Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 9 Issue 01, January-2020

Fraud Detection on Social Media using Data Analytics

Ms. Archna Goyal[1,]
Assistant Professor,
Department of Computer Science Engineering,
Aryabhatta College of Engineering and Research,
Ajmer, India

Ms. Surbhi Singh[2,]
Assistant Professor,
Department of Computer Science Engineering,
Aryabhatta College of Engineering and Research,
Ajmer, India

Er. Saurabh Sharma[3]
HOD,
Department of Computer Science Engineering,
Aryabhatta College of Engineering and Research,
Ajmer, India

***Abstract*: Data Analytics is one of the newest and emerging technologies. An application of artificial intelligence (AI), data analytics provides the ability to learn and improve systems without being programmed but from experience automatically. It centers around the advancement of computer programs that can get to information and use it learn for themselves. Machine learning is used in number of areas. One of major area is detecting fake news on social media. Internet based life for news utilization is a twofold edged sword. From one perspective, its ease, simple access, and fast spread of data lead individuals to search out and expend news from online networking. Then again, it empowers the wide spread of fake news, i.e., low quality news with purposefully bogus data. The broad spread of phony news has the potential for very negative effects on people and society. In this way, counterfeit news identification via web-based networking media has as of late become a developing examination that is drawing in gigantic consideration. Counterfeit news recognition via web-based networking media presents remarkable attributes and difficulties that make existing discovery calculations from conventional news media in active or not relevant. To start with, counterfeit news is purposefully composed to deceive users to accept bogus data, which makes it troublesome and nontrivial to identify dependent on news content. We directed this study to additionally encourage explore on the issue. In this overview, we present counterfeit news portrayals on brain research and social speculations, existing calculations from an information mining point of view, assessment measurements and delegate datasets.**

Keywords: Fraud News, Social Media, Neural Networks, Naive Bayes, Robust Fraud Detection, Entity Link Analysis, Social Network Analysis, Graph Database.

I. INTRODUCTION

Over the most recent twenty years Social media has developed quickly. Huge number of clients pulled in for multi distinctive which had been sorts of internet based life made numerous exercises, social media experience the ill effects of spreading a ton of fake records and fake news made. In addition, the fake records are utilizing their profiles for various objectives, for example, spreading bits of gossip influencing a decided financial matters or even the general public as a bigger section. Recognizing misrepresentation news is an expanding nonstop testing task [1].

Twitter, Facebook, Instagram and others are significant sort of internet organizing perhaps contains enormous information that opens new degree for content examination. There are Challenges in conveying innovation face extortion location, actually, 74% of respondents state the greatest test is either" Lack of IT assets" or" convincing cost/advantage investigation". Notwithstanding those difficulties, it appears that innovation is being embraced at an expanding rate. The greater part of the back up plans overviewed state they have started utilizing against misrepresentation innovation arrangements inside the most recent five years, numerous inside the most recent two years.

It has been proved that the force compelling of internet based life on the consistency of U.S. presidential decisions by means of Twitter, Individuals give more consideration on postings identified with these occasions what's more, will in general effectively accept the tweets [3]. Shockingly, there are malevolent clients who know the propensity, and post and distribute counterfeit tweets, for example, phony and spam data. For model, when tropical storm Sandy occurred, terrible clients posted pertinent messages with counterfeit pictures [4].

Online networking sites (or stages) experience the ill effects of the extending number of phony records that have been made, counterfeit accounts imply that the records don't match to genuine people. Fakes can show counterfeit news and spam. Web based life administrators right now use unique and decided assets to identify, physically affirm, and close phony records [5].

News were re-tweeted by numerous clients who believed re tweeting the news would help the unfortunate casualties influenced by the Hurricane Sandy. Furthermore, individuals give cautious thought on postings related to these crises and tend to easily trust the substance of the postings. [8]

Tragically, there are phony customers who know the conclusion, and post and induce deception, for instance, phony and spam information. For example, when storm Sandy occurred, counterfeit customers posted noteworthy messages with counterfeit pictures [4]. These messages were re tweeted by numerous customers who thought re-tweeting the messages would offer the losses some help with influencing by the Hurricane Sandy [5]. As systems PCs are significant of science and economy huge measure of machine ready to be perused gotten accessible. There are evaluating about 85% of business data lives as content [6]. Lamentably, the standard rationale based programming worldview has extraordinary troubles in catching questionable relations in content archives.

