

Rumor Detection System for Twitter (A Micro-Blogging Site)



Sakshi Yadav, Anuradha Purohit

Abstract— Micro-blog provides a platform for the users to transfer their thoughts and information in limited words more expressively. Its concise and easy to access nature makes it popular among every age group. Inspite, of all its pros and popularity, some people use it to achieve their bad motives i.e. to misguide people and create violence. To overcome this problem a system is required that will help to detect fake tweets in a limited amount of time. In this paper, a feature based approach for rumor detection has been proposed. The proposed approach utilizes 9 features which shows author as well as readers reaction to identify rumor tweets which may differ for different users in different situations. For experimentation synthetic data and from Pheme has been utilize. A comparative study of the approach for the datasets has been done on the basis of evaluation parameters Recall, Precision and fmeasure. Satisfactory results have been obtained for Pheme data with less number of features as compare to synthetic dataset.

Keywords: Micro-blogs, Social Media, Rumor, Machine learning algorithms.

I. INTRODUCTION

Nowadays, micro-blog systems are more popular. The reason for the popularity is fast transfer rate of information. By using a micro-blog system a user can share information and views by publishing a post, re-post and re-post adding own comments [1]. Some popular micro-blog systems are Twitter, Tumblr and Sina weibo. One of the most popular micro-blog system in India is Twitter, from businessman to politicians and common man to popular personalities all are active on Twitter. As popularity increases the probability of the number of fake tweets also increases. Credibility matters a lot when it is all about information. As information can create violence on the other side it will help to solve many serious issues. The main motto of these fake tweets is to create violence and misguide people. Rumor is a piece of information whose sources are untrustworthy.

These are likely to be generated under crisis and extremity, causing public panic, disrupting the social order, decrease trust on government and directly These fake tweets are known as rumors and affects security of the nation.

Manuscript published on November 30, 2019. * Correspondence Author

Anuradha Purohit, Associate Professor, Computer Engineering Department, S.G.S.I.T.S. Indore, <u>anuradhapurohit78@gmail.com</u>

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an <u>open access</u> article under the CC-BY-NC-ND license <u>http://creativecommons.org/licenses/by-nc-nd/4.0/</u>

Retrieval Number: D4308118419/2019©BEIESP DOI:10.35940/ijrte.D4308.118419 Journal Website: <u>www.ijrte.org</u> For example, in June 2016, after banknote demonetization was officially announced. RBI declared the message of invalid 10 rupee coin; it was spread so quickly on social networking mostly in the area of metro cities like Delhi. The declaration became the reason for not accepting 10-rupee coin by shopkeepers, rickshaw drivers and creating confusion among people. This rumor became a great issue among people. After all this RBI confirmed that, who are not accepting the currency will have to face legal action.

Credibility is a major concern for researchers. That's why researcher focus on the reliability of the information which spread through online platform using features extracted from tweets [2]. Some researchers make use of previously done survey by applying k-nearest neighbor and Naive Bayes classifier which are machine learning algorithm and helps to improve the efficiency of existing approach. Many researchers have shown interest in automatic rumor detection method on the online social platform. These methods can be classified into two categories: classification-based approach and propagation-based approach [3]. In the direction of automatic rumor detection, a classification method has been proposed [4] which treats rumor detection as a binary classification problem and make use of a combination of implicit features and shallow features of the messages. As the popularity of micro-blogs increases the amount of data also increases. So, finding a recent trend topic becomes a tough task. So, their importance may vary with time as well as the situation. Since the feature based identification is more reliable and dependable, the proposed approach is based on the feature- based rumor detection system. In this paper, an approach has been proposed which is based on features where the behavior of the user is treated as hidden clues to find rumor posts. Proposed approach works in three phases: 1) based on collected micro-blogs of Twitter, features of user's behavior have been gathered. 2) Three most popular algorithms have been used to train classifiers for rumor detection. These are SVM (support vector machine), RF (random forest) and MaxEnt (maximum entropy). 3) Trained classifier from phase two used to predict whether a post is a rumor or not. Experiments are conducted on two types of the dataset of Twitter, one of them is synthetic dataset which is based on sentiments and another one is Pheme dataset which consists of rumors and normal post related to 5 breaking news to show the performance of the designed system. Evaluation parameters such as, precision, recall and f-measure are calculated which in return shows that fewer features can improve performance.

