Detection Of Social Interaction With Rumor Through Social Network Using NLP And Random Forest Classifier

Ammar Adl
Faculty Of Computer And Information System
Beni-Suef University
ammar@fcis.bsu.edu.eg

Shimaa Eid
Faculty Of Computer And Information System
Beni-Suef University
shimaaeid_sd@fcis.bsu.edu.eg

ABSTRACT-- The rumors detection problem on the social network has attracted considerable attention in recent years. As the rumors propagation has been increasing every day in online social networks (OSNs), it is important to analyze and understand this phenomenon. Damage caused by the rumors propagation is difficult to handle without a full understanding of the dynamics behind it. We question this belief and argue that the rumors differ in how they engage their audience, how they are spread and by whom, the type of users who interact with them, and how they evolve over time. One of the central steps of understanding rumors propagation is to analyze social interaction with rumors, we implement this general approach using social network data. We find that we can reliably classify users into one of five categories: (1) "Generally support a rumor", (2) "Generally deny a rumor", or (3) "Generally joke about the rumor". (4) "View Generally neutral". (5) "Request more information (asking) "Combining text mining techniques, such as text analysis, machine learning (NLP), (Random Forest Classifier), and social network analysis, we aim to identify and classify those user categories which interact with rumor automatically and provide a more holistic view of rumors propagation in OSNs.

Keywords— Social Network Analysis, Content Analysis, Text Mining, Machine Learning.

1- Introduction

Social networking sites have become a major platform for spreading rumors in recent times and this rumor effect on individuals, false rumors can spread fear, hate, or even euphoria. They may lead to defamation, protests or other undesirable responses, Rumors differ in how and where they originate, how they spread, and how users interact with them, I need to detect the rumor through social network, In this Paper I study how people interact with the rumor through those points [1] How user interact with rumor. [2] Analyzing user characteristics belief and rumor propagation.[3] Some support, deny, joke, neutral, asking. [4] Understanding user beliefs to rumor. This research is done to achieve Detecting the social interaction with rumor through social network by following this [1]

Collecting and analyzing user posting behaviors in the Social network about a specific rumor. Based on users' interaction, determine if there is a group of users that are actively spreading rumors. [2] Using social network analysis, content analysis, and text mining techniques, the system classifies the active rumor-spreading users into one of the five categories: [1] Support, [2] Deny, [3] Joke, [4] Neutral, and [5] Asking. This paper covers Detecting the social interaction with rumor through social network using NLP And Random Forest Classifier and this research will achieve accuracy 90%. Combining the machine learning technique [NLP, Random Forest Classifier] in the text processing pre-learning stage, adding a social network analysis.

To date, most of the work in this emerging area has been conducted to detect rumors, limit the spread of rumors, and identify the source of rumors. However, in order to develop effective methods for rumor detection and prevention in OSNs, we first need to understand who spreads rumors online and why. This motivates us to propose the following research statements:

1. Based on user activities on Facebook and Twitter, could we determine if there is a specific group of users that is greatly interested in discussing and spreading rumors?
2. Based on user activities in Facebook could we determine if there is a rumor propagation personality type in Facebook and Twitter who, for example, “Generally supports a false rumor” (Support), “Generally refutes a false rumor” (Deny), or “Generally jokes about a false rumor” (Joke) ,or "Generally have neutral opinion"(Neutral), and "Generally asking about more information"(Asking).
3. Will visualizing rumor spread in Facebook provides better insight into how users interact with rumors?

THIS PAPER MAKES THE FOLLOWING CONTRIBUTIONS:

1. Collecting and analyzing user posting behaviors on Facebook and Twitter about a specific rumor. Based on users' interaction, and determine if there is a group of users that are actively spreading rumors.
2. Using social network analysis, visual analysis, content analysis, text mining techniques, and machine learning, the system classifies the active rumor propagation users into one of the five categories:: [1] Support, [2] Deny, [3] Joke, [4] Neutral , and [5] Asking.
3. The experimental results using text mining techniques and machine learning will confirm and support our approach.