Furthermore, individuals give cautious thought on postings related to these crises and will in general easily trust the substance of the postings. Tragically, there are phony customers who know the estimation, and post and cause misrepresentation, for model, phony and spam information. For example, when storm Sandy occurred, counterfeit customers posted huge messages with counterfeit pictures [8]. These messages were re-tweeted by numerous customers who thought re-tweeting the messages would offer the setbacks some help with influencing by the Hurricane Sandy. Specialists intrigue recognize identify counterfeit news by means of web based life for example, [7], [8] examinations a few cases study as 2013 Moore Tornado and Hurricane Sandy which it was spread by means of microblogs as twitter their methodologies had been founded on recognize validity of picture with some element of tweet where Gupta et al, manages tweet through two measure classifications the principal classification related with recognize counterfeit picture second some component of tweet (16 feature)[8] characterizes counterfeit twitter as it isn't be content with some measure as occasion (place, time, data, picture interface Wrong area identified with the occasion).

The quantities of profile highlights have been a decreased by perceive ten traits for discovery, Be that as it may, as remind in this exploration, the outcome was not confident for perceiving counterfeit records with progressively hopeful view that it is ready to distinguish counterfeit tweets with higher precision by the backing of diagram methods. Gupta et al in [10] apply classifier approach (Nave Bayes, Decision Tree) on a similar way had been applied NB tree classifier to Identify spam and phony messages through twitter with picture and barely any component of tweet content which disregard nearly the substance of. our investigation takes a shot at raise up the precision classifier by contrast our work and their informational index by utilizing closeness approaches strategies and apply extraordinary grouping techniques for confirm our precision where result Nave Bayes 91.52% Decision Tree 96.65% likewise classifier Exactness F-measure Precision Recall NBTree 96.43%.

Then again, likewise Castillo et al in [15] centre around programmed techniques for surveying the trustiness of a given arrangement of tweets and posts. In particular, Castillo et al, investigate smaller scale blog postings related to "drifting" subjects, and arrange them as trust or phony, in light of highlights separated from them. They use highlights from clients' posting and reposting ("retweeting") conduct, from the content of the posts, and from references to outer source.

II. FRAUD DETECTION METHOD

Conventionally fraud detection techniques, for example, a deviation from ordinary or anticipated examples, focus on discrete information instead of the associations between them. Albeit discrete strategies are helpful for discovering fraudsters acting alone, they miss the mark in their capacity to identify sorted out wrongdoing rings. Further, discrete strategies are inclined to bogus positives, which make undesired symptoms in consumer loyalty and lost income opportunity. Gartner proposes a layered model for fraud prevention, which can be seen below:

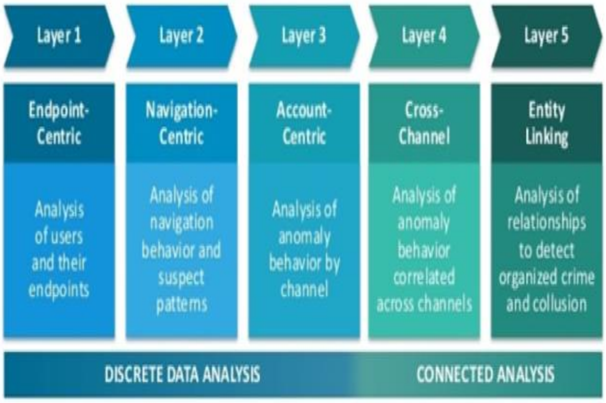
Layer 1
Layer 2
Layer 3
Layer 4
Layer 5
Endpoint-Centric
Navigation-Centric
Account-Centric
Cross-Channel
Entity Linking
Analysis of users and their endpoints
Analysis of navigation behavior and suspect patterns
Analysis of anomaly behavior by channel
Analysis of anomaly behavior correlated across channels
Analysis of relationships to detect organized crime and collusion
DISCRETE DATA ANALYSIS
CONNECTED ANALYSIS

Fig.1: Gartner's Layered Fraud Prevention Approach

It begins with customary discrete strategies (at the left), and advances to progressively expound "enormous picture" sorts of examination. The right most layer, "Entity Link Analysis", use associated information so as to identify sorted out misrepresentation. Intrigues of the sort depicted above can be effectively revealed—with a high likelihood of exactness—utilizing a diagram database to complete element interface investigation at key focuses in the client lifecycle.