Published By: Blue Eyes Intelligence Engineering & Sciences Publication



12287

Sakshi Yadav*, M.E. Scholar, Computer Engineering Department, S.G.S.I.T.S. Indore, <u>sakshiyad06@gmail.com</u>

Also, a comparative study has been done on datasets having a different number of attributes.

This paper is organized into six sections. Section II gives a detailed survey of work done in the field of rumor detection. In Section III, the problem of rumor detection has been explained. Section IV consists of a description of the proposed approach for rumor detection. Experiment results are presented in Section V. Finally, conclusion of the paper is discussed in Section VI.

II. LITERATUREREVIEW

A. Survey of Rumor Detection Approaches

There is various rumor detection approaches used for identification of fake news. The major categories identified are: machine learning based algorithms, feature based methods and application based.

Xiamo Liu et al. [5] proposed an approach which is concentrate on identification of a rumor. It treats rumor as an event that may comprise of one or more conflicting microblogs. This work includes continuous monitoring of rumor event and generates real time updates dynamically based on any additional information received.

Sardar Hamidian et al. [6] proposed an approach which makes comparison in between a single-step (SRDC) 6 way classification with a two-step classification (TRDC). Itworks in two-fold: (i) Proposed an automated TRDC pipeline that employs the results from the rumor detection step and performs the classification task upon data and gives better results than SRDC. (ii) It employed a set of meta linguistic and pragmatic features, which performs the experiments with and with out pre-processing on the textual content.

Majed Al Rubaian et al. [7] has been proposed a multistage reliablity analysis framework which identifies rumors in Twitter. Nave Bayes classifier is used and it is enhanced by considering the relative importance of the used features to improve the classification accuracy. Classifier examined with 1000 unique tweets along with 700 accounts.

Carlos Castillo et al. [8] assess the credibility of a given set of tweets by focusing on automatic methods. Credibility refers to the reliability. Specifically, it works for trending topics and classify them as credible or not credible by making use of features like posting and re-posting extracted from them.

Qiao Zhang et al. [9] proposed a rumor detection method based on the combination of new proposed implicit features and shallow features (like span, URL, multimedia, refer) of the messages. The proposed implicit features obtained by deep mining, it includes popularity orientation, internal and external consistency, sentiment polarity and opinion of comments, opinion re-tweet influence and match degree of messages.

Jaya prakash et al. [10] proposed a rumor identification scheme by applying five new features which are based on user's behaviors and combine the new features with the existing user behavior based features, such as followers comments and re-posting to predict whether a micro-blog post is a rumor or not.

Castillo et al. [11] proposed an approach which extract 68 features from posts of twitter and categorized them in to four types: 1) content-based features, which consider characteristics of the tweet content; 2) user-based features, which consider traits(qualities) of Twitter users; 3) topicbased features, which are computed from message-based and user-based features; 4) propagation-based features, which consider features related to the propagation tree of a post.

Raveen Dayani et al. [12] proposed an algorithm for the purpose of pre-processing on tweet content is to retain key information which is to be passed on to learning algorithm to obtain improved results as far as rumor detection is concerned. For this researcher, investigate in retrospect a dataset on which rumor detection was done in the past in 2009.

Yasin N. Silva et al. [13] focused their research on cyber bullying for which they proposed a model. These model identify the degree of cyber bullying as well as a Facebook app notifies the parents when it happen with their children.

Shubhi Mittal et al. [14] presented a report which consist responses of 100s of users and consumers which helps to show the effect of sentiments on e-marketing as well as social networking sites.

Dr. S. Vijayarani et al. [15] discussed about various text mining processes and their preprocessing techniques.

Jing Ma et al. [16] proposed a method for identification that learns continuous micro-blog events representation. The proposed approach identify the variation in contextual information of relevant posts over time by learning hidden representations. In this model recurrent neural network has been used. It detects rumors more accurately and quickly than existing techniques.