This paper has the following structure: Section II reviews related work, Section III Methodology, Section IV Dataset, and Section V analyzes and discusses the results.

2- Related Work

Previous work in this area is concentrated in three main areas: mining online social networks, rumor analysis, and visualizing rumor spread in OSNs. It is important to highlight that the research focuses on rumor spreading on social media. It is implemented for Reddit data, and illustrated with the "Obama is a Muslim" rumor.

Boidiou (Boidiou et al. 2014) created a corpus of tweets around big events, focusing on the ones linking to images the reliability of which could be verified by independent online sources. Castillo et al. (Castillo, Mendoza, and Poblete 2011) studied the credibility of users and proposed a set of features that were able to retroactively predict the users' credibility. Finn (Finn, *Metaxas, and* Mustafa raj 2014) offered Twitter Trails, a web-based tool, that allows users to investigate the origin and propagation characteristics of a rumor and its reputation, if any, on Twitter. Ratkiewicz et al. (Ratkiewicz et al. 2011) created the 'Truthy' service identifying misleading political memes on Twitter using tweet features.

Many studies have focused on identifying false rumors, but very few papers have concentrated on analyzing them. Friggeri et al. (Friggeri et al. 2014) inspected various topics discussed on Facebook. Liao and Shi (Liao and Shi 2013) analyzed the statements users made in response to an infamous rumor that spread through Sina Weibo and observed that throughout the lifetime of the rumor, different response types could become more popular depending on their functional role. Maddock et al. analyzed four rumors that spread through Twitter after the 2013 Boston Marathon Bombings (Maddock et al. 2015).

In a similar vein, we present our study of rumors and their semantic and functional characteristics. Compared to previous literature, our research is based on a much larger data set, comprising more than 1000 rumors.

3- Methodology.

We define a rumor as a "circulating story of questionable veracity, which is apparently credible but hard to verify, and produces sufficient skepticism and/or anxiety so as to motivate finding out the actual truth" (Zubiaga et al., 2015b), or "item of circulating information which not verified from veracity at the time of posting". when posting important information, users seek to get more data about this information and increase their comments on the publication, We will analyze this publication after a certain period of user response. While breaking news unfolds, gathering opinions and evidence from as many sources as possible as communities react becomes crucial to determine the veracity of rumors and consequently reduce the impact of the spread of misinformation. We describe our general approach to collecting, visualizing, and analyzing rumor related data using Facebook and Twitter as a specific implementation example. To study the spread of rumors on Facebook and Twitter, we need the following elements: 1) A rumor. 2) The truth about this rumor. 3) Posts about this rumor. and 4) Comments about this rumor.

This research is very useful for detecting social interaction with rumors through the social network used NLP and Random Forest Classifier algorithm, Data analysis for examining the meaning of textual data manually to identify and assess themes patterns, natural language processing, Data Mining by mining the data analysis, and social network to know how to deal with data.

Approach To Analysis:

After collecting the data, we adopted social network analysis, content analysis, and Data mining techniques, to analyze and visualize these contents.

Social network analysis (SNA) refers to the use of network theory for understanding social network data. Because of the analogy between online social networks and the structure of social hierarchy and stratification. We focus on how users interact with other users about this rumor using our *visualization* tool to explore the data, particularly connections between users and a longitudinal assessment of the prevalence and re-appearance of the rumor over the 9 years in our dataset.
Content analysis is a qualitative method that examines the meaning of textual data manually to identify and assess themes and patterns. We focus on the characteristics and content quality of user posts in each category of user (Support, Deny, Joke, Neutral, and Asking) and manually review each comment in all five categories to identify typical patterns and themes.
We explored data in order to bring important aspects of that data into focus for further analysis. We used an ordinal variable which is a type from the categorical variable. An ordinal variable has a clear ordering. For Example. In our research, social interaction with rumors divides into 5 categories (Support, Deny, Joke, Neutral, and Asking).

