GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES
A REVIEW ON THE PRESENT STATE-OF-THE-ART OF ASTROTURFING AND ASTROTURFING GROUP DETECTION IN ONLINE SOCIAL NETWORKS

T. Ramdas Naik[*1] & P. Premchand[2]
[*1]Dept. of Computer Science Engineering, College of Engineering, Osmania University, Hyderabad, India
[2]Department of Computer Science & Engg, University College of Engineering, Osmania University.
Hyderabad, India

ABSTRACT
Reviews on Online Social Networks (OSNs) became crucial in decision making. People of all walks of life are using them to make decisions especially while purchasing products or services. The decision making paradigm has been shifted from traditional offline approach to technology-driven approach. The former has limitations in information while the latter is abundant in information. Dependence on the ratings or reviews on products online became higher than ever due to the availability of information and other advantages like saving time and effort. Digitalization in various countries in the contemporary era is also one of the reasons on the increased reliance on the online reviews. Through the technology –driven approach has many advantages; it also carried evil of probability of making inaccurate decisions due to the intentional spreading of fake reviews known as astroturfing. There is organized effort and an astroturfing group (group of fake reviewers) works behind spread of fake reviews to promote or demote services. In this context, this paper reviews the present state of the art of astroturfing (spreading fake reviews), detection of astroturfed reviews and astroturfing groups. The insights provided in this paper are useful to know different aspects of fake review detection research.

Keywords: *Online social networks, fake review detection, astroturfing, astroturfing group detection.*

I.   INTRODUCTION

Fake review detection became an essential requirement in the context of spreading fake reviews intentionally over OSNs. As online reviews are influencing people at large to make their decisions on purchasing products and services, this paper reviews the present state of the art of fake review detection methods and its related topics.

A. History of Online Social Networks and Emergence of Online Review Web Sites
As studied in [59] Online Social Networks (OSNs) that prevail now has their roots from the social media which can be traced back to 1978. From bulletin board system online to the emergence of online web sites that can be used for reviews on products and services (foursquare, YELP etc.) With the emergence of online review web sites, the users of OSNs make use of them to make decisions. Therefore reviews are able to influence the decisions of people.
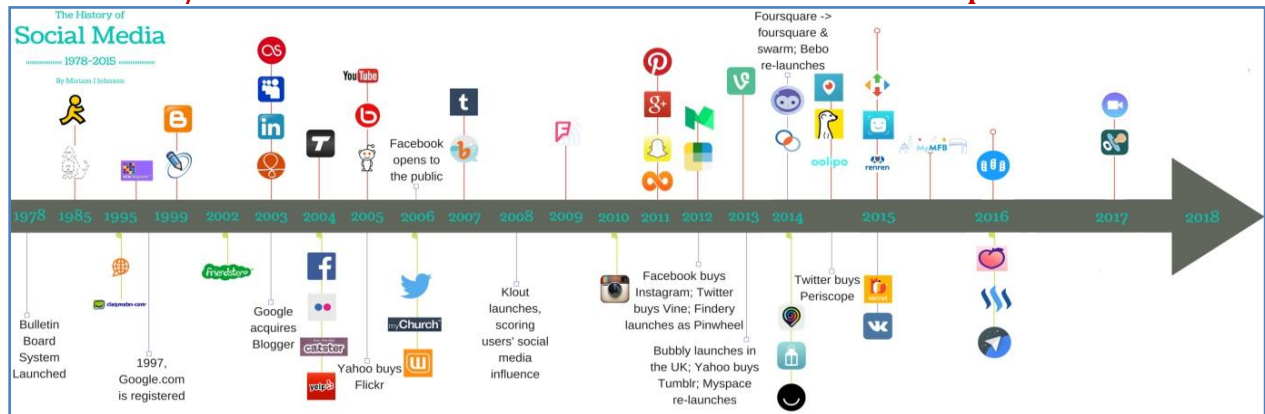
Figure 1: History of social networks from 1978 to 2017

As presented in Figure 1, it is evident that OSNs have evolved over a long period of time. Now we are in the era of Facebook, Twitter, and Instagram to mention few OSNs that are widely used. In the world of restaurants and hotels, YELP.com and foursquare.com are widely used OSNs that provide online reviews and micro-reviews respectively. OSNs, threats and possible solutions can be found in [2]. The potential of OSNs in the present communication era is studied in [8] and [16]. Privacy and security challenges in OSNs are explored in [12]. Importance of Facebook in social networking and its profile elements are investigated in [13]. The concept of online personas in social networks is reviewed in [14]. Corporate world and its usage patterns of OSN data is investigated in [37]. Social media provides many advantages to enterprises. For instance, social media data is analyzed and used for political activities in the form of sentiment knowledge [53]. Recommender system [17] is built for providing personalized recommendations to OSN users.