Zhe Zhao et al. [17] used a technique to identify rumors related to trending issues. It works in several steps: search for the inquiry phrases, clusters of similar posts form together, then collects related posts that do not contain these simple phrases, than rank the clusters by their likelihood of really containing a disputed factual claim.

Samantha Finn et al. [18] introduces Twitter Trails which is a semi-automatic web-based tool that will help users to investigate the origin of the fake news and propagation characteristics of a rumor, its reliability on Twitter. It is especially designed for amateur and professional journalists investigating recent and breaking stories. It made use of tweet relevancy algorithm, brustiness algorithm, and negation classification and propagation metrics.

Paul Resnick et al. [19] introduces an useful interactive tool Rumor Lens which is semi-automatic and designed to help journalists to identify new rumors on Twitter.

Through this automated tool, learning and computation makes it feasible to engage journalists and the broader public to run a continuous rumor-

Juno and Engineents IJRTE IJRTE

Retrieval Number: D4308118419/2019©BEIESP DOI:10.35940/ijrte.D4308.118419 Journal Website: <u>www.ijrte.org</u>

12288

Published By: Blue Eyes Intelligence Engineering & Sciences Publication

monitoring service.



Anh Dang et al. [20] introduced a prototype system, called Rumor Flow. The proposed approach is supported by a set of analytical tools which includes rumor spread models, sentiment analysis, along with Wikipedia entity linking and text semantic similarity, to help attain a visual summarization of the many facets of rumor spread, and provide various levels of granularity through an exploratory interface. This work introduces an approach to tackle multi-component problem by adapting state-of-the-art visualization techniques and providing a Visual Analytics (VA) tool for understanding and analyzing how rumors are disseminated and discussed by OSNusers.

Arkaitz Zubiaga et al. [21] discussed about SUPER which was a 3-year project funded by the European Commission that studied the use of Social sensors for security assessments and Proactive Emergencies management, in part dealing with crowd sourced annotation of rumors.

Chengcheng Shao et al. [22] introduce Hoaxy which is a platform for the collection, detection and analysis of online misinformation. It focuses on two aspects: the temporal relation between the spread of misinformation and fact checking, and the differences in how users share them.

Stuart E. Middleton [23] suggested a credibility model which was able to process real-time evidence which are extracted by making use of a combination of natural language processing, image analysis, social network analysis and semantic analysis. Middleton along with Krivcovs suggests approaches to track down the location of social media users.

Aditi Gupta et al. [24] present a semi-supervised ranking model named as TweetCred for scoring tweets according to their credibility. It's a real-time system that assigns a credibility score to tweets in a user's time-line. An extensive set of 45 features is used to determine the credibility score for each of the tweets.

III.BACKGROUND

In this section, background of the system has been discussed. Text plays a major role for communication over micro-blogs which needs to classify according to their content. Any kind of text retrieval system requires text classification as it helps to respond user query and extract data. There are several methods available which has been discussed in research [25]. In case of web documents genetic algorithms produces satisfactory results efficiently [26].In case of social platform, Micro-blog provides ease to the user so that they can share information and express their views on every topic. On this platform there is a network of users having certain relationships. As there are two ways by which users communicate on Twitter either by publishing own post or by publishing someone else post (adding or without adding own comments). This relationship can be understood as, suppose there are three users namely A, B and C. If user A follow user B then A can see posts posted by B but for the same B needs to follow back user A. As in the same way C can't see posts by A as they do not follow each other. A and B can comment on each other's post and re-tweet each other's tweet.

The relation of micro-blog and user can be explain by taking user, attributes and features in the scenario. Suppose there are n numbers of users having i number of attributes. These attributes can be date of registration, number of followers, number of followee etc. For the calculation of values of features, let it be j whose value vary from 0 to 8, so it has been mapped as value of feature j for user n by making use of attributes i=1,2,...m where is number of attributes of a user. Final result will be obtained by making use of this relationship.

These calculations and implementation become possible practically by using RStudio framework, which is a freely available environment for statistical computing and GUI. Its environment provides a convenient and easy to use tool for organizing packages, workspace and more.

