Impact Factor Value 4.046 e-ISSN: 2456-3463
International Journal of Innovations in Engineering and Science, Vol. 4, No.4, 2019
www.ijies.net

A System to Analyze Credibility of User on Social Media Like Twitter Using WSRI

C.R. Pote[1],Bhagyashree durgekar [2], Pranoti patil [3], Dharmendra dhapade[4], Damini Gahane[5]

[1]Assistant Professor,[2,3,4,5,6,]Students
Priyadarshini College of Engineering, Nagpur, India, 440019

Abstract *–Today, social media provide the means by which billions of people experience news and events happening around the world. However, Data generated on social media has become a rich source for various data mining tasks. Those data analysis tasks that are dependent on the post semantics, such as sentiment analysis, emotion mining, and rumors detection. Information credibility on Social media has been a topic of interest among researchers in the fields of both computer and social sciences. Social media has made it increasingly possible to offer near-real-time transfer of information in a very cost effective manner. This is a platform which delivers timely content in a tailored manner that makes it possible for users to obtain news regarding their topics of interest. Consequently, the development of techniques that can verify information obtained from social media has become a challenging and necessary task. In this paper, we propose a new credibility analysis system for assessing information credibility on social media to prevent from fake or malicious information. The proposed system consists of three integrated components: Credibility analysis, sentimental analysis, WSRI algorithm. The components operate together in an algorithmic form to analyze and assess the credibility of social media, post and users. Using certain pre-processing technique - Replace Negations with Antonyms, Handling Negations, Remove Stop words, Remove Punctuation.*

Keywords-*Data mining, social media, text mining, user, WSRI algorithm, credibility analysis*

I- INTRODUCTION

Credibility can be defined as the quality of being trusted and believed in or the quality of being trusted and believed in," or "the quality of being convincing or believable." The base word of credibility is "credible," which can be convincing," or "capable of persuading people that something will happen or be successful". Latterly, the most correlated synonyms of credibility have been "trustworthiness" and "believability."

For the purpose of our research on Twitter, we define three classes of credibility with respect to two levels of credibility assessment as follows.

Definition 1 post- (tweet-) level credibility, denoted by c is a numerical score for tweet that reflects how believable and reliable the tweet is , and thus how likely it is that the tweet conveys acceptable information , regarding a certain event.

Definition 2 User- (account-) level credibility, denoted by is a numerical score for user that reflects the trustworthiness of a user in an online social network. The lower the trustworthiness of a user is, the more likely it is that the information disseminated by that user is not credible.

Definition 2 User reputation level

User' reputations are based on popularity measures. We describe user as one who is recognized by other user on a similar network. The measures include the Follower-Rank and the Twitter Follower-Follow Ratio (TFF). In addition, we consider replies and re-tweets as measures of a user's popularity.

Among the many challenges related to studying credibility on social networks and the web are the following:

1. The complexity of social networks and the web creates difficulty in identifying resources for use in studying and assessing credibility.

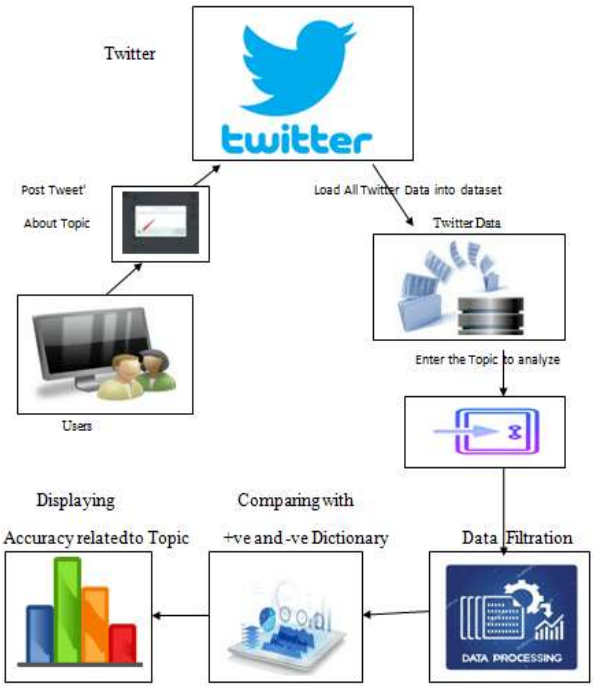
2. OSNs by their very nature evolve dynamically over time and become very large in size, with various structures that make it difficult to obtain the information needed to discern the credibility of users.