User Category	**Description**
Support	People who believe the rumor is true and is usually very short and had no backup evidence or explanation
Deny	People who believe the rumor is false and many comments are very thoughtful and providing the in-depth explanation.
Joke	Usually made a sarcastic comment or joke to refute this rumor.
Neutral	Not dining or supporting either side in an opinion.
Asking	People want to get more information

Table 1: User categorization.

Finally, text mining techniques, such as data classification, visualization, and sentiment analysis are used to validate if the characteristics of each user group found from social network analysis and content analysis could be classified automatically

.

The framework were as follows:

1- Data Collection.
2- Data Preprocessing.
3- Data Exploration & Visualization.
4- Model Building.
5- Model Evaluation.

The preprocessing stage this is the longest phase and take the longest time in implementation for In order to purify data from impurities. Our proposed rumor detection method is a typical classification problems, and it mainly contains 3 parts which are data cleaning, feature extraction and model training.

Text Preprocessing:-

We used Text Preprocessing framework 1) **Tokenization**: is a step which splits longer strings of text into smaller pieces or tokens. Larger chunks of text can be tokenized into sentences, sentences can be tokenized into words, etc. 2) **Normalization**: text needs to be normalized. Normalization generally refers to a series of related tasks meant to put all text on a level playing field: converting all text to the same case (upper or lower), removing punctuation, converting numbers to their word equivalents, and so on. Normalization puts all words on equal footing and allows processing to proceed uniformly.

Normalizing text can mean performing a number of tasks, but for our framework, we will approach normalization in 3 distinct steps: (1) stemming, (2) lemmatization, and (3) everything else**.**

There are, however, numerous other steps that can be taken to help put all text on equal footing, many of which involve the comparatively simple ideas of substitution or removal. They are, however, no less important to the overall process. These include:

- Set all characters to lowercase.
- Remove numbers (or convert numbers to textual representations).
- Remove punctuation (generally part of tokenization, but still worth keeping in mind at this stage, even as confirmation).
- Strip white space (also generally part of tokenization).
- Remove default stop words (general English stop words).

3) Noise removal: Noise removal continues the substitution tasks of the framework. While the first 2 major steps of our framework (tokenization and normalization) were generally applicable as-is to nearly any text chunk or project noise removal is a much more task-specific section of the framework.
These include:

- Remove text file headers, footers.
- Remove HTML, XML, etc. markup and metadata.
- Extract valuable data from other formats, such as JSON, or from within databases.
- Remove URLs and username with '@' tags.
- Leave important clues such as hashtags and some special characters. Question marks and exclamation marks.

Random Forest Classifier:

To build the random forest algorithm we are going to use the dataset which collected from Facebook and Twitter. We are going to build a random forest classifier to predict users categories which interact with rumors.
Random forest algorithm is an ensemble classification algorithm. Ensemble classifier means a group of classifiers. Instead of using only one classifier to predict the target, In ensemble, we use multiple classifiers to predict the target.
In case, of random forest, these ensemble classifiers are the randomly created decision trees. Each decision tree is a single classifier and the target prediction is based on the majority voting method.
The majority voting concept is same as the political voting. Each person votes per one political party out all the political parties participating in elections. In the same way, every classifier will votes to one target class out of all the target classes.
To declare the election results. The votes will calculate and the party which got the most number of votes treated as the election winner. In the same way, the target class which got the most number of votes considered as the final predicted target class.

System Architecture:

1. Collection dataset from other sources.
2. Input will be textual rumor (posts and comments) and will apply text preprocessing on it (Tokenization, Normalization, and Noise Removal).
3. Appling text mining technique and content analysis.
4. Creating bag of words model.
5. Feature extraction by classifying users comments.
6. Splitting the data into training set and testing set.
7. Filling random forest classification to the training set.
8. Predicting the test set result.

9. Making the confusion matrix.
   Tell now the dataset is classified into 5 categories (Support, Deny, Joke, Neutral, and Asking). After that
10. Calculate the count in each category.
11. Calculate the percentage in each category relative to the number of total comments.
12. Visualize the results on a graph.

