NOVATEUR PUBLICATIONS
International Journal of Research Publications in Engineering and Technology [IJRPET]
ISSN: 2454-7875
VOLUME 3, ISSUE 7, July-2017

DECIDING METRICS FOR DETECTING FALSE NEWS AND INFLUENCERS ON TWITTER

SHANTANU KUMAR
Lovely Professional University, Jalandhar, Punjab Email: shantanu01497@outlook.com

ABSTRACT:

The era of the social media brings with it quite some information and news that is untrue and misleading to the masses. This issue is such that needs to be addressed as people today are highly dependent on their mobile devices and online news portals that they tend to make uninformed decisions based on these pieces of false information they get.

The purpose of this paper is to help people and industries alike, to differentiate good quality sources from bad ones based on a certain distinction of metrics and methodologies. This paper probes into current trends adopted by fake information promulgators and how to spot and work around those.

INDEX TERMS: Fake influencers, False news, Misleading, Social Media

I. INTRODUCTION:

The major cause of wrong decisions and misleading information in today's generation is the social media. Most times when we read something on the internet, we don't verify its source and skip right to the information which is fake and sculpted.

This can cause havoc in societies where the use of social media does not entirely correlate to literacy. Twitter has over 320 million users active and if every person began to post even one tweet everyday whilst being connected to an average of 100 people, there would be a lot of opinion and a lot of information from different viewpoints that will be hard to categorize into the brackets of real and fake.

This paper will demonstrate based on some metrics chosen to decide how we can spot fake social media influencers using a dataset drawn from the popular social media giant, Twitter.

II. METRICS CHOSEN:

A. NUMBER OF TWEETS FROM A GIVEN USER:

This would help us analyze if they are an active user or not, and posting often about any topics in their field.

B. NUMBER OF TIMES THE USER HAS TWEETED WITH AN AUTHORITY RATING OF OVER 7:

We have an authority rating of every post out of 10, that tells us how influential each post was, with 10 being highly influential and having widespread reach. This will help us visualize how every author fared with his posts. We can have users posting multiple times but if none of the authority ratings go over a certain bar, we will know that their posts have no influence or importance and we can flag it as a source.

C. USAGE OF SIMILAR WORDS IN TWEETS:

The usage of similar words in topics shows that the relevance of tweets is maintained and not deviating.

D. NUMBER OF TIMES MENTIONED IN RETWEETS:

The more the user is mentioned in relevant tweets, the higher the influence they have. If the user belongs to the textile industry and is mentioned in retweets that are related to the textile market, then we can assume he's influencing the topic in some way.

E. THE TIMES THEY WERE A PART OF A TOPIC IN A CLUSTER:

If a user is mentioned in topics related to their field, this shows they are valued in those discussions and thereby making this a strong metric for influencer analysis.

F. THE SENTIMENT ANALYSIS OF RETWEETS THEY WERE MENTIONED IN:

This is one of the most important metrics for the analysis. We have chosen retweets as a basis because they give the highest information about a user. When someone retweets a user with a negative sentimental feedback, there's a high chance they are not an influence in the field and vice versa.

III. ANALYZING METRICS:

The first metric will help us set the hypothesis about our influencers.

Figure 1: Top Authors Based on Number of Tweets

Visualizing a word-cloud based on top authors from Figure 1, we see that one of the most influential authors/users based on metric A is “assamgreentea”. A word-cloud is a visualization that brings the maximum frequency terms to the center of the plot and expands it in size, amplifying their presence. This helps us believe that the user is an influencer based on the number of times they tweet.

Figure 2: Top Content Based on Number of Times Used in Tweets

The above visuals from Figure 2show the top used content, which would bring us to the conclusion that the authors who talk about topics like “green”, “tea’, “weight” tend to be the ones who have the maximum number of tweets as well.Now according to these, “assamgreentea” looks to us as an influencer. Let’s look at metric B to see if we can find similar users.

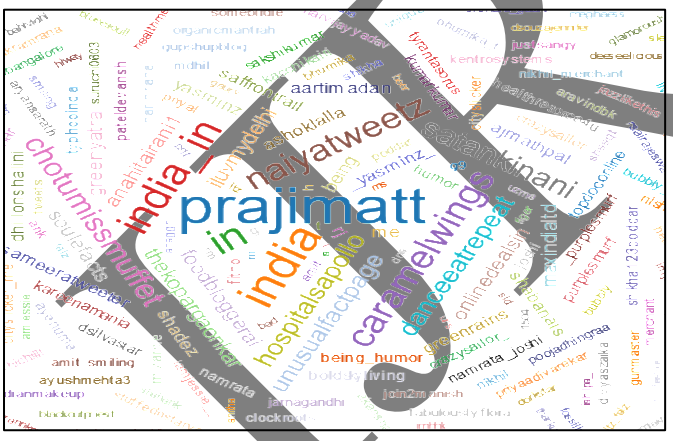

Figure 3: Top Authors based on Average Authority Rating of Over 7

The word-cloud in Figure 3 shows us that the authors that were actually tweeting the most in Figure 1were not the ones that were most influential. The user, “assamgreentea” is nowhere to be seen in the cloud of users with an average authority rating of over 7 in their tweets.