B. Significance of Online Reviews
As explored in [26], online reviews became very important source to make decisions for most of the people. The traditional approach in decision making is transformed into a technology dependent approach. The traditional approach is offline and it was based on asking other human beings before making a decisions. It has limitations in finding known people. With the emergence of OSNs, it is changed as millions of people are online and they do provide ratings or reviews to services or products.
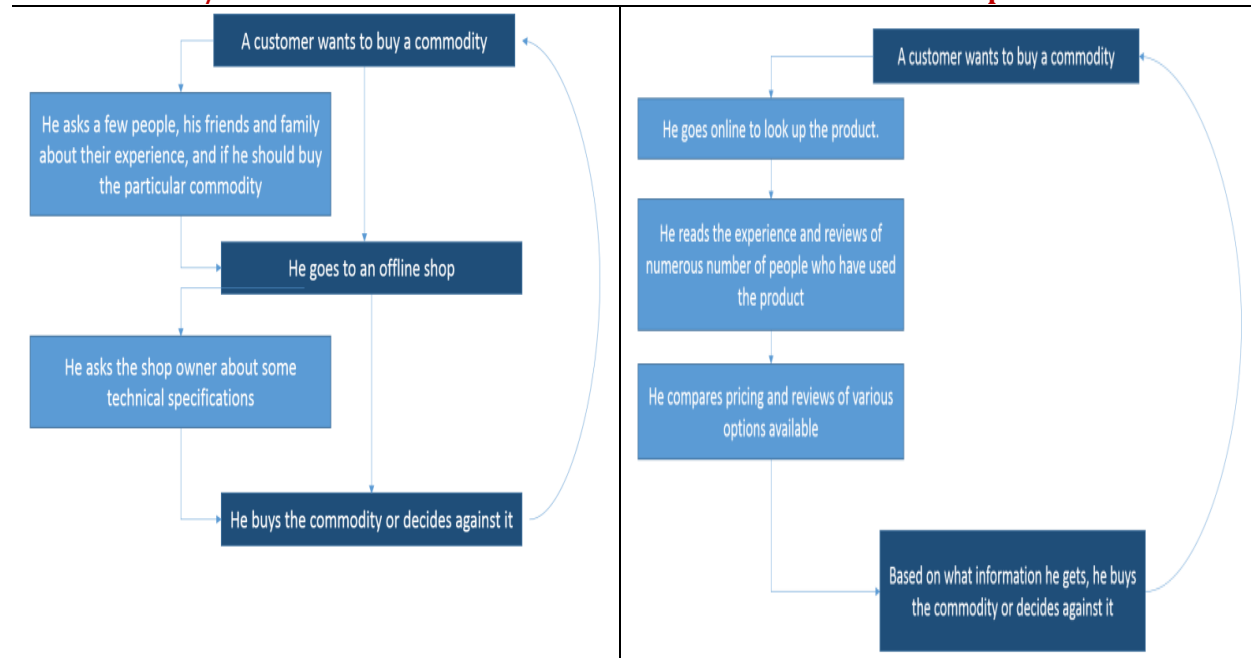
Figure 2: Tradition approach (a) and technology-driven approach (b) in decision making

As shown in Figure 2, there is difference in decision making in traditional and technology-driven approaches. In the technology driven approach, there are many advantages such as huge number of reviews to support decision making, quick and instant understanding of products and the trends in customer loyalty to such products or services to mention few. Time and effort of people are saved due to the online presence of knowledge used to make decisions with ease. At the same time there are many drawbacks when decisions are made on the reviews online. These issues are due to fake reviews made online with malicious intentions to promote or demote products or services.