3. The credibility of a user is influenced continuously by various factors, such as changes in the social topography, other users' behavior, preferences, and context.

4. Malicious activities can evade existing spam filters through various means. For example, in Twitter, malicious users can purchase followers or use tools to automatically generate fake ac-counts and post tweets with the same meaning but different words.

5. The process of evaluating solutions has also been a problem in terms of resources, given that most researchers are limited in terms of the extent to which they can test their

II- METHOLOGY



A. DATA MINING

Data mining is the process of sorting through large dataset to identify patterns and establish relationship to solve problems through data analysis. Data mining tools allow enterprise to predict feature trends.in data mining, association rule are created by analyzing data for frequent patterns, then using the support and confidence criteria to locate the most important relationship within the data. Support is how frequently the items appear in the database, while confidence is the number of times if/then statement are accurate.

B. TEXT MINING

Text mining is a process to extract interesting and significant patterns to explore knowledge from textual data sources. Text mining is a multi-disciplinary field based on information retrieval, data mining, machine learning, statistics, and computational linguistics. Figure 1 shows the Venn diagram of text mining and its interaction with other fields. Several text mining techniques like summarization, classification, clustering etc., can be applied to extract knowledge. Text mining deals with natural language text which is stored in semi-structured and unstructured format. Text mining techniques are continuously applied in industry, academia, web applications, internet and other fields. Application areas like search engines ,customer relationship management system, filter emails, product suggestion analysis, fraud detection, and social media analytics use text mining for opinion mining, feature extraction,

C. PRE-PROCESSING TECHNIQUE

Replace Negations with Antonyms: It is an approach that has not been used by many researchers and is presented in . We search in each sentence for the word 'not' and then, we check if the next word has an antonym. If yes, we replace both words with the antonym. For example, the phrase 'not good' will be replaced with the word 'bad', using WordNet.

Handling Negations: When text analysis is performed in a word level, it is very challenging to handle negation. One method that is widely used by researchers is the detection of words that imply negation and the addition of the prefix 'NOT ' in every word after them until the first punctuation mark.

Remove Stopwords: Stopwords are function words with high frequency of presence across all sentences. It is considered needless to analyze them, because they do not contain much useful information. The set of these words is not completely predefined and it can be changed by removing or adding more to it, depending on the application. In our implementation, we used the standard stopwords provided by NLTK.

Remove Punctuation: In many works, it is common to remove punctuation signs in pre-processing. However, many times the presence of punctuation marks denotes the existence of some sentiment. For example, an exclamation mark may mean an intense positive or negative sentiment. So if we remove them we might decrease the accuracy of classification

D. WSRI algorithm

WSRI stand for word segmentation of review information. Word segmentation is the problem of dividing a string of written language into its components words. Word splitting is the process of parsing concatenated text (i.e. text that contains no spaces or other word seprators) to infer where word breaks exit.

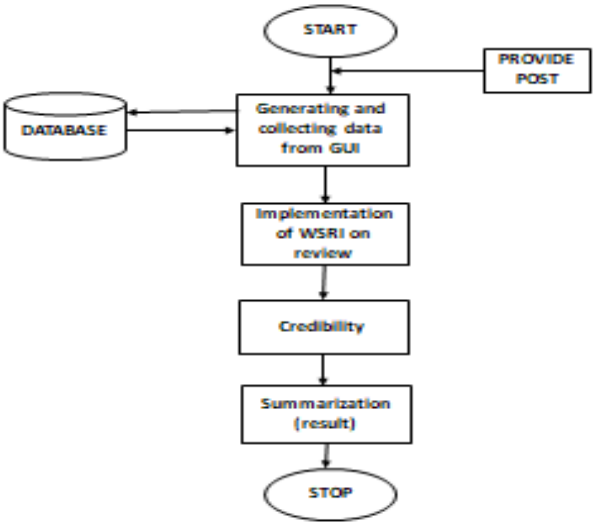


Fig 2 Credibility analysis by using WSRI

III- PROBLEM FORMATION

The architecture of our proposed system is illustrated in Figure 1. It consists of five major procedures labeled as follows: 1) tweet collecting and repository, 2) credibility scoring technique, 3) reputation scoring technique, 4) user experience measuring technique, and 5) trustworthiness value, the last of which is an output of the preceding three techniques. In principle, all these mechanisms together represent an iterative process that combines an automated-based methodology for achieving better credibility or trustworthiness results with sophisticated accuracy.