A. Classical Approach for Fraud Detection

The traditional way to deal with fake data depends on making of express principles (IF-THEN-ELSEIF-…) in light of the suggestion of specialists. These rules are created and altered through their aggregate field encounters. All things considered, over time, because of the dynamic and complex nature of the fakes, the guidelines become complex and hard to keep up and execute (except if they are normally refreshed).[16] This is likewise a very work escalated approach requiring human intercession at each phase of assessment, recognizable proof, and observing.[16],[17]

The accessibility of information from different sources, furthermore, the capacity of present frameworks to process also, investigate this information have given new open doors for recognizing extortion. As is evident from Figure 2 given below. Fraud Analytics System, the utilization of various information sources to distinguish designs is one of the foundations of an information mining way to deal with misrepresentation recognition. Extortion investigation likewise gives a potential to computerize various phases of the extortion location, observing, and mediation phases of a commonplace cycle.

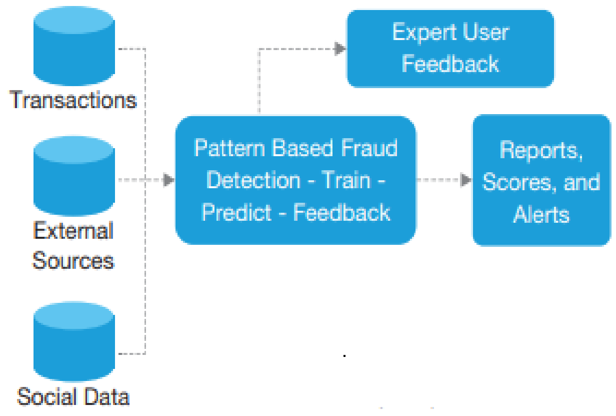
Transactions
External Sources
Social Data
Expert User Feedback
Pattern Based Fraud Detection - Train - Predict - Feedback
Reports, Scores, and Alerts

Fig.2: Fraud Analytics System

HyperGraf™ joins information from different sources, including financial assessments, endeavour value-based information, and internet based life to distinguish furthermore, examine extortion. One of the key strategies utilized in HyperGraf™ is arrange investigation for extortion recognition and the accompanying segment features a portion of its key angles.

B. Entity Link Analysis Using Graph Database

Social databases require datasets to be displayed as a lot of tables and segments. To reveal rings in such a situation requires completing a progression of complex joins and self-joins. Such questions won't just be exceptionally unpredictable to manufacture yet additionally costly to run and will present critical specialized difficulties on scaling. The full size of this issue turns out to be clear as one considers the combinatorial blast that happens as the ring develops alongside the all out dataset.

Diagrams are intended to express connections between information. Diagram databases can reveal designs that are generally hard to recognize utilizing customary portrayals, for example, tables. Since they are intended to inquiry mind boggling associated systems, diagram databases can be utilized to recognize extortion rings in a genuinely clear manner.

C. Robust Fraud Detection by Social Network Analysis (SNA)

At whatever point we consider interpersonal organization examination (SNA), the primary thing that strikes our brain is online life. In any case, SNA is past just Facebook, Twitter, LinkedIn or Google Plus. Informal organization is a system of elements all associated with a specific goal in mind. The elements can be charge cards, organizations, shippers, fraudsters, or others. It can incorporate value-based information, for example, online exchanges and banking information, internet based life information, call conduct information, IP address data, and geospatial information and so on.

This information is regularly put away in unstructured configurations in conditions like internet based life, telecom vaults, instalment doors or bank servers. Fortunately techniques exist to test such huge systems of connections and set up suspicious examples of conduct through diagram database innovation that has been explicitly created to work with enormous datasets that have associations and connections. Putting away and recovering interconnected data in a local 'organize diagram' arrangement can convey intelligent system perceptions to find concealed structures, find bunches and examples, recognize interfaces in exchange chains, and apply particular calculations to distinguish suspicious examples.

NOSQL diagram databases store and recover information in a local system position. Neo4J is a market driving chart database which can be quickly executed and is profoundly versatile. Progressed examination techniques, for example, AI are as of now applied to distinguish false exchanges. Alongside such logical strategies, SNA with chart databases can altogether diminish the bogus positive proportion in extortion recognition.