Our contributions in this paper include the study of fake review detection literature and reviewing it besides providing useful insights. The remainder of the paper is structured as follows. Section 2 presents different types of online reviews. Section 3 focuses on fake online reviews and their negative impact. Section 4 provides characteristics of fake reviewers and reviews. Astroturfing versus astroturfing group is explored in Section 5 while Section 6 throws light into Amazon Mechanical Turk (AMT) which is used to generate fake reviews for conducting research in this area. Section 7 provides the need for fake review detection techniques while section 8 reviews various fake review methods. Section 9 presents datasets and SVM classification results. Section 10 concludes the paper besides giving directions for future work.

II. TYPES OF ONLINE REVIEWS

Online reviews are of two types. They are known as reviews and micro reviews. Reviews are the opinions given by people on products or services. On the other hand micro-review is a short review (may have a single line or two) or opinion. The popular web site for reviews is YELP.COM [39] while the micro-reviews can be found in foursquare.com [34]. The reviews online accumulate huge amount of data. And it leads to social big data which is voluminous and exhibit other characteristics of big data in general.

III. FAKE ONLINE REVIEWS OR ASTROTURFING AND THEIR IMPACT

In the contemporary era, product or service quality plays an important role which is crucial to generate demand among consumers. This psychological aspect of customers is considered by enterprises to generate fake reviews and

84

attract customers in making purchasing decisions. As people believe in other customers' ratings or reviews, fake reviews often attract and mislead people in making inaccurate decisions. Fake opinion spreading has its negative impact on the consumers and service providers in the long run. Its impact is huge as people of all walks of life started using ratings and reviews online before making decisions. Misleading human readers is the main reason for the fake reviews. There are many challenges in detecting fake reviews programmatically online. The reason is that it is not easier to say a review is fake or not. Section 4 presents characteristics of fake reviewers and reviews which can help in making detection models effectively.

IV.   CHARACTERISTICS OF FAKE REVIEWERS AND REVIEWS

From the study of [9] it is understood that fake reviewers or astroturfers and their reviews have certain characteristics. They include burstiness, negative ratio of reviews, average content similarity, maximum content similarity, early time frame, rate deviation, ratio of exclamation sentences and number of first person pronouns.
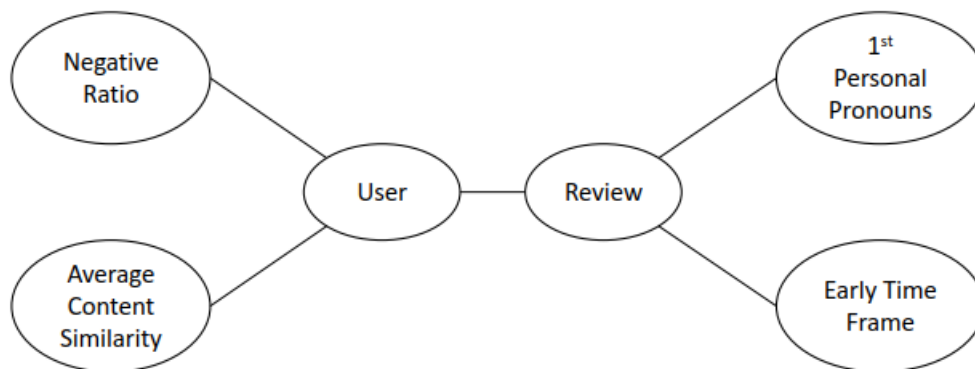

Figure 3: Shows characteristics of reviewers and reviews

Average content similarity indicates that the content of reviews is likely to have much similarity. The similar kind of reflection is there with maximum content similarity. The fake reviewers tend to write negative comments on the product of competitor. They use ! symbol as much as possible to generate curiosity among readers. They also use 1st personal pronouns and their work starts very early to be ahead in influencing decisions of people.

V.   ASTROTURFING VS. ASTROTURFING GROUP

Astroturfing is the process of making fake reviews online. Towards this end people who give fake reviews make use of web sites in social media such as YELP.com [39] and foursquare.com [34] to mention few. The process of spreading fake reviews is called astroturfing while the group of people behind the astroturfing process is called astroturfing group. There are many characteristics of astroturfing and reviews online as explored in this paper.

VI.   ROLE OF AMAZON MECHANICAL TURK IN THIS RESEARCH AREA

When online review web sites do not have filters, it is not possible to generate training datasets for fake review detection methods for supervised learning. To overcome this drawback, as explored in [3], Amazon developed a framework known as AMT for generating fake reviews. These fake review datasets can be used to have training set and build classifiers.