I followed some of the stages to solve this problem [1] Collect the data from the social network.[2] This data leads to our machine. [3] Prepressing to this data. [4] Get users which (support, deny, joke, neutral, asking) specific rumor. [5] Calculate the count in each part. [6] Calculate the percentage in each part. [7] Show all through the graph.

Classifying User Comments:

The analysis of Comments refers to the degree of acceptance of the comments in the message. Therefore, opinion analysis carried out on the comments can be used to obtain the credibility of the rumor.

Using the content analysis, we observe that content characteristics in each rumor-spreading user group have its own characteristics. As a result, in this section, we explore if we could determine the user rumor-posting behavior automatically based on its content. For each user that has more than 10 comments, we transform them using the TF-IDF vectors, which reflect how important a word is in a document or a corpus (stop words are removed). Each user is represented by a vector:

After transforming each user comment data into a TF-IDF vector, we apply NaivesBayes classifier to those vectors and classify each user into one of the five groups: Support, Deny, Joke, Neutral, and Asking. Through various parameter settings, we achieve the best result with 90% accuracy using 10 fold cross-validation and the dimension of the vector is 200. The classification result agrees with the manually classified data based on the two human assessors and further supports our hypothesis about the intrinsic content characteristics of each user group.

We classified the rumors into 4 types based on the percentage of each category.

Category	Rumor Type
Support	Real Rumor
Deny	False Rumor
Joke	Unknown
Neutral	Unverified
Asking	Unverified

Table2: Rumor Types

Algorithm

```
If (maximum == 'Support')
  Print ('Real Rumor')
Else if (maximum == 'Deny')
  Print ('False Rumor')
Else if (maximum == 'Joke')
  Print ('Unknown Rumor')
Else if (maximum == 'Neutral')
  Print ('Unverified Rumor')
Else
  Print ('Unverified Rumor with More Question')
```

1- Examples:-

[1] U1: Mohammed went to school today and then fell off from his bicycle and was wounded in the leg #Yes I visited him today and he is injured– [support] u2: @u1 apparently a hoax. Best to take Tweet down. [Deny] u3: @u1 He attended the camp and was not injured. [Deny] u4: @u1He is a cheerful and lovable boy. [Joke] u5: @u4 I called him to make sure, I cannot believe this. [Neutral] u4: @u5 I need more details about Mohamed's injured. [Asking]

[2] Searched the keywords "Mohamed & injured" from Facebook. [2] Collected 195 submissions, 421 comments from 125 users. [3] The number of users with more than 10 comments is 163. Reviewed comments of these 163 users and categorize them into one of the five categories Support, Deny, Joke, Neutral, and Asking.

Frequency Table.

Rumor-discussing Users	User Count	Percentage
Support	95	22.5%
Deny	155	36.8%

Joke	**60**	**14.2%**
Neutral	**100**	**23.7%**
Asking	**90**	**21.3%**

TABLE ٣: RUMOR-DISCUSSING USERS ABOUT THE "MOHAMED IS INJURED".

Beyond understanding the users and their interactions, we also sought to analyze the textual content of submissions in each category. As most users that are actively engaged in this rumor do not believe it is true, we revisited the original dataset, which includes users who commented fewer than 10 times. We found that users in Support usually posted only one or two comments about this rumor. All of these comments were usually very short and had no backup evidence or explanation. Here are a few examples: "He is infected" "I think He is infected".

2- DATASET

The dataset used in this research contains 2000 post as life input for the algorithms. Training dataset contains 1000 training post and 1000 training post. Training post is of 120 words. The training data is collected from different sources. 1- Collecting and analyzing user posting behaviors in the Social network about a specific rumor. Based on users' interaction, determine if there is a group of users that are actively spreading rumors. 2- Using social network analysis, content analysis, and text mining techniques, the system classifies the active rumor-spreading users into one of the five categories: (1) Support, (2) Deny, (3) Joke, (4) Neutral, and (5) Asking.