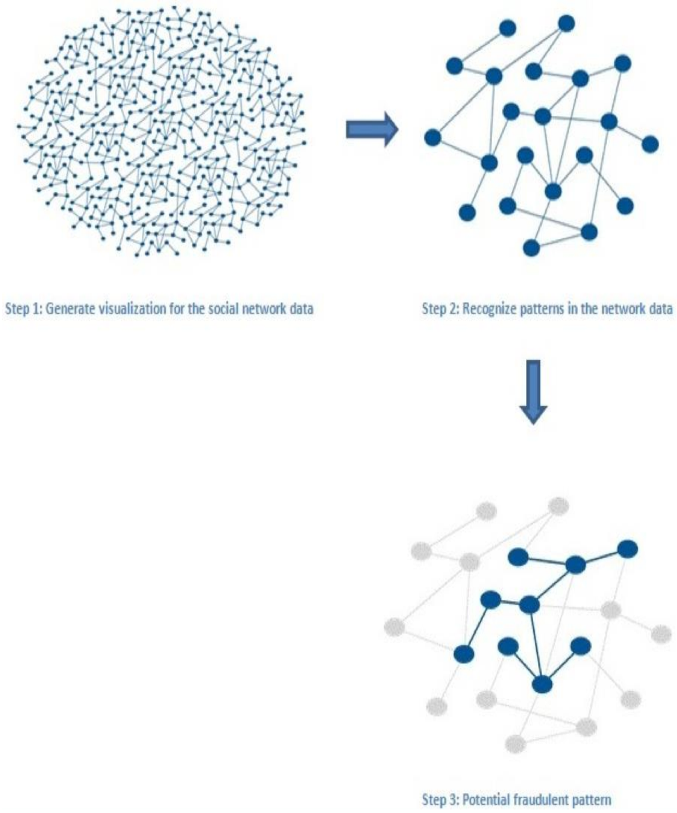
Step 1: Generate visualization for the social network data
Step 2: Recognize patterns in the network data
Step 3: Potential fraudulent pattern

Fig.3: Steps for Fraud Detection by Social Network Analysis

III. PROPOSED APPROACH

The general methodology configuration will be introduced in this proposed approach, which incorporates the methodology design also, brief depiction of the usefulness of each capacity in our approach.

A. The Proposed Approach

The proposed approach is to distinguish Fraud news for Online Informal organization, it comprises of two stages: recognize counterfeit records what's more, recognize counterfeit substance news as appeared in figure 4.

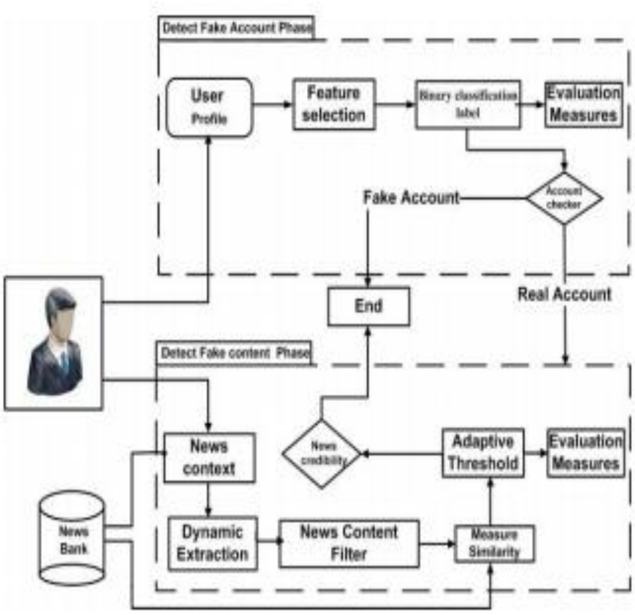

Fig.4: Stages in Proposed Approach

B. Recognize Fake Accounts Phase

This stage plans to viably identify the phony records on the informal community with the conceivable least arrangement of properties. The proposed technique comprises of three principle steps, the initial step to decide the primary factors that impact a right identification of counterfeit records; second step is to apply arrangement calculations that utilization the decided factors in stage one via web-based networking media represents finding counterfeit records; and the third step points to gauge the presentation of a classifier. This may give bogus outcomes in the recognition task. This stage plans to distinguish the phony records on the social media by diminished quantities of profile highlights [11]. The proposed technique comprises of three primary advances, the main stage to decide the fundamental factors that affect of discovery of fake accounts; organize two is to run grouping calculations that utilization the decided highlights in stage one via web-based networking media represents finding counterfeit records; and the third step means to quantify the exhibition of a classifier. This stage means to propose the least arrangement of properties that can recognize the phony clients with the most noteworthy precision. In spite of the fact that the past research in [5], displayed many profile highlights; be that as it may, by playing out an by and large investigation of these qualities, it is uncovered that the greater part of these characteristics are not utilized by the vast majority of the clients and have been left in default mode which may give bogus outcomes in the discovery task.