VII.   NEED FOR FAKE REVIEW DETECTION METHODS

Online reviews are widely used to make decisions. They are used by both organizations and individuals. The reviews that are genuine can have positive impact on the people who make decisions based on them. In fact there are many advantages of having reviews (truthful opinions) given by consumers of products or services. It leads to

saving time, fame and financial gain as well. Unfortunately there has been evidence of fake reviews being posted in OSNs that mislead people. This is called opinion spamming. This problem as become widespread as studied in [1] and [3]. Online social networks may have polluted opinions that may lead to inaccurate decision making. Therefore it is inevitable to detect fake reviews and the people behind spreading fake reviews besides making further efforts to ensure that only genuine reviews are spread.

VIII.   FAKE REVIEW DETECTION METHODS

According to [4] it is important to take fake online reviews seriously. Online credibility evaluation and heuristic approaches are studied in [5]. The detection of YELP reviews fraud is the focus in [6]. The significance of securing online content rating systems in order to avoid fake opinions from spreading is explored in [7]. Matching accounts of OSN users is the study in [20] while [21] show multi-model features in order to detect user accounts across the OSN. Hot topic detection [22] is studied and different topics models came into existence as discussed in [23]-[30]. Risk patterns are mined in [31]. Tourist related data is used for such research as in [32], [34], [35], [38], and [42]. Hotel guest reviews are explored for feature extraction in [33]. Different approaches on online fake review detection are explored in [45]-[58].

A. Summary of Fake Review Detection Methods

Table 1: Shows summary of fake review detection methods

Author	Technique	Advantages	Limitations	Datasets
Akoglu et al. [1]	FraudEagle framework with Network effects and scoring algorithm	Scalable, efficient, reveals fraud-bots	More priors to be explored.	SWM and synthetic datasets
Mukherjee et al. [3]	YELP fake review filtering	Filters deceptive opinion spam.	Accuracy needs to be improved.	YELP dataset
Shehnepoor et al. [9]	NetSpam frame work	Effective spam detection	Finding spammer communities is not explored.	YELP dataset
Kukherjee et al. [10]	Classification and analysis	Fake review detection	Ensemble method can be studied.	AMT generated dataset and YELP dataset.
Lai et al. [18]	Inferential language modelling and association mining.	Online review spam detection.	Improvements on the approach are required.	Amazon dataset.
Benerjee et al. [40]	Supervised learning	Detecting fake reviews	Needs further improvement	YELP dataset.
Shojaee et al. [41]	Fake review annotation.	Fake review detection.	Needs to detect fake review communities.	YELP dataset.
Azman et al. [44]	Aspect based contradiction	Detection of fake reviews.	Needs further improvement.	YELP dataset.

IX.   DATASETS AND CLASSIFICATION RESULTS

Datasets for fake review detection research are available in [39]. In the same fashion, micro reviews can be obtained as described in [34]. Effects of online review datasets are investigated in [11]. Datasets and SVM classification results [3] are as provided in Figure 4, Figure 5 and Figure 6.
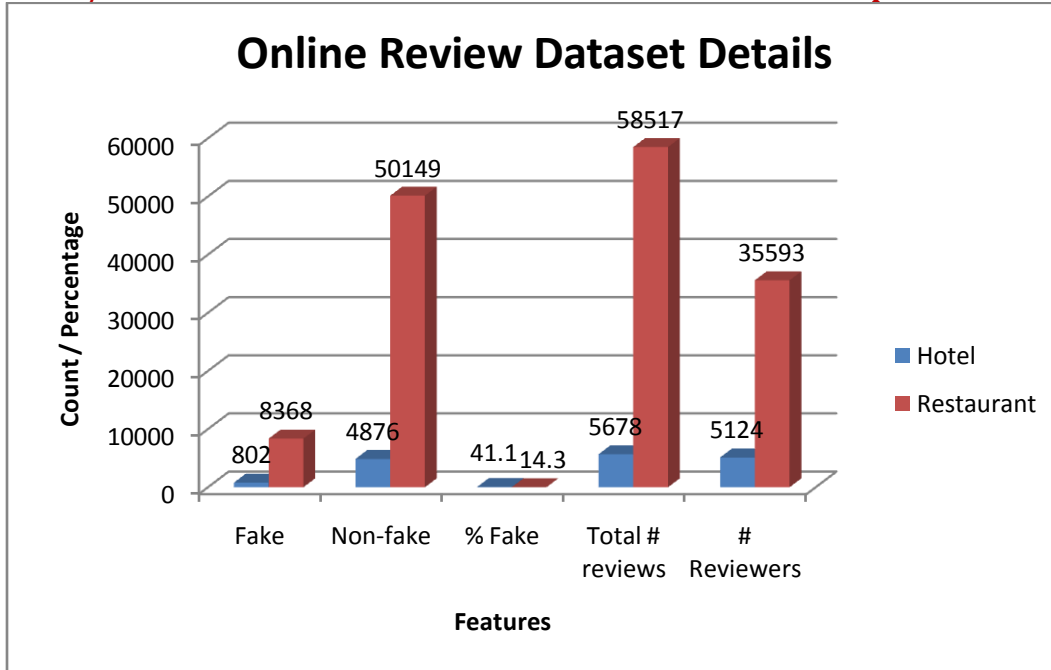
Figure 4: Dataset used for SVM classification