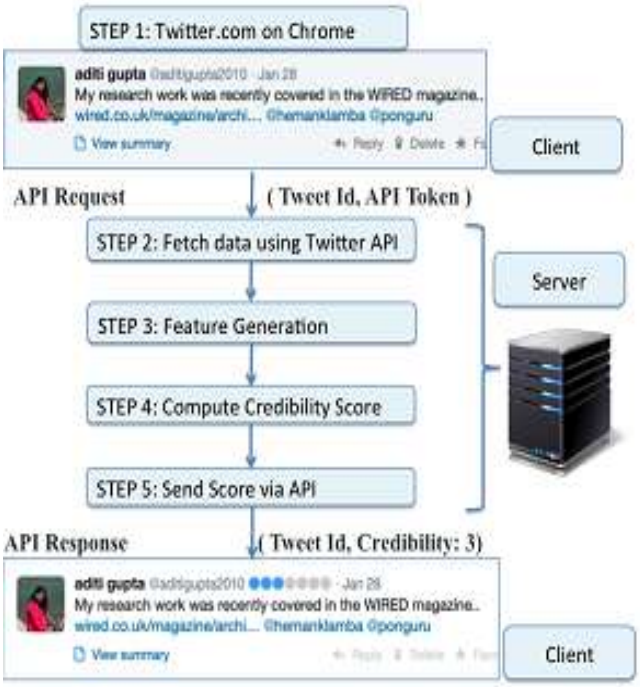


Fig 3: Data flow steps of the TweetCred extension and API.

Tweets are collected using two different Twitter application programming interfaces (APIs): a streaming API and an API for searching for tweets regarding different events. The streaming API is used to collect datasets on given events. The search API is used to collected users' tweets histories simultaneously. On a database server, the data are organized, processed (Step 2), and made available for analysis (Step 3). The prepared data are divided into three groups: tweet content, users who post that content, and the histories of those users (Step 4). These groups of data are passed as inputs to the three techniques to look for signals of truth and credibility (Step 5). The reputation-based technique (Step 6) does not consider aspects such as message content features but does consider factors such as the structure of the network in its model. The credibility technique (Step 7) relies on machine learning methods that are based on training with established ground truth while user expertise method applying both techniques in establishing the reliability of users. Finally, all of the scores obtained using the three techniques are combined to obtain the trust-worthiness value of a given tweet (Step 8). In this paper, we focus particular attention on how to extract and clean data in Step 2, how to calculation reputation scores in Step 6, the credibility assessment mechanism in Step 7, and finally and most importantly, how the users' experience is calculated.

.

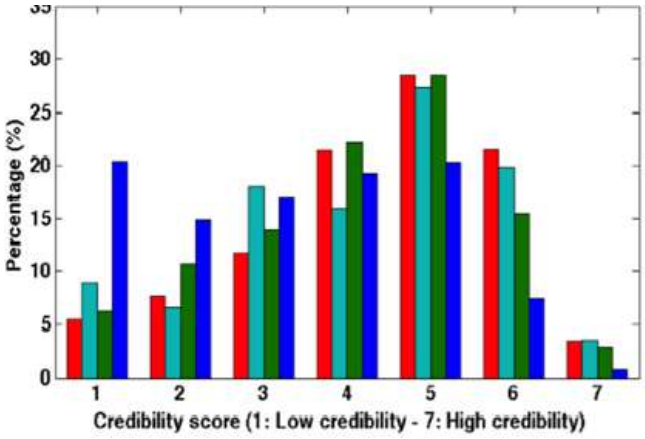


Fig 4: Distribution of credibility scores (1=low, 7=high) as given by TweetCred. We observe that during crisis events there are more tweets with high credibility.