IV.PROPOSEDAPPROACH

The system architecture of the proposed approach has been discussed in this section. Also, it describes working of the system, calculation of features, brief of levels of architecture and relations among levels. Fig.1 describes the system architecture of the proposed approach.

In Fig.1, The system architecture of the proposed system is divided into 5 levels from bottom to top. In level 1, the basic need of any system that is input data is collected. Here, input tweets in the form of data are used which consist of attributes and there corresponding values like a source of URL, number of followers, number of posts etc. On moving on level 2 make use of this dataset and convert it into preprocessed and normalized data by the using tf-idf (for preprocessing) and min-max (for normalization). For getting values in a specific format. Level 3 deals with the calculation of values of features corresponding to a tweet by making use of attributes given in dataset for a particular tweet. After calculation of values, level 4 deals with classifiers namely, SVM(support vector machine), RF(random forest) and MaxEnt (maximum entropy) for the purpose of training and testing. In level 5, the final output will be shown regarding whether a tweet is a rumor or not.



Retrieval Number: D4308118419/2019©BEIESP DOI:10.35940/ijrte.D4308.118419 Journal Website: www.ijrte.org

Published By:

& Sciences Publication



Fig.1 System Architecture of Proposed Approach

Proposed approach is based on feature based identification. Hence selection of feature is a very important part of a system. As it can improve performance or degrade it. There are 9 features used in the proposed system which are based on author, reader and available features (known as implicit features). These features are based on human behavior and used as hidden clues. As it has been seen that behavior of human and pattern of post may vary for rumor and normal post, as an example rumor post are questioned more than normal post.

Calculation for features, there categories and importance is mentioned below:

1) Features Based on Micro-blog's Author

In the proposed work, features based on micro-blog's author refer to features extracted from behavior of author who published post. It includes average number of followees per day, average number of posts, verified user or not and number of follower probability. The same are derived as,

a) Average number of followees per day:

A followee is someone who is followed by others. If someone follows others than there is an option of follow back. Mostly, it has been seen that popular personalities follow back to few people but they have large number of followers. Rumor-monger wants to spread their fake news among many people so they increase their followee by following large number of people. The value of average number of followees per day calculated by dividing number of followees by user register day, value of user register day can be extracted through data. b) Average number of posts:

Average number of posts means how many posts a user posted. It has been seen that rumormongers who want to share fake news post micro-blog in less number. Also they use different accounts so that they can't be trapped. It can be calculated by dividing number of post by user register day.

c) Verified user:

Verified tag is an important feature for rumor detection. By this tag one can easily identify whether the microblog is posted by authenticate user or not. Example, a post related to a company can be verified by checking who has posted that micro-blog either an authenticate person or an employ.

- d) Calculate number of follower probability: Follower is a person who follows other person. A person can follow you only by your permission. Higher number of follower shows that user is popular among others like celebrities. The more number of followers, more people will receive author's post. So, rumor-mongers mostly post fake information when followers' number reaches to high.
- Features Based on Micro-blog's Reader In the proposed work, features based on micro-blog's reader refer to features extracted from behavior of reader when they read post. It includes number of re- posts/ retweets and number of questioned comments. The same are described as,
- a) Number of reposts/re-tweets:

Micro-blog platform gives a facility of repost and retweet. It shows how many people respond to the microblog post. Repost means posting someone's post either by adding or not adding comments and re-tweet means retweet it in favor of it. Mostly, rumors spread are related to popular topics no matter whether sources are reliable or not. For example, end of the life on earth in 2020. Therefore, value of these features is more in case of rumors than normal post.

b) Number of questioned comments:

Twitter provide a facility of comment to user where he/she can comment on posted micro-blog. Many researchers have been analyzed that false information is much more questioned than truth. This feature check the number of questioned words in comments like, what, how, when, where, whom, why, who etc.