As presented in Figure 4, it is evident that two domains such as hotels and restaurants are considered for detection of fake reviews. The dataset statistics reveal the details of fake and non-fake hotels and restaurants, total number of reviews and total number of reviewers.
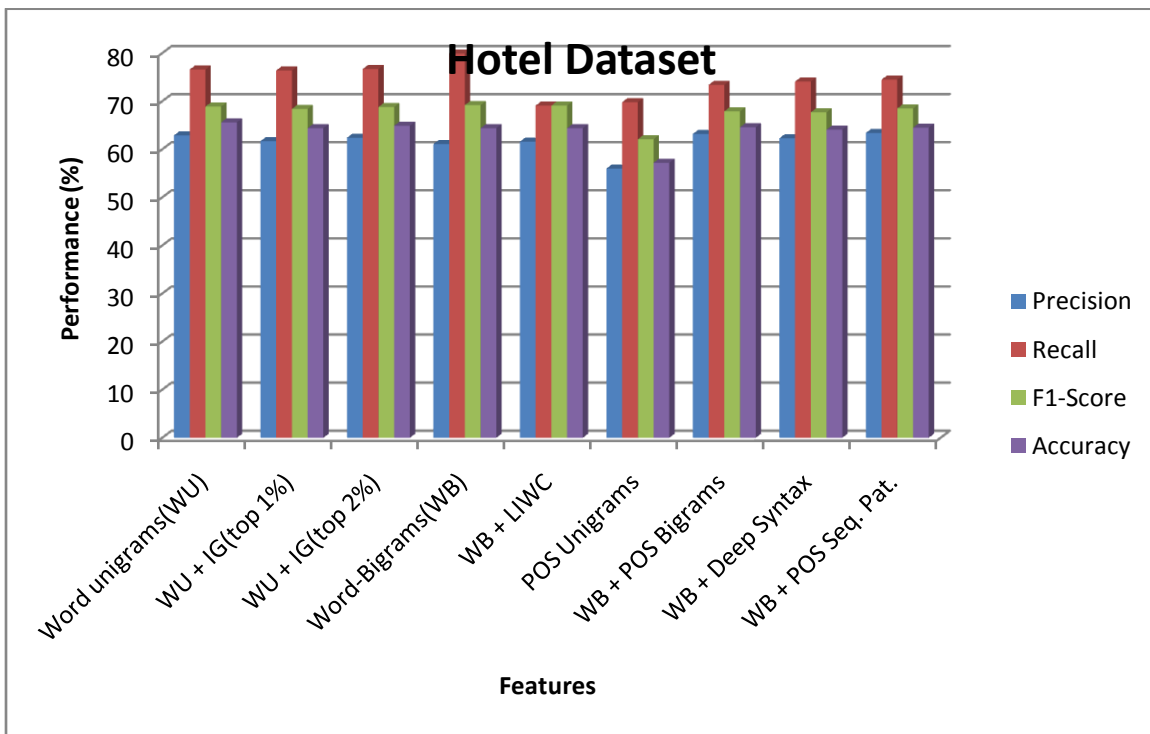

Figure 5: SVM five-fold cross validation results in Hotel domain

87

Figure 6: SVM five-fold cross validation results in Restaurant domain

As shown in Figure 5 and Figure 6, the results related to various measures like precision, recall, F measure, and accuracy are shows against different features. The results revealed the dynamics of the dataset in terms of two domains and associated fake and non-fake reviews.

X. CONCLUSIONS AND FUTURE WORK

In this paper, we made a review of fake review detection and methods to detect fake reviews. Fake reviews are made to promote products and services. The process of spreading fake reviews is known as astroturfing. Often it is done by well organized group of people hired by enterprises. Thus it is possible to detect the group behind the fake reviews. This group is known as astroturfing group. Astroturfing exhibits certain characteristics and it is possible to detect groups behind this activity. Different topic models are explored in the literature besides other fake review detection methods. It also shows the summary of detection methods, their techniques, merits, demerits and datasets used for experiments. The generative models provided in the literature are not sufficient to model the astroturfing groups. Therefore in our future work, we intend to enhance author topic models in order to have new generative framework for effective detection of astroturfing and astroturfing groups.

REFERENCES

1. *Leman Akoglu,Rishi Chandy And Christos Faloutsos, "Opinion Fraud Detection In Online Reviews By Network Effects", International Aaai Conference On Weblogs And Social Media, 2013,P1-10.*
2. *Michael Fire,Roy Goldschmidt, And Yuval Elovici, "Online Social Networks: Threats And Solutions", Ieee Communication Surveys & Tutorials,16 (4), 2014,P1-18.*
3. *Arjun Mukherjee, Vivek Venkataraman, Bing Liu And Natalie Glance, "What Yelp Fake Review Filter Might Be Doing", International Aaai Conference On Weblogs And Social Media, 2013,P1-10.*
4. *Justin Malbon, "Taking Fake Online Consumer Reviews Seriously", Springer Science+Business Media New York, 2013,P1-22.*

5.   Miriam J. Metzger, Andrew J. Flanagin, & Ryan B. Medders, "Social And Heuristic Approaches To Credibility Evaluation Online", Journal Of Communication. 60 , 2010,P413–439.
6.   Michael Luca,Georgios Zervas, "Fake It Till You Make It: Reputation, Competition, And Yelp Review Fraud", Acm, 2015,P1-36.
7.   Arash Molavi Kakhki,Chloe Kliman-Silver And Alan Mislove, "Iolaus: Securing Online Content Rating Systems", Acm, 2013,P1-10.
8.   Romina Cachia, Ramón Compañó, Olivier Da Costa, "Grasping The Potential Of Online Social Networks For Foresight", Technological Forecasting & Social Change, 2007,P1179–1203.
9.   Saeedreza Shehnepoor, Mostafa Salehi, Reza Farahbakhsh, Noel Crespi, "Netspam: A Network-Based Spam Detection Framework For Reviews In Online Social Media", Ieee, 2017,P1-10.
10. Arjun Mukherjee, Vivek Venkataraman, Bing Liu, Natalie Glance, "Fake Review Detection: Classification And Analysis Of Real And Pseudo Reviews", University Of Illinois At Chicago, 2013, P1-10.
11. Michael Anderson,Jeremy Magruder, "Learning From The Crowd: Regression Discontinuity Estimates Of The Effects Of An Online Review Database1", Acm, 2011,P1-20.