This should raise concern as a very important aspect of influencing users is for your posts to have some value given by them, which is not the case here. Let’s look at some content that got an average authority rating of over 7.

Figure4: Top Content Terms Gaining Average Authority Rating of Over 7

We see from the above cloud that the content gaining maximum popularity is related to green tea, weight loss, home delivery etc.

Analyzing metric C, we will go over associativity rules to see if the words in a topic remain in sync to the topic. For example, if a tea vendor is talking about tea and then starts talking about textiles, we will know that this could be a misleading influencer or an advertiser and can flag it accordingly.

ASSOCIATIVITY RULES
Rule 1: contents includes tea <-> contents includes green [Coverage=60.86% (84732); Support=58.13% (80931); Confidence=95.51%; Leverage=0.22559; Lift=1.63417; p-value=0]
Rule 2: contents includes tea -> 3 < authority <= 5 [Coverage=60.86% (84732); Support=16.09% (22396); Confidence=26.43%; Leverage=0.0293; Lift=1.2227; p-value=0] Rule 3: contents includes tea -> contents includes weight [Coverage=60.86% (84732); Support=5.95% (8282); Confidence=9.77%; Leverage=0.02016; Lift=1.51267; p-value=0] Rule 4: contents includes tea -> contents includes deliver [Coverage=60.86% (84732); Support=4.96% (6911); Confidence=8.16%; Leverage=0.01907; Lift=1.62381; p-value=0]
Rule 5: 3 < authority <= 5 -> contents includes green & contents includes tea[Coverage=21.62% (30096); Support=15.59% (21706); Confidence=72.12%; Leverage=0.02956; Lift=1.23396; p-value=0]

Figure 5: Clusters Based on Topics

Rule 1 tells us that if contents include either green or tea, they include the other word as well and that covers roughly 61% of the data.

Rule 2 states that the content including the word tea again has a low authority rating of between 3 and 5 in majority of the tweets, with a coverage of 61%.

Rule 3 tells us that content with the word tea includes some relevance around a person's weight.

Rule 4 says that when a person talks about tea, he also looks for delivery options.

Rule 5 states something similar to Rule 2, but reinforces that green tea in specific contains a low authority rating of between 3 and 5.

The visualization of specific topics on the right will clearly specify trending topics and what the content was in them.

Running through certain topics, we can see that the first topic talks about the health benefits of certain products.

Topic #4 shows us the mention of "Starbucks" in a topic cluster, which would help us reassure that it is an influence in the field it belongs to.

Topic #10 shows us a similar mention of "Barista" in the topics telling us that "Starbucks" and "Barista" may be competing for influence in the target group of their cluster.

Topic #19 shows us the health associated problems and solutions to a specific issue, like inflammation, and how to get rid of it, antioxidants. The clusters formed of related topics link them by using the current set of metrics.

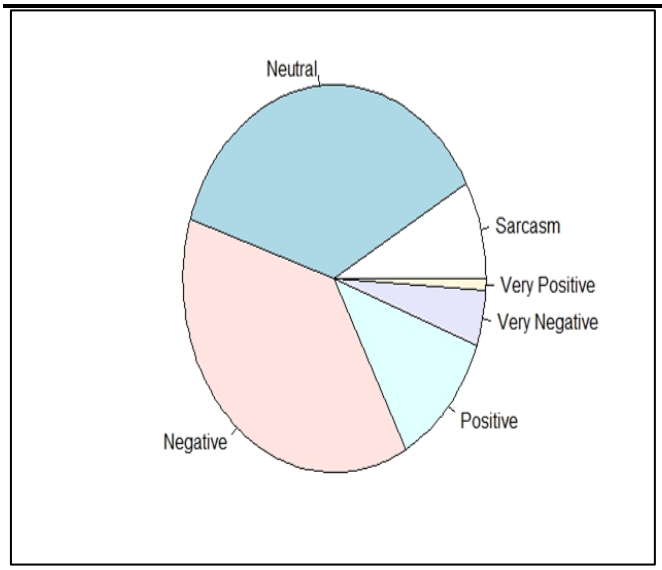

Figure 6: Sentiment Analysis of Retweets

Coming to a very important and the final metric F, we have now already seen a couple of measures and signs to look out for while classifying fake news and influencers.

The final metric analyses the retweets that the influencers have been mentioned in. The retweets hold for very high importance in the study as when a person takes out the effort to say something about somebody's tweet or work, it tells more about the original tweet than any other metric may.