3) Implicit Features

Implicit features refer to the features which were extracted through deep mining of existing features. It includes popular orientation, social influence and opinion re-tweet influence. The same are described as,

a) Popular orientation:



Retrieval Number: D4308118419/2019©BEIESP DOI:10.35940/ijrte.D4308.118419 Journal Website: <u>www.ijrte.org</u>

12290

Published By: Blue Eyes Intelligence Engineering & Sciences Publication



When there is bulk amount of data present it becomes difficult to get statements which are hot topics. In this situation the popular orientation feature matches the similarity among the words and sorts them out. The popular orientation de- fined as, $max(sim_i(w, t1), sim_i(w, t2))$

where, w= keywords set of message,

ti= certain category of popular words.

b) Social influence:

For finding a rumormonger along with internal some related factors are also responsible like, friends, follower and bi-follower. It has been seen that mostly people follow those accounts which are authentic. Even if there are more friends in common these also decrease the possibility of rumor. It can be calculated as,

Log((folnum-bifolnum)/frinum+1) where, folnum

= number of followers,

bifolnum = common in both friends as well as followers, frinum = number of friends.

num = number of finance.

c) Opinion re-tweet influence:

It shows the effect of re-tweet. It means how re- tweet will make influence on opinion of others. For this, it will make use of re-tweet number and status count. Status count can be obtained from data. Value of opinion re-tweet influence can be calculated by dividing re-tweet number by status number.

Five fold cross validation method is used which is available as a library in RStudio framework. In 5 fold cross validation data is divided into 5 folds with the same size, then training can be done on 4 folds of the data and testing can be done on remaining folds.

V. EXPERIMENT

This section includes description of datasets and results obtained by the system.

A. Datasets:

For the evaluation of the results obtained by proposed approach there are two types of datasets taken namely, synthetic dataset and Pheme dataset.

Synthetic dataset has been manually designed. It uses sentiment analysis dataset obtained from git-hub. It is used to check the system performance for 9 features. Data available in synthetic dataset taken from different fields which consist of 16 attributes. Source: https://github.com/vineetdhanwat/twitter-

analysis/blob/master/datasets/Sentiment

Pheme dataset contains a collection of Twitter rumors and non-rumors posted during breaking news. It works for 4 features. These five breaking news are related to Chrlie Hebdo, Ferguson, German wings Crash, Ottawa Shooting and Sydney Siege. It consists of rumors and non-rumors in its sub directories. It was used in paper "Learning reporting dynamics during breaking news for rumor detection in social media".

B. Evaluation Metrics:

To evaluate the results of proposed approach, there are three evaluation parameters used. These are precision, recall and f-measure.

Precision is the fraction of the correctly predicted rumor micro-blogs to all rumor micro-blogs identified. Recall is the proportion of correctly predicted rumor micro- blogs to all the rumor micro-blogs. Whereas Fmeasure can be considered as the harmonic mean of recall and precision.

C. Results:

In this result section, experiment results will be discussed for two datasets: Pheme dataset and synthetic dataset. During experiment it has been seen that, comparison among both the datasets in terms of accuracy- Pheme returns best results with fewer features.

As it has been discussed that there is a difference in behavior of author and reader when they read rumor post or normal post, so a system has been proposed which makes use of these features and identifies whether a post is a rumor or normal. In order to train and test the proposed system, Three classifiers have been used: 1) Support vector machine; 2) Maximum Entropy and 3) Random Forest.



Fig.2 Evaluation parameters of synthetic dataset Fig.2 shows the result of evaluation parameters for synthetic dataset. The result of precision, recall and f-measure for SVM algorithm is 0.56, 1 and 0.66 respectively. For maximum entropy it returns values of 0.57, 1 and 0.62 while for random forest values of precision, recall and fmeasure are 0.61, 1 and 0.75 respectively.



Retrieval Number: D4308118419/2019©BEIESP DOI:10.35940/ijrte.D4308.118419 Journal Website: <u>www.ijrte.org</u>

12291

Published By: Blue Eyes Intelligence Engineering & Sciences Publication





Fig.3 shows the result of evaluation parameters for Pheme dataset. The result of precision, recall and f-measure for SVM algorithm is 0.75, 1 and 0.81 respectively. For maximum entropy it returns values of 0.83, 1 and 0.88 while for random forest values of precision, recall and f- measure are 0.86, 1 and 0.90 respectively. Results shows that, proposed system works better by using less amount of time and return results to the user within seconds.