We collected dataset for all emotions and set the target to help us and set the target for each one. Second, this is dataset for (post, comment count, comments, and target)
Target = (Support, Deny, Joke, Neutral, or Asking)
We collected it from Facebook and Twitter, and analysis it to set the target..

Collecting and analyzing user posting behaviors on Facebook and Twitter Based on users' interaction, this data collected from Facebook and Twitter in range of 2012 to 2018, contain 2000 record, this data from public pages which contain (post, comment count, comments, and target)

٥. RESULTS AND DISCUSSION

1- Visualizing the result on a Gephi tool.
2- Show users which (support, deny, joke, neutral, asking) in a graph.

USERS INTERACTION CATEGORIES.

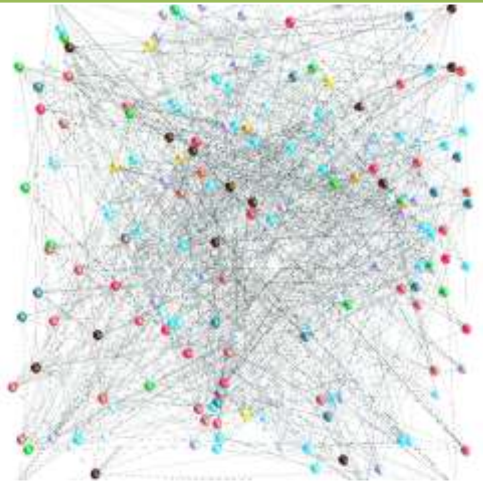

Fig.2: an example of a user interaction graph. Blue nodes: users in SUPPORT category, Red nodes: users in DENY category, Yellow nodes: users in JOKE category, Green nodes: users in NEUTRAL category, and White nodes: users in ASKING category.

3- Calculate the count in each category.

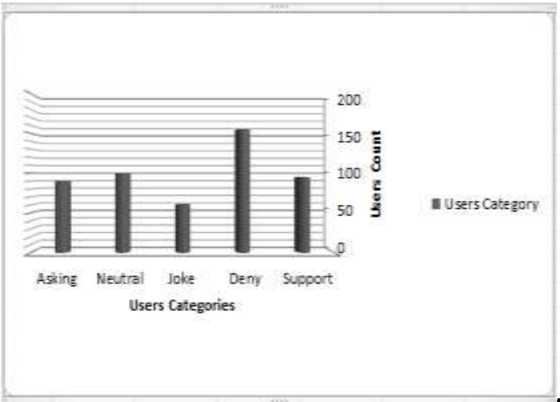

FIG.3: USERS INTERACTION IN EACH CATEGORY.

4- Calculate the percentage in each category

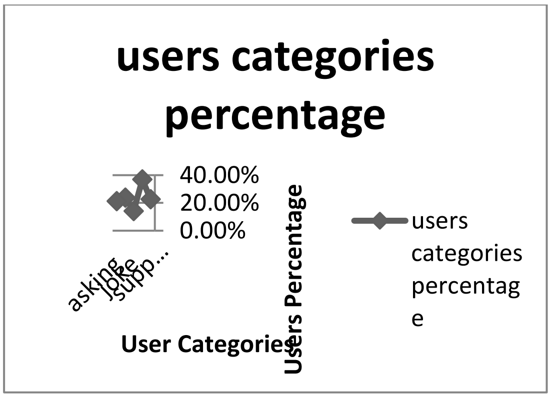

Fig.4: Users Percentage in each Category.

5- User Categories (User interaction with "Mohammed is injured"):-

Category	**Comments**
Support	“Yes, I visited him today and he is injured".
Deny	" Apparently a hoax"?" " He attended the camp and was not injured"
Joke	" He is a cheerful and lovable boy"
Neutral	"Maybe yes maybe not, I called him to make sure I can't believe this. “
Asking	" I need more details about Mohamed's injured"

Table ٤: Examples of User categories.