IV- TWITTER DATASET

1) 476 million twitter tweets: this datasets is estimated to comprise about 20-30% of all public tweets posted over the 7-month period between june 1 and December 31, 2009.
2) Sentiment140: with emoticons removed and six formatting categories, this collection of 160,600 tweets is particularly useful for brand management and polling purposes.

3) Customer support on twitter: this dataset on kaggle includes over 3 million tweets and replies from the biggest brands on twitter.

4) cheng-caverlee-lee September 2009~January 2010 twitter scrape: this datasets is collection of scraped public twitter updates used in coordination with an academic project to study the geolocation data related to twittering.

5) Follow the hashtag: twitter analysis tool that includes a section where different, complete large datasets are regularly uploaded in a ready-to-use format.

V- CONCLUSION

This paper presents the results of a study of the problem of assessing information credibility on twitter. The issue of information credibility has come under security, especially in social networks that are now being used activity as first sources of information. Twitter and other social networks have become widely used in disaster mitigation in cases of high impact events because they make it possible for relevant parties to obtain important information sufficiently quickly to coordinate countermeasures to such events. Positive and negative ranking play important role to find the credibility. So it find the credibility according to the comments by scoring and give the final result weather it is credible or not after finding the credibility the main aim is to check whether the user is trustworthy or not.

REFERENCES

[1] *Lorek , K., Suehiro-Wiciński , J., Jankowski-Lorek , M., & Gupta, A. (2015). Automated credibility assessment on twitter. Computer Science, 16(2)), 157-168.*

[2] *El Ballouli , R., El-Hajj, W., Ghandour , A., Elbassuoni , S., Hajj, H., & Shaban, K. (2017). CAT: Credibility Analysis of Arabic Content on Twitter. In Proceedings of the Third Arabic Natural Language Processing Workshop (pp. 62-71).*

[3] *Gupta, M., Zhao, P., & Han, J. (2012, April). Evaluating event credibility on twitter. In Proceedings of the 2012 SIAM International Conference on Data Mining (pp. 153-164). Society for Industrial and Applied Mathematics.*

[4] *Popat, K. (2017, April). Assessing the Credibility of Claims on the Web. In Proceedings of the 26th International Conference on World Wide Web Companion (pp. 735-739). International World Wide Web Conferences Steering Committee.*

[5] *Shariff, S. M., Zhang, X., & Sanderson, M. (2017). On the credibility perception of news on Twitter: Readers, topics and features. Computers in Human Behavior 75,785-796*

[6] *Mitra, T., & Gilbert, E. (2015, May). CREDBANK: A Large-Scale Social Media Corpus With Associated Credibility Annotations. In ICWSM (pp. 258-267).*

[7] *Mitra, T., Wright, G., & Gilbert, E. (2016). Credibility and Dynamics of Collective Attention. arXiv preprint arXiv:1612.08440.*

[8] *Fairbanks, J., Fitch, N., Knauf, N., & Briscoe, E. (2018). Credibility Assessment in the News: Do we need to read?.*

[9] *Ross, J., & Thirunarayan, K. (2016, October). Features for ranking tweets based on credibility and newsworthiness. In 2016 International Conference on Collaboration Technologies and Systems (CTS) (pp. 18-25). IEEE.*

[10] *Gupta, A., Kumaraguru, P., Castillo, C., & Meier, P. (2014, November). Tweetcred: Real-time credibility assessment of content on twitter. In International Conference on Social Informatics (pp. 228-243). Springer, Cham.*

Details of All Authors (Optional)

Sr.No	Photo	Details
1		Name:- Bhagyashree Durgekar Ph.no:– 9156580175 Email ID:- bhagyashreedurgekar@gmail.com
2		Name **:-** Pranoti Patil Email-Id **:-** 7249179234 Contact No **:-** pranotipatil02@gmail.com
3		Name **:-** Dharmendra dhapade Email-Id**:-** 8793893391 Contact No **:-** dharmendradhapade@gmail.com
4		Name **:-** Damini Gahane Email-Id **:-** damini15198@gmail.com Contact No**:-** 8408913183
5		Name:- Prof .C. R Pote Ph.no:– 7719966016 Email ID:-scpote@gmail.com