VI.CONCLUSION

In this paper proposed a feature based identification system for rumor detection has been proposed. The approach treats user's behavior as hidden clues and uses three classifiers for training and testing purpose.

The experimental results shows that, system works efficient for fewer features. Different evaluation parameters lie in the range of 0.55 to 0.80 for synthetic dataset and 0.75 to 0.90 for Pheme dataset. The results obtain are more efficient in terms of time and accuracy for Pheme dataset. The development of the system is a contribution in the field of security on social networking sits.

REFERENCES

- Gang Liang, Wenbo He, Chun Xu, Liangyin Chen, Jinquan Zeng Rumor identification in microblogging system based on users behavior in IEEE Transactions on computational Social Systems, 2015, Pages 99-108.
- Carlos Castillo, Marcelo Mendoza, Barbara Poblete,"Information Credibility on Twitter", Proceedings of the 20th international conference on World wideweb, Pages675-684.
- Xing Zhou, Juan Cao, Zhiwei Jin, FeiXie, Yu Su, Dafeng Chu, Xuehui Cao, Junqiang Zhang, "Real Time news certification on sina weibo "Proceedings of the 24th International Conference on World Wide Web, Pages 983-988,2015.
- Qiao Zhang, Shuiyuan Zhang, Jian Dong, Jinhua Xiong and Xueqi Cheng Automatic detection of rumor on Social Network, Proceeding NLPCC 2015 Proceedings of the 4th CCF Conference on Natural Language Processing and Chinese Computing - Volume 9362, Pages 113-122.
- Xiamo Liu, Arminesh Nourbakhsh, Quanzhi Li, Rui Fang, Sameena Shah, "Real Time rumor debunking on twitter" Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, Pages 1867-1870,2015.
- 6. Sardar Hamidian and Mona Diab Rumor Detection and classification
- 7. for twitter data, SOTICS 2015 :The Fifth International Conference on Social Media Technologies, Communication, andInformatics.