6- Performance:-

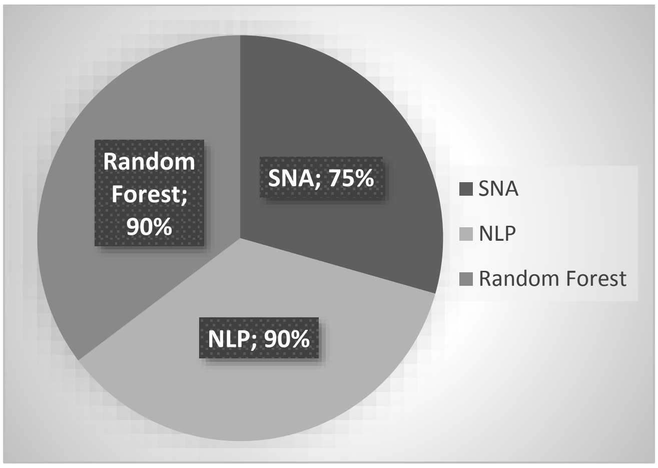

FIG.4: PERFORMANCE FOR EACH TECHNIQUE.

7- Conclusion:-

Detecting and verifying rumors is a critical task and in the current media landscape, vital to populations so they can make decisions based on the truth. This shared task brought together many approaches to fixing veracity in real media, working through community interactions and claims made on the web. Many systems were able to achieve good results on unraveling the argument around various claims, finding out whether a discussion supports, denies, questions or comments on rumors. In this research, we developed a framework to detect social interaction with rumor through the social network. Using a set of 2282 distinct rumors, we organized the study around capturing various semantic aspects of rumors. We proposed a methodology to characterize Belief and captured how it evolves with time. We characterized rumor usage and determined the roles various user-types play, and how they vary with respect to Beliefs. Ultimately, our characterization covers usage analysis of rumors and aids in creating systems that uncover news from the Social network by eliminating rumors automatically.

8- Reference:

1- Dang, A., Smit, M., Moh'd, A., Minghim, R. and Milios, E., 2016, August. Toward understanding how users respond to rumors in social media. In Advances in Social Networks Analysis and Mining (ASONAM), 2016 IEEE/ACM International Conference on (pp. 777-784). IEEE.
2- C. Buntain and J. Golbeck, “Identifying social roles in Reddit using network structure,” in Proceedings of the Companion Publication of the 23rd International Conference on World Wide Web Companion. International World Wide Web Conferences Steering Committee, 2014, pp. 615–620.
3- Boididou, C.; Papadopoulos, S.; Kompatsiaris, Y.; Schifferes, S.; and Newman, N. 2014. Challenges of computational verification in social multimedia. In Proc. of the 23rd International Conference on World Wide Web Companion, 743–748.
4- Liao, Q., and Shi, L. 2013. She gets a sports car from our donation: rumor transmission in a Chinese microblogging community. In Proceedings of the 2013 Conference on Computer Supported Cooperative Work, 587–598. ACM.
5- Arkaitz Zubiaga, Maria Liakata, Rob Procter, Geraldine Wong Sak Hoi, and Peter Tolmie. 2016b. Analyzing how people orient to and spread rumors in social media by PLoS ONE 11(3):1–29..
6- J. Ratkiewicz, M. Conover, M. Meiss, B. Gonçalves, S. Patil, A. Flammini, and F. Menczer, “Truthy: mapping the spread of AstroTurf in microblog streams,” in Proceedings of the 20th International Conference Companion on World Wide Web. ACM, 2011, pp. 249–252
7- Rob Procter, Farida Vis, and Alex Voss. 2013b. Reading the riots on Twitter: methodological innovation for the analysis of big data. International journal of social research methodology 16(3):197–214
8- Dervis Karaboga · Bahriye Basturk A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm
9- Dervis Karaboga and Bahriye Basturk Artificial Bee Colony (ABC) Optimization Algorithm for Solving Constrained Optimization Problems
10- Ramos, Juan. "Using tf-idf to determine word relevance in document queries." Proceedings of the first instructional conference on machine learning. Vol. 242. 2003.

.

.