12. Chi Zhang , Jinyuan Sun,Xiaoyan Zhu,Yuguang Fang, "Privacy And Security For Online Social Networks: Challenges And Opportunities", Ieee Network, 2010,P1-6.
13. Cliff Lampe, Nicole Ellison, Charles Steinfield, "A Familiar Face(Book): Profile Elements As Signals In An Online Social Network", Acm, 2010,P1-10.
14. Michael Sirivianos,Kyungbaek Kim And Xiaowei Yang, "Facetrust: Assessing The Credibility Of Online Personas Via Social Networks", Ieee, 2010,P1-6.
15. Gemabello-Orgaz,Jasonj.Jung,Davidcamacho,   "Socialbigdata:Recent   Achievements   And   New Challenges", Elsevier, 2015,P1-15.
16. George Pallis, Demetrios Zeinalipour-Yazti, And Marios D. Dikaiakos, "Online Social Networks: Status And Trends", Springer , 2011,P213–234.
17. Xujuan Zhou, Yue Xu, Yuefeng Li, Audun Josang And Clive Cox, "The State-Of-The-Art In Personalized Recommender Systems For Social Networking", Springer , 2012, P1-15.
18. C.L. Lai, K.Q. Xu, Raymond Y.K. Lau,Yuefeng Li And Dawei Song, "High-Order Concept Associations Mining And Inferential Language Modeling For Online Review Spam Detection", 2010,Ieee, P1-8.
19. Susan Athey And Guido W. Imbens, "The State Of Applied Econometrics: Causality And Policy Evaluation", Journal Of Economic Perspectives. 31 (2), 2017,P3–32.
20. Oana Goga, "Matching User Accounts Across Online Social Networks: Methods And Applications" Hal, 2015,P1-151.
21. Ehtesham Hassan, Santanu Chaudhury, M Gopal And Vikram Garg, " A Hybrid Framework For Event Detection Using Multi-Modal Features", Ieee, 2011,P1-6.
22. Hui-Fang Mat, Yue-Xin Sunt, Mei-Hui-Zi Jiat And Zhi-Chang Zhang, "Microblog Hot Topic Detection Based On Topic Model Using Term Correlation Matrix", Ieee, 2014,P1-5.
23. Siqi Huang, Yitao Yang, Huakang Li And Guozi Sun, "Topic Detection From Microblog Based On Text Clustering And Topic Model Analysis", Ieee, 2014,P1-5.
24. Bin Xu And Guoliang Fan , "Multimodal Topic Modeling Based Geo-Annotation For Social Event Detection In Large Photo Collections", Ieee, 2015,P1-5.
25. Yuzhong Chen,Wanhua Li, Wenzhong Guo And Kun Guo, "Popular Topic Detection In Chinese Micro-Blog Based On The Modified Lda Model", Ieee, 2015,P1-6.
26. Tian Nie, Yi Ding, Chen Zhao, Youchao Lin, Takehito Utsuro And Yasuhide Kawada, "Clustering Search Engine Suggests By Integrating A Topic Model And Word Embeddings", Ieee, 2017,P1-6.
27. Olga Isupova, Danil Kuzin And Lyudmila Mihaylova, "Learning Methods For Dynamic Topic Modeling In Automated Behavior Analysis", Ieee, 2017,P1-14.
28. Jie Zoub, Ling Xua,B, Mengning Yanga,B, Meng Yanb, Dan Yangb And Xiaohong Zhanga, "Duplication Detection For Software Bug Reports Based On Topic Model", Ieee, 2016,P1-6.
29. S. Sendhilkumar, M. Srivani And G.S. Mahalakshmi, "Generation Of Word Clouds Using Document Topic Models", Ieee, 2017,P1-3.