C. Client Profile

Internet based life systems are currently a well known route for clients to convey what needs be, and share multi data. Clients frequently post a profile, comprising of highlights like age, occupation, and number of companions, geographic area, interests, and schools visited. Such profile data is utilized on the destinations as a reason for gathering clients, for sharing substance, and for recommending clients who may profit by association. Be that as it may, by and by, not all clients give these characteristics [1].

D. Feature Selection

This progression intends to figure a base weighted list of capabilities that impacts the location of the phony records via web-based networking media by an addition proportion measure, which gain proportion process the weight of highlight which result the component choice. Highlight choice preparing overlooks any property estimation under effective condition agreeing worldwide limit. Highlight determination strategies rely upon complete Search for the ideal include as per the assessment work utilized among 2 N subset information, different techniques heuristic or arbitrary hunt strategies endeavour lessen highlights to decrease intricacy execution [12].

E. Age Iteration

Age cycle present all cleaning unique list of capabilities from invalid worth (physically erase invalid columns). Settling the invalid esteem with any arrangements is unsafe for online internet based life information, since NULLs are terrible and hard. For instance, if client profile didn't contain the quantity of companions, taking care of this issue will be creating all the more cheating.

F. Weighted Attributes

Web based life information highlights produce 2N from N includes, a channel capacities are utilized for free criteria highlight subset choice. It doesn't contain any learning calculation assessed utilizing certain assessment criteria. Autonomous online networking information lead us to apply Information Measures as opposed to remove measures and reliance Measures. Since separation measures well known as dissimilarity to gauge distinction between two highlights additionally Dependence Measures measure connection between highlights [12].

G. Information Measures

Proportions of data intend to diminish time and work finished with the making of tremendous internet based life information, it gets helpful data gain from N highlights, (for example Addition proportion measure). Increase proportion: measure plans to least list of capabilities of records via web-based networking media, this measure a proportion of information increment to the natural information. It is used to diminish a rate towards multi-regarded attributes by considering the number and size of branches while picking a property [Karegowda, et al., 2010].

Data Gain proportion = entropy (parent) – [entropy (children)] (1)

Where Entropy originates from data hypothesis. The higher the entropy the more the data content.

$$\text{Entropy}=\sum -P_i \text{Log}_2 P_i \qquad (2)$$

Where P_i is the likelihood of class.

The notion of Gain presented before help traits that have countless qualities. Which increase apportion process weight of highlight which result the element choice. The proposed model had gathered all proposed highlights in the dataset and applied the Gain Ratio measure on the preparation dataset to create weighting for all properties dependent on the idea that the characteristics' weighting decides the adequacy of the property in the order task. The Gain proportion result is analyzed in the Stopping criteria.

- Stopping criteria (Global Threshold)

Stop criteria expects to make edge to express the helpful include information Global Threshold condition (3)

$$T = \frac{(\text{maxvalue}+\text{minvalue})}{2}$$

Where T= Global Threshold

H. Binary Classification mark

This progression expects to foresee the phony records by apply a arrangement calculation to the weight determined traits. There are two sorts of arrangement Binary Classification. Parallel Classification plans to arrange two classes agreeing Administered Learning, Some models incorporate fake identification (eg. Credit Card), Medical Diagnosis, Spam Identification. Presently there are different calculations that are utilized for learning parallel classifiers, which incorporate, Decision Trees, Neural Networks Bayesian Classification, and Support Vector Machines. Along these lines, distinguish counterfeit record apply Binary Classification, theory applies an order calculation for identify counterfeit records that utilizing the determined weighting for the characteristics, the five well known order calculations have been applied again on the dataset [Proctor, 2006].This advance applied the well known order calculations utilizing the weighted characteristics that are resolved in the initial step. These calculations are Random Forest, Decision Tree, Naive bayes, Neural Network, and Support Vector Machine. These calculations are characterized as the best calculations Double Classification in International Conference on Data Mining recognized calculations.