Analyzing these, we have noted that a lot of people tend to negatively respond to tweets on an average, bringing us to a conclusion that most of the influencers and information promulgators on Twitter seem to be delving away from reality, angering followers. These can also be advertisements or personal opinions so we can flag the user depending on the number of tweets we analyze for a particular user.

IV. RESEARCH METHODOLOGY:

A dataset containing 1,32,915 rows was drawn from Twitter using the API and package on R (twitteR). Since the data was raw and unclear to process, we had to strip a lot of the text down to the exact word meanings, remove URLs and stem the document completely. Term Document Matrices were used to calculate term frequencies and for sentiment analysis, the package R Sentiment was used. Clustering the data is one of the most important parts, and the K-Means algorithm was used to define the clusters. There are two types of clustering algorithms that were tried out, namely K-Means and G-Means. The K-Means defines a specific k number of clusters and then attempts to split the data into Voronoi cells.

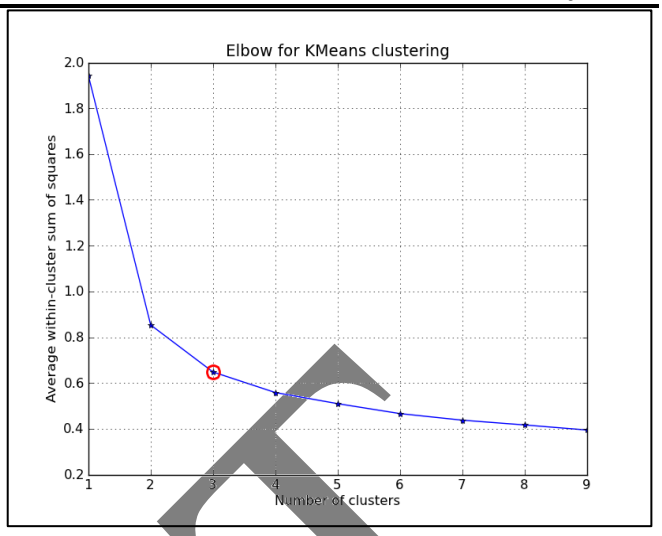

Figure 7: The Elbow Method Depiction to Find Best Number of Clusters for A Dataset

Source: https://stackoverflow.com/questions/6645895/calculating-the-percentage-of-variance-measure-for-k-means On the other hand, the G-means algorithm (Gaussian-means) is a sort of recursive K-means where it starts off with k=1 and builds on. This is a more useful algorithm as it helps us when we do not know how many clusters our algorithm will fit best in.

V. CONCLUSION:

The need to check our sources for any kind of information derived from the internet today is more than ever. With organizations paying huge amounts to influencers to gain post reach about their product or service, more and more people are aiming to breach the bracket of social media influencers. This has caused a certain imbalance between real information and fake rants, making it harder to differentiate between the two.

Comparing each metric and its visualization, the first thing we notice is that the content in authority ratings of over 7 is similar to the content in the top used terms. This shows us that metric A and B are in sync when it comes to content, but not in the case of authors.

This tells us that some people know about the terms that usually amass a lot of reach in the media, and are frantically trying to make an impact by posting repeatedly about the same terms.We will flag "assamgreentea" as a fake influencer based on metric A and B.

Coming to the rules, we will note that if a user belongs to the category of either rule #2 or rule #5, he is posting frequently about influential topics but still not gaining any authority. Such a user will be flagged as a fake influencer too.

The topic modeling will show us a very important part of flagging fake influencers. If they are talking about a particular topic and then hop onto

another one, this might tell us something about their behavior and help us flag them as a possible fake influencer.

The topic cluster tells us about some correlated topics and if a user is flagged as a possible faker, we can verify that here. If he switches topics between correlated clusters, then we can remove his flag; but if not, we have to confirm his flag as a fake influencer.

The sentiment analysis brings special insights to our analysis. If a user is retweeted and the sentiment is negative or similar, we will flag him as a possible faker and go over other metrics to verify this claim.

Combining these metrics will help us find a possible user on Twitter or any other Social Media Platform, that is trying to become an influencer without really possessing the required knowledge or skills; thereby promulgating misleading information that may harm readers.

REFERENCES:

1) Vandana Korde et al *Text classification and classifiers*:" International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.2, March 2012".
2) Dash, M. & Liu, H. (1997) "*Feature Selection for Classification".* Intelligent Data Analysis, Vol.1, no.3, pp. 131-156.
3) Fraley, C. and Raftery, A. (2002). *Model-based clustering, discriminant analysis, and density estimation*. Journal of the American Statistical Association, 97(458):611–631.
4) Blei, D.M., and Lafferty, J. D. —*Dynamic Topic Models*||, Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, PA, 2006.
5) *"Semantic Word Cloud Visualization - Description",* Wordcloud.cs.arizona.edu, 2017. [Online]. Available: http:// word cloud. cs. arizona.edu/description.html.