- Majed Al Rubaian, Muhammad Al-Qurishi, Mabrook Al-Rakhami, SkMd Mizanur Rahman, Atif Alamri A Multistage Credibility Analysis Model for Microblogs, 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining.
- Sardar Hamidian and Mona T Diab Rumor identification and belief investigation twitter (www.aclweb.org/anthology/W16-0403),SanDiego, California,June12-17,2016,Pages3-8.
- Zhiwei Jin, Juan Cao, Han Guo, Yongdong Zhang, JieboLuoMultimodal fusion with recurrent neural networks for rumor detection on microblogs, Proceedingsofthe 2017 ACMonMultimedia Conference, Pages 795-816.
- Jayaprakash. S ,Sumathi. P, Suganya. S, "Identification of rumor spreading platform in microblogging systems" Published by AENSI Publication 2017 Special 11(6):Pages167-173.
- C. Castillo, M. Mendoza, and B. Poblete, Information credibility on twitter, in Proc. 20th Int. Conf. World Wide Web, 2011, Pages675684.
- Raveena Dayani, Nikita Chhabra, Taruna Kadian and Rishabh Kaushal Rumor Detection in Twitter: An Analysis in Retrospect, Advanced Networks and Telecommuncations Systems (ANTS), 2015 IEEEInternational Conference on Advanced Networks and Telecommuncations Systems(ANTS).
- Yasin N.Silva, Christopher Rich, Deborah Hallz"BullyBlocker:Towards the Identification of Cyberbullying in Social Networking Sites", in 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM) 978-1-5090-2846-7 2016 IEEE.
- Shubhi Mittal, AshnaGoel,RachnaJain,"Sentiment Analysis of E-Commerce and Social Networking Sites "in Computing for Sustainable Global Development (INDIACom), 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom) March2016.
- Dr.S.Vijayarani, Ms.J.Ilamathi, Ms.Nithya, Preprocessing techniques for text mining-An overview, International Journal of Computer Science and CommunicationNetworks, Vol5(1), Pages 7-16.
- Jing Ma, Wei Gao, Prasenjit Mitra, Sejeong Kwon, Bernard J.Jansen, Kam-Fai Wong, Meeyoung Cha, "Detecting Rumors from Microblogs with Recurrent Neural Networks" in IJCAI'16 Proceedings, Twenty-Fifth International Joint Conference on Artificial Intelligence, Pages 3818-3824.
- Zhe Zhao, Paul Resnick, Qiaozhu Mei, "Enquiring Minds: Early Detection of Rumors in Social Media from Enquiry Posts" in Proceeding WWW'15 Proceedings of the 24th International Conference on World Wide Web, Pages1395-1405.
- Samantha Finn, Panagiotis Takis Metaxas, Eni Mustafaraj, "Investigating Rumor Propagation with TwitterTrails", partially supported by NSF grant CNS-1117693 and by the Wellesley Science Trustees Fund.
- 20. Paul Resnick, Samuel Carton, Souneil Park, Yuncheng Shen, Nicole Zeffer, "Rumor Lens: A System for Analyzing the Impact of Rumors and Corrections in Social Media", partially supported by the National Science Foundation under grant numbers IIS-0968489. This work is also partially supported by Google.
- 21. Anh Dang, Abidal rahman Mohd, Evangelos Milios, Rosane Minghim, "What is in a Rumour: Combined Visual Analysis of Rumour Flow and User Activity", Proceeding CGI '16 Proceedings of the 33rd Computer Graphics International, Pages17-20.
- 22. Arkaitz Zubiaga, Ahmet Aker, Kalina Bontcheva, Maria Liakata and Rob Procter, "Detection and Resolution of Rumours in Social Media: A Survey", Journal ACM Computing Surveys (CSUR)Surveys Homepage archive, Volume 51 Issue 2, April 2018 Issue-in-Progress Article No.32.
- Chengcheng Shao, Giovanni Luca Ciampaglia, Alessandro Flammini, Filippo Menczer, "Hoaxy: A Platform for Tracking Online Misinformation", Proceeding WWW '16 Companion Proceedings of the 25th International Conference Companion on World Wide Web, Pages 745-750.
- Aditi Gupta, Ponnurangam Kumaraguru, Carlos Castillo, Patrick Meier, "Tweet Cred: Real-Time Credibility Assessment of Content on Twitter", International Conference on Social Informatics SocInfo 2014: Social Informatics, pp228-243.

Published By: Blue Eyes Intelligence Engineering & Sciences Publication



Retrieval Number: D4308118419/2019©BEIESP DOI:10.35940/ijrte.D4308.118419 Journal Website: <u>www.ijrte.org</u>

12292



- Anuradha Purohit, Deepika Atre, Payal Jaswani, Priyanshi Asawara: "Text Classification in data mining", International Journal of Scientific and Research Publications 2015: Volume 5, Issue 6, pp.1-7.
- Shalini Bansal, Anuradha Purohit, Kirti Arora, Neela Pandit, Smita Sharma, "Genetic Algorithm for Classification of Web Documents", International Conference on Advanced Computing, Communication and Networks, 2011, pp. 1298-1301.

AUTHORS PROFILE



Sakshi Yadav, has completed her B.E. degree in Computer Science Engineering from Mahakal Institute of Technology, Ujjain in 2016. She has received her M.E. Degree in Computer Engineering from Shri G. S. Institute of Technology And Science, Indore. Her field of research is machine learning.



Anuradha Purohit, received the B.E. degree in Computer Engineering from Vikram University, Ujjain in 2000. She has received her M.E. Degree and PhD Degree in Computer Engineering from Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal. She has 19 years of teaching experience and is currently working as Associate Professor in Department of Computer Engineering at Shri G. S. Institute of Technology and Science, Indore. Her research interests include Soft Computing and Software Engineering.



Retrieval Number: D4308118419/2019©BEIESP DOI:10.35940/ijrte.D4308.118419 Journal Website: <u>www.ijrte.org</u>

12293

Published By: Blue Eyes Intelligence Engineering & Sciences Publication