30. Stanislav Sobolevsky, Iva Bojic, Alexander Belyi And Izabela Sitko, Bartosz Hawelka, "Scaling Of City Attractiveness For Foreign Visitors Through Big Data Of Human Economical And Social Media Activity", Ieee, 2012,P1-10.
31. Yi-Ting Cheng, Yu-Feng Lin, Kuo-Hwa Chiang And Vincent S. Tseng, "Mining Disease Sequential Risk Patterns From Nationwide Clinical Databases For Early Assessment Of Chronic Obstructive Pulmonary Disease, Ieee, 2016,P1-4.
32. Alvin Chua A, Loris Servillo A, Ernesto Marcheggiani B, C And Andrew Vande Moere, "Mapping Cilento, Using Geotagged Social Media Data To Characterize Tourist Flows In Southern Italy", Tourism Management, 2016,P1-16.
33. Astrid Dickinger And Josef A. Mazanec, "Significant Word Items In Hotel Guest Reviews, A Feature Extraction Approach", Ieee. 40 (3), 2015,P1-12.
34. Ana Paula G. Ferreira, Thiago H. Silva And Antonio A. F. Loureiro, "Beyond Sights, Large Scale Study Of Tourists' Behavior Using Foursquare Data", Ieee, 2015,P1-8.
35. Juan Carlos García-Palomares, Javier Gutierrez And Carmen Mínguez, "Identification Of Tourist Hot Spots Based On Social Networks, A Comparative Analysis Of European Metropolises Using Photo-Sharing Services And Gis", Applied Geography, 2015,P1-10.
36. Yue Huang, Yang Chen, Qian Zhou, Jin Zhao And Xin Wang, "Where Are We Visiting, Measurement And Analysis Of Venues In Dianping", Ieee, 2016,P1-6.
37. Thomas Aichner And Frank Jacob, "Measuring The Degree Of Corporate Social Media Use", International Journal Of Market Research. 57 (2), 2014 P1-19.
38. Huy Quan Vu A, Gang Li A, Rob Law B And Ben Haobin Ye, " Exploring The Travel Behaviors Of Inbound Tourists To Hong Kong Using Geotagged Photos", Tourism Management, 2015,P1-11.
39. Yelp.Com , Retrieved From Https://Www.Yelp.Com/Sf, (2017).
40. Snehasish Banerjee And Alton Y. K. Chua, "Using Supervised Learning To Classify Authentic And Fake Online Reviews", Acm, 2015,P1-8.
41. Somayeh Shojaee, Azreen Azman, Masrah Murad, Nurfadhlina Sharef And Nasir Sulaiman, "A Framework For Fake Review Annotation," Ieee, 2015,P1-6.
42. M. Geetha, Pratap Singha And Sumedha Sinha, "Relationship Between Customer Sentiment And Online Customer Ratings For Hotels - An Empirical Analysis", Tourism Management, 2017,P1-12 .
43. Cyril Labbé, Dominique Labbé And François Portet, "Detection Of Computer Generated Papers In Scientific Literature", Acm, 2015, P1-20.
44. Siti Nuradilah Azman, Iskandar Ishak, Nurfadhlina Mohd Sharefand Fatimah Sidi , "Towards An Enhanced Aspect-Based Contradiction Detection Approach For Online Review Content", Journal Of Physics, 2017,P1-11.
45. Tilly, Roman And Cologne, "An Approach To Derive User Preferences From Multiple-Choice Questions In Online Reviews", Twenty-Third European Conference On Information Systems, 2015,P1-11.
46. Shaswat Rungta, "Detection Of Opinion Spam In Online Reviews", Ieee, 2015,P1-35.
47. Enrico Palumbo And Giuseppe Rizzo, "Predicting Your Next Stop-Over From Location-Based Social Network Data With Recurrent Neural Networks", Acm, 2017, P1-48.
48. Snehasish Banerjee And Alton Y. K. Chua, "Authentic Versus Fictitious Online Reviews: A Textual Analysis Across Luxury, Budget And Mid-Range Hotels", Information Science, 2015,P1-14.
49. Mugdha More And Bharat Tidke, "A F Ramework For S Ummarization Of O Nline O Pinion Using W Eighting S Cheme", Advanced Computational Intelligence: An International Journal. 2 (3), 2015,P1-9.
50. Akshay Kangale, S. Krishna Kumar, Mohd Arshad Naeem, Mark Williams And M. K. Tiwari, "Mining Consumer Reviews To Generate Ratings Of Different Product Attributes While Producing Feature-Based Review-Summary", Systems Science, 2016,P1-17.
51. Lawrence Hoc Nang Fong, Soey Sut Ieng Lei And Rob Law, "Asymmetry Of Hotel Ratings On Tripadvisor: Evidence From Single- Versus Dual-Valence Reviews", Ieee, 2016,P1-38.
52. H.Almagrabi And A. Malibari, "A Survey Of Quality Prediction Of Product Reviews. International Journal Of Advanced Computer Science And Applications", 6 (11), 2015,P1-10.
53. Naresh Kumar And Aditi Sharma, "Sentimental Analysis For Political Activities From Social Media Data Analytics", Icetetsm, 2017,P1-10.

54. *Bagnera And Suzanne, "An Examination Of Online Ratings On Hotel Performance Indicators: An Analysis Of The Boston Hotel Market", Ieee, 2017,P1-252.*
55. *Seongsoon Kim, Seongwoon Lee, Donghyeon Park And Jaewoo Kang, "Onstructing And Evaluating A Novel Crowdsourcing-Based Paraphrased Opinion Spam Dataset", 2017,Acm, P1-10.*
56. *Wing Lok Yeung, "A Review Of Data Mining Techniques For Research In Online Shopping Behaviour Through Frequent Navigation Paths", Ieee, 2016,P1-21.*
57. *Anna-Theres Pieper, "Detecting Review Spam On Amazon With Reviewalarm", Acm, 2016, P1-16.*
58. *Khalid Mahmood, "Correlation Between Perception-Based Journal Rankings And The Journal Impact Factor (Jif): A Systematic Review And Meta-Analysis", Serials Review, 2017,P1-12.*
59. *Timeline Of Social Media , Retrieved From Http://Www.Booksaresocial.Com/Timeline-Social-Media-2017/. [Accessed: 10 March 2018].*