I. DISTINGUISH FAKE CONTENT NEWS

This progression intends to introduce approach ventures for a practical extortion news distinguishing framework, recognize counterfeit substance comprises of five

Steps: elements web extraction, news content channel, likeness measures, classifier calculations, measure the presentation of a classifier. Online networking is a strategy for web clients to arrange, store, oversee and scan for labels, bookmarks (likewise as known as social bookmarking) of assets on the web. Where client created catchphrases and labels have been proposed method for improving portrayals of online data assets, and improving their entrance through more extensive ordering. "Social labelling" alludes to the act of freely marking or characterization assets in a common, online condition. Dissimilar to record sharing, the assets themselves aren't shared, only the labels that depict them or bookmark that reference them. The ascent of social labelling administrations displays a potential extraordinary arrangement of information for mining valuable data on the web. The clients of labelling administrations have made an enormous volume of labelling information which has pulled in ongoing consideration from the examination network [13]. For instance follow the spread a "hashtag" over the system, follow the spread of a specific url, retweets posts .social labelling fabricates information base about the occasion .in the following stage approach will check the presence of these news.

J. News Content Filter

This stage means to encourage the investigation procedure by getting ready information base about the occasion news substance with it keep a structure of news content. The channel preparing of substance news manage blog structure (tweets, posts), news substance comprise of four segments (content, date, picture, url). At the initial step checker intends to recognize presence of hashtag occasion, checker utilizes the well known news web index on the planet web wide. Proposals motor interests gathering news from multidiverse trust assets. The Most Famous Popular site in USA American Cable And Satellite TV station Contain Huge Search Engine For News Via World (http://www.C-Span.Org). Tokenization is the strategy of breaking a flood of substance into words, states of course other significant parts called tokens channel message by utilizing tokenization. For instance, Unique character, Punctuation, whitespace was not being remembered for the subsequent of tokens, the purpose of the tokenization is the examination of the words in a sentence.

Stemming: Snowball stemming strategies is the procedure of decreasing curved (or here and there determined) words to their promise stem, base or root structure commonly a composed word structure. For model, ("stems", "stemmer", "stemming", "stemmed" as in view of "stem").

Measure Similarity: The inconsistency can be particular number and time positions. Next stage after substance channel is measure similitude by TF and Euclidean Distance measure. Term Frequency quantifies how regularly a term happens in internet based life setting (posts, tweets).

Every news is short terms; it is conceivable that that a term would appear to be fundamentally a larger number of times in long blogger than shorter ones. In this way, the term repeat is every now and again as a technique for institutionalization:

TF (t) = (Number of times term t shows up in a blogger) / (Complete number of terms in the blogger) (4)

IV. CONCLUSION

This paper has talked about a proposed approach for illuminating the fake news issues by distinguishing the validity of news in two stages is distinguish counterfeit record ,recognize fake substance of news ,the first stage recognize the phony clients which overlook the news that gives by fake clients however on the off chance that client isn't phony, at that point go to the subsequent stage identify the validity of the news content by utilizing the likeness measures and AI calculations that improve the validity than different calculations This postulation has talked about a proposed approach for tackling the phony news issues by identifying the believability of news in two stages ,In the first stage, counterfeit client was distinguished. by using the Naive Bayes calculation with a decreased arrangement of properties the present methodology rate barely contrasts.. This work commitment can be abridged as follows: -

1. The base arrangement of properties for distinguishing the phony records on Twitter has been resolved and tried.

2. Upgrade the precision of characterization for distinguishes fake news by utilizing likeness measure.

3. Utilizing the content news just to arrange the news is phony or genuine without need different properties, for example, (picture, url, date, and so on). The same methodology will be applied to multi dialects and other internet based life stages.

REFERENCES

[1] Fang Jin, Edward Dougherty, Parang Saraf, Yang Cao, Naren Ramakrishnan,”Epidemiological Modeling of News and Rumors on Twitter”. The 7th SNA-KDD Workshop 13 (SNA-KDD13), August 11, 2013.

[2] Aditi Gupta, Hemank Lamba, Ponnurangam Kumaraguru, Anupam Joshi, ”Faking Sandy: characterizing and identifying fake images on Twitter during Hurricane Sandy”, In Proceedings of the 22nd international conference on World Wide Web companion, 729-736,2013.

