

Detecting Fake News in Social Media Networks : A Review Paper

Akanksha Prakash Rathod

Department of Information Technology

Sir Sitaram and Lady Shantabai Patkar College of Arts and Science, Mumbai, Maharashtra, India

Abstract: *As the use of the internet increases day by day, fake news and hoaxes are increasing too. In order to attract more viewers the articles or news are fabricated to deceive the readers. In general, the goal behind the spreading of fake news is profiting via click baits. Click baits basically attract users with their flashy and luring headlines or design links such that revenue is generated on every click. Curious users often become a target for such advertisements. The purpose of this research paper is to run some analysis and come up with a solution wherein users can detect if the news is true and genuine or contains some false or misleading information.*

Keywords: Fake News, Clickbait, Social Media.

I. INTRODUCTION

It has become really easy for people to post whatever they desire. Posting something online which is completely fake may lead to a panicking situation for the readers, manipulating the news or promising fake gifts or vouchers by clicking certain links, or anything else that may lead to long-lasting repercussions. With the internet resources increasing every day, it becomes very difficult and almost impossible to differentiate between genuine information and misleading one. This leads to the problem of fake news among everybody who uses the internet. Another way of fooling users are clickbaits[2]. They are phrases or designs constructed specifically to attract the attention of curious users. When a user clicks on the link, they are redirected to a web page that is not the intended page. More are the clicks, more is the revenue generated for the publishers.

Major tech companies like Google, Facebook, and Twitter have attempted to solve this issue, however, their contribution had very fewer effects on solving this problem. Due to this, users continue to deal with this problem where they come across such misleading information via click baits and this affects their ability to get in touch with the actual news [8]. The reason for the engagement of major tech companies like Facebook and Twitter in this problem is that all the social media platforms contain a share option which can make fake news viral in minutes across a large number of people.

II. IMPACT OF FAKE NEWS

Social media is helpful in tremendous ways, however, when we compare the quality of news provided by social media with the traditional news organizations, it is quite low. The main reason is that it is cheaper. The spread quickly among groups of people. This news is generally fake and is spread intentionally with false information in order to gain financial and political attention.

Fake news can spread across large volumes and can have a serious negative impact on the people:

1. It is can break the authenticity of the news ecosystem.
2. It forces peers to accept false and biased beliefs.
3. Fake news is mostly manipulated by political parties or influencers to get the attention of a large group of people.
4. Fake news can impact the way a person would respond to real news. If a person heard fake news which triggered his distrust, it may confuse him when he would listen to real news. He will not be able to differentiate which is real news and which is not.

III. FAKE NEWS CHARACTERIZATION ON SOCIAL MEDIA

Below given are the two ways we can categorize the spread of fake news:

3.1 Malicious Accounts on Social Media for Propaganda

Social media was created for connecting people and sharing information around the globe. While there are genuine and legitimate accounts on social media, there are fake and malicious accounts too. Creating accounts on social media can be created for free or at a very low cost. This encourages people to create fake social media accounts, sometimes a social bot, a cyborg, or a trolling account [1]. A social bot or a cyborg is a social media account that is like a robot. It is not controlled by a human but by a computer program or an algorithm. It produces content automatically[1]. Also, it can interact with other users or bots online. These bot accounts are used specifically for sharing fake news online or for manipulating users with false information.

3.2 Echo Chamber Effect

The news that appears on consumer's news feed on social media accounts is selectively exposed to them because of the people they follow online, and the groups they are added into. For example, users on Facebook follow accounts of like-minded people, and hence, they receive news that favors them. Users on social media, therefore, form groups that contain like-minded people, which polarizes their thoughts and opinions resulting in an echo chamber effect. Due to this, there are more chances of people consuming fake news and believing them.

IV. NEWS CONTENT FEATURES

Features of news content can describe the information related to the news[5,8]

4.1 Linguistic-Based Feature

Research has utilized both common linguistic features and domain-specific linguistic features to capture differences between fake and real news. Documents are often represented by linguistic features for various tasks in natural language processing.

4.2 Visual-Based Feature

People tend to trust or believe something that they see. For any fake news propaganda, visual cues have proved to be the biggest manipulator. Fake news publishers often focus on images that can provoke anger or emotional response from the users which may exploit their vulnerabilities. Images or videos can be categorized as visual-based features used to capture different aspects of fake news.

4.3 User-Based Feature

Fake news is spread mostly by non-human accounts such as social bots or cyborgs as explained earlier. User profiles can be used to provide useful information in order to detect fake news.

4.4 Post-Based Feature

People tend to express their opinion or emotions via social media posts. Fake news may include skeptical opinions, sensational reactions which may trigger one's belief, etc can be expressed via these posts. Therefore, it is wise to extract post-based features to find out the reaction that the general public has expressed to the post.

4.5 Data Set for Fake News Detection

News that is online can be collected from different means such as social media, search engines, and news agency homepages. Manually determining whether the news is genuine or fake can become a tedious and challenging task, hence, we can refer to some publicly available datasets that can be very helpful[8]:

- **BuzