

A Data Mining Approach of Detection of Fake News on Social Media

B.Umamaheswari, Dr. Vijeta Kumawat
CSE Dept, JECRC

Abstract:

The advantage of easy accessibility, less expensive and faster reach to more people in less time has made the news on social media gaining popularity in recent years but at the same time it faced the problem on quality of news spread compared to other standard traditional media such as newspaper and TV channels. How the people are victimized by the fake news is the major concern. Also identifying such fake news has become almost a challenge. In this paper we will see how data mining is providing solution to identify fake news on social media.

Keywords— Fake news, Traditional media, social media , detection

INTRODUCTION

Social media originated in 1970 with the introduction of emails and slowly gained popularity when six degrees the first social media site was launched in 1997. The lack of interaction in traditional media also fuelled the growth of social networking sites. The social media is less expensive and it facilitates easy interaction with others. Literally speaking nowadays people are much more interested in accessing social media than traditional media. This created fake news to spread easily across the social networking sites. The main purpose of this paper is summarized as follows:

- There is no proper knowledge of fake news among online media users. Knowingly or unknowingly they are

spreading the fake news. At the same time they are the victims of the same. Educating the common people about the fake news is the primary task.

- There is no standard method for detecting fake news. Whatever prevailing is in development level only.

Before detecting the fake news we need to give proper definition of fake news and explain its characteristics. After that we can provide some approaches to detect them by considering some metrics. We can discuss related areas followed by issues in detection and how to solve those issues. Finally we will conclude the review.

FAKE NEWS DEFINITION AND FEATURES

Here defining the fake news is important in both traditional and social media. Fake news is the deception information under the semblance of genuine news. Even now there is no customary way to define the fake news. Everyone tried to define it in terms of legitimacy and purpose. Fake news always contains some counterfeit information about the original news. It may be related to some person, event, etc. Also the main intension is to damage someone identity. It is created to delude the reader.

The following ideas don't seem to be fake news as per our study:

- a) Sarcasm news with proper context that has no intent to mislead or deceive consumers and is unlikely to be mis-perceived as factual
- b) Rumors that didn't originate from news events
- c) Conspiracy theories that can be demonstrated as true or false
- d) Information that is created unintentionally and finally
- e) Hoaxes that are solely impelled by fun or to scam targeted individuals

FAKE NEWS ON TRADITIONAL MEDIA

Fake news is common even from traditional media itself. Traditional media normally target psychology of the people to spread the fake news. People who follow the particular news media for years strongly believe the news published is always the authentic news. Any negative comment against the media is normally neglected by them. They

normally support the media who go along with their ideology.

The next way to spread the fake news in traditional media is to influence the society we belong. People believe the news only in the way they wanted. Other way we can say if the news is benefit for them then they believe them. The theory behind fake news can be better understood considering the publisher and reader. The publisher has two motives one is to make profit and second giving unbiased news to maintain their reputation. Similarly the reader has two motives one is to receiving the true and unbiased news and second is to satisfy their social need

FAKE NEWS ON SOCIAL MEDIA

Fake news is even more common in social media because it is easy to create an account in social media. Also malevolent account can be created with fake identity. In that case user need not be real human being. They are normally social bot, cyborg and trolls. A social bot [2] refers to a social media account that is controlled by a computer algorithm to automatically produce content and interact with humans (or other bot users) on social media. Social bots [3] can turn into malevolent unit designed specifically with the purpose to do harm, such as manipulating and spreading fake news on social media.

Trolls, real person who aim to deliberately disrupt online communities [6] and incite clients into an emotional response, are also playing an important role in spreading fake

news on social media. Trolling behaviors are highly affected by people's mood and the context of online discussions, which enables the easy dissemination of fake news among otherwise "normal" online communities. The effect of angling is to trigger people's inner negative emotions, like anger and fear, resulting in doubt, distrust, and irrational behavior.

Finally, cyborg users can easily unfold fake news in a way that blends automated activities with human input. Usually cyborg [1] accounts are registered by human as a camouflage and set automated programs to perform activities in social media. The easy switch of functionalities between human and bot offers cyborg users distinctive opportunities to unfold fake news. In a nutshell, these highly active and partisan malicious accounts on social media become the powerful sources and proliferation of fake news

In Social media news will usually seen by intended users. Not all groups of people see the same news like traditional medium. People with same ideology can easily form groups and spread their opinion across. As a result, this creates segmented, homogeneous communities with a very limited information ecosystem. Research shows that such communities become the primary driver of information transmission that further strengthens division.

FAKE NEWS DETECTION USING DATA MINING

In traditional news medium fake news detection totally depend on the content of the news. While in social media in addition

to the news content milieu plays an important role in detecting fakenews. Milieu is nothing but social and physical environment in which people reside. There are two stages in detection. First extract the features [5] based on content of news and milieu. Second construct the model based on features extracted.

Features extraction

In the news content, features are nothing but meta information that are available in the news. First is the source who originated the news. Second catchy title given to those news. Then comes the essence of original content of the news. Finally other supporting media like images, audio and video clips. From these meta information feature representation can be constructed to retrieve fake news characteristics. Also fake news are always intentionally introduced mostly for exploit image of someone or for political gain. They are available as either hyperlinks, clickbaits or as cartoon. Linguistic features of language and visual media are given much more importance because they can capture fake news.

Linguistic features enable us to extract features in the form of characters, words, sentence etc. Visual features can be easily extracted from images and videos. Faking images were identified based on various user-level and tweet-level hand-crafted features using classification framework. In recent times, assorted illustration and statistical features are being extracted for news verification. In social context, features are user related and post related. In user based, features of individual users or group

of users are captured from their profiles and used for fake news detection. Individual features are important for user reliability and group features help to analyze how the members of group react to any particular news. In the post based emotional reaction of people towards any news are recorded for extracting features.

Model Construction

Based on the content of the news, the news content models are created based on the fact of the news. They are mainly for checking the truthfulness of news in the form of fact checking. The fact is checked either by some expert groups who can confirm whether the news is true or not. Also if news is supported or rejected by a large number of people then in that case their opinion is also considered for model construction [6]. After the news content the style or presentation of the news are given much more importance for model construction.

CONCLUSION

Nowadays social media has become the prime source regarding consumption of news compared to traditional news media. But it also faced the problem of spreading fake news. Such things affect not only individual member but social communities also. In this paper we are analyzing fake news characteristics and its problem. Also we reviewed how to detect them from data mining point of view.

REFERENCES

[1] Zi Chu, Steven Gianvecchio, Haining Wang, and Sushil Jajodia. Detecting automation of twitter accounts: Are you a human, bot, or

cyborg? IEEE Transactions on Dependable and Secure Computing, 9(6):811–824, 2012.

[2] Alessandro Bessi and Emilio Ferrara. Social bots distort the 2016 us presidential election online discussion. First Monday, 21(11), 2016.

[3] Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. The rise of social bots. Communications of the ACM, 59(7):96–104, 2016.

[4] Justin Cheng, Michael Bernstein, Cristian Danescu-Niculescu-Mizil, and Jure Leskovec. Anyone can become a troll: Causes of trolling behavior in online discussions. In CSCW '17.

[5] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang and Huan Liu. "Fake news Detection on Social Media: A Data Mining Perspective" ACM SIGKDD Explorations Newsletter Volume 19 Issue 1, June 2017 Pages 22-36

[6] Kai Shu, H. Russell Bernard, Huan Liu. "Chapter 3 Studying Fake News via Network Analysis: Detection and Mitigation" , Springer Nature America, Inc, 2019