[3] Karegowda, A. G., Manjunath, A. S., & Jayaram, M. A. (2010). Comparative study of attribute selection using gain ratio and correlation based feature selection. International Journal of Information Technology and Knowledge Management, 2(2), 271-277.

[4] Hu, X., & Liu, H. (2012). Text analytics in social media. In Mining text data (pp. 385-414). Springer, Boston [14] Niwattanakul, S., Singthongchai, J., Naenudorn, E., & Wanapu, S. (2013, March). Using of Jaccard coefficient for keywords similarity. In Proceedings of the International MultiConference of Engineers and Computer Scientists (Vol. 1, No. 6).

[5] Cresci, S., Di Pietro, R., Petrocchi, M., Spognardi, A., Tesconi, M.,”A Fake Follower Story: improving fake accounts detection on Twitter”, IIT-CNR, Tech. Rep. TR-03, 2014.

[6] Vahed Qazvinian Emily Rosengren Dragomir R. Radev Qiaozhu Mei, ”Rumor has it: Identifying Misinformation in Microblogs”, Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, p.p. 15891599, Edinburgh, Scotland, UK, July 2731, Association for Computational Linguistics,2011

[7] El azab, A., Mahmood A. Mahmood, El-Aziz, A.,”Effectiveness of web usage mining techniques in business application”, web usage mining techniques and application across industries, p.p.324-350,igi global, 2017.

[8] M. Dash, H. Liu,”Feature Selection for Classification”, Intelligent Data Analysis, Vol 1, p.p. 131156, 1997.

[9] Kazem Jahanbakhsh, Yumi Moon,”The predictive power of social media: On the predictability of US Presidential Elections using Twitter”, Social and Information Networks, arXiv: 1407. 0622, 2014.

[10] Adamic L., ZhangJ., Bakshy E., and Ackerman M., ZhangJ .BakshyE., ”Knowledge sharing and yahoo answers: everyone knows something”. Processed in 17th international conference on World Wide Web, ACM, pp 665-674, 2012.

[11] Baesens, Bart, Véronique Van Vlasselaer, and Wouter Verbeke. 2015. “Fraud Analytics Using Descriptive, Predictive & Social Network Techniques.” https://lirias.kuleuven.be/handle/123456789/500346

[12] Adamic L., ZhangJ., Bakshy E., and Ackerman M., ZhangJ .BakshyE., ”Knowledge sharing and yahoo answers: everyone knows something”. Processed in 17th international conference on World Wide Web, ACM, pp 665-674, 2012.

[13] Carlos Castillo, Marcelo Mendoza, Barbara Poblete, ”Information Credibility on Twitter”, the 20th international conference on World wide web ACM, 675-684,2011.

[14] Rajdev, Meet,”Fake and Spam Messages: Detecting Misinformation during Natural Disasters on Social Media”. All Graduate Theses and Dissertations. Paper 4462, 2015.

[15] Supraja Gurajala, Joshua S White, Brian Hudson, Brian R Voter, Jeanna N Matthews, ”Profile characteristics of fake Twitter accounts” in SM Society ’15, July 27- 29, Toronto, ON, Canada, 2015.

[16] Rajdev, Meet,”Fake and Spam Messages: Detecting Misinformation during Natural Disasters on Social Media”. All Graduate Theses and Dissertations. Paper 4462, 2015.

[17] Supraja Gurajala, Joshua S White, Brian Hudson, Brian R Voter, Jeanna N Matthews, ”Profile characteristics of fake Twitter accounts” in SM Society ’15, July 27- 29, Toronto, ON, Canada, 2015.

[18] Fabricio Benevenuto, Gabriel Magno, Tiago Rodrigues, and Virgilio Almeida,”Detecting Spammers on Twitter”, CEAS 2010 - Seventh annual Collaboration, Electronic messaging, AntiAbuse and Spam Conference July 13-14, 2010, Redmond, Washington, US.

[19] A. Talukdar, “Mitigating and Detecting Financial Fraud with Social Network Analysis and Graph Database”, resources.zaloni.com.

[20] http://www.cio.com/article/2422376/infrastructure/fighting-fraud-with-social-network-analysis.html
http://www.propertycasualty360.com/2016/07/25/8-ways-social-networks-help-identify-fraud
https://www.globalbankingandfinance.com/insurance-fraud-detection-are-we-ignoring-social-media-at-our-peril/