determined and inserted into the database. In the seventh level the count percentage, n gram percentage and positive percentage of each user is determined and are added to get the total percentage threshold. If any user's review surpasses the total percentage threshold, the user is considered fake and their data is inserted into the meta fake user table.

V. RESULTS AND DISCUSSION

The results are depicted in the form of reports and graphs. There are 5 categories of reports namely:
1. One User One Product Report: This report lists the review made by one user for one specific product.
2. One Product All User Report: This report lists the reviews made by multiple users only for one specific product
3. One User All Product Report: This report lists the reviews made by one specific user for multiple products.
4. Fake Reviews by All Users: This report lists the spam reviews made by multiple users
5. Meta Data Fake Review Table: This report lists all the data like IP address, date, time etc. for all the fake reviews.

IPAddress	Date	Time	Product	Reviews	Userid
192.168.1.34	16-05-2016	12:10:12	B00171APVA	View Review	A1UQRSCLF8GW1T
192.168.1.34	16-05-2016	12:10:12	B00171APVA	View Review	A1UQRSCLF8GW1T

Fig 5.1: Report 1

IPAddress	Date	Time	Product	Reviews	Userid
192.168.1.16	16-05-2016	13:05:26	B00171APVA	View Review	A3SGXH7AUHU8GW
192.168.1.16	16-05-2016	13:05:27	B00171APVA	View Review	A3SGXH7AUHU8G1
192.168.1.16	16-05-2016	13:05:28	B00171APVA	View Review	A3SGXH7AUHU8G2
192.168.1.19	16-05-2016	13:05:29	B00171APVA	View Review	A3SGXH7AUHU8G3
192.168.1.10	16-05-2016	13:05:26	B00171APVA	View Review	A1D87F6ZCVE5NK

Fig 5.2: Report 4

There are 2 categories of graphs namely the pie chart and the bar graph that are generated by the system.

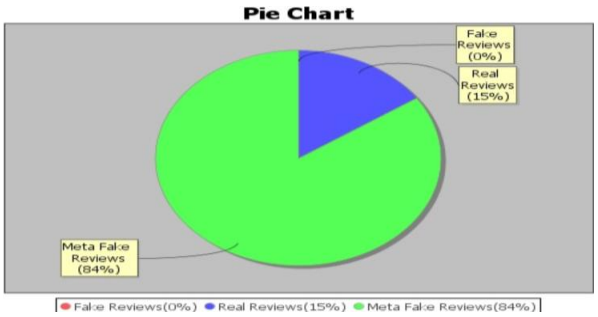

Fig 5.3: Pie chart

The pie chart in the figure 5.3 shows the representation of the amount of genuine reviews and fake reviews in a circular statistical form. Here, 84% of the reviews are fake and 15% are genuine.

Fig 5.4: Bar graph

From the bar graph in figure 5.4 we can infer that that the second product in the list has the maximum number of fake, meta fake and genuine reviews.

VI. CONCLUSION

Spam detection was simplified for clients in terms of use by providing a web-based model compared to windows-based models implemented before. Reviews could be efficiently classified and spam as well as fake reviews were clearly demarcated using the proposed methodology. The results generated gave useful and clear insights to the user with regard to percentage of fake reviews and most importantly the meta fake review table that contained the list of users who have been categorized as spammers. Various figures generated gave the pictorial representation of data analyzed and results obtained. The bar graph indicated the product which got the maximum number of real, fake and meta fake reviews which directly helps the organization understand which reviews are to be considered as real feedback. Unlike other systems, it proved to be more user-friendly and straightforward as well as accurate to its use. This web-based spam detection and review analysis can benefit companies and help them analyze their product, which gives them further scope for improvement whereas the customer can decide whether they should buy the product based on the genuine positive/negative reviews thereby making it a great platform for business and advertising.

Malicious users create havoc and try to spread fake news. The similar algorithm can detect fake news and block IP addresses and users from spreading fake news and analyzing the content of the news and segregating it into fake and genuine news.

REFERENCES

1. Amir Karam and Bin Zhou, University of Maryland Baltimore County,
2. "Online Review Spam Detection by new Linguistic features" Somyeh Shojaee and Azreen Azman, University Putra, Malaysia, "A Framework for Fake Review Annotation"
3. Abdullah Talha Kapakus and Resul Kara, International Journal of Advanced Computer Science and Applications, "A Survey of Spam Detection Methods on Twitter",
4. Geli Fei, Arjun Mukherjee, Bing Liu, Meichun Hsu, Exploiting Burstiness in Reviews for Review Spammer Detection"
5. Bimal Viswanath, Muhammad Ahmad Bashir, Muhammad Bilal Zafar, Simon Bouget, Saikat Guha, Krishna P. Gummadi, Aniket Kate and Alan Mislove, "Strength in Numbers: Robust Tamper Detection in Crowd Computation"
6. Congrui Huang, Qiancheng Jiang and Yan Zhang, "Detecting Comment Spam through Content Analysis"
7. Ashish Sureka "Mining User Comment Activity for Detecting Forum Spammers in YouTube"
8. Rohini D.Warkar and I.R.Shaikh, "Detection of Spam Comments using NLP Algorithms"
9. Dhinaharan Nagamalai , Beatrice Cynthia Dhinakaran and Jae Kwang Lee, "Bayesian based comment spam defending tool"
10. Satish Tukaram Pokharkar, Ajit Jaysingrao Shete, Vishal Dyandeo Ghogare, "Survey in Online social media Skelton by network based spam"
11. Tulio C. Alberto, Johannes V. Lochter, Tiago A. Almeida, "TubeSpam: Comment Spam Filtering on YouTube"
12. Draško Radovanović, Božo Krstajić, "Review spam detection using machine learning"
13. A. Heydari, Mhd. A. Tavakoli, N. Salim, Z. Heydari ,"Detection of review spam: A survey"
14. M. Crawford, T.M. Khoshgoftaar, J.D. Prusa, A.N. Richter, H. Al Najada , "Survey of review spam detection using machine learning techniques"
15. M. Ott, C. Cardie, J. T. Hancock,"Negative deceptive opinion spam"

AUTHORS PROFILE

Anam Jawaid Anam Jawaid is a final year graduate student pursuing Bachelor of Computer Science and Engineering from Ramaiah Institute of Technology. She is the author of "PATTERNS THAT DON'T EXIST: Study on the effects of psychological human biases in data analysis and decision making", International Conference on Computational System and Information Systems for Sustainable Solutions (CSITSS), IEEE. Her areas of interest include full-stack development, web development and software engineering.

Saima Dev Saima Dev is a final year graduate student pursuing Bachelor of Computer Science and Engineering from Ramaiah Institute of Technology. She is the author of "PATTERNS THAT DON'T EXIST: Study on the effects of psychological human biases in data analysis and decision making", International Conference on Computational System and Information Systems for Sustainable Solutions (CSITSS), IEEE. Her areas of interest include computer networks and compiler design.

Radhika Sharma Radhika Sharma is a final year graduate student pursuing Bachelor of Computer Science and Engineering from Ramaiah Institute of Technology. She holds good command over various prgramming languages and data structures and has a track record of exceptional delivery of code. Her areas of interest include algorithms, competitive programming and android application development.

Dr. Veena G.S Veena G.S. is working as an Assistant Professor in Computer Science, Department of Computer Science and Engineering at Ramaiah Institute of Technology. Her areas of interest include, image processing, embedded systems, mathematical modeling, cognition, IOT. She is the author of various publications including " Discovering Frequent Itemsets over Event Logs Using ECLAT Algorithm", International Conference on Soft Computing , Intelligent systems and Application (ASISA 2016) a Springer aise and "Kannada handwritten Word Conversion to Electronic Textual Format Using HMM Model", International Conference on Computational System and Information Systems for Sustainable Solutions(CSITSS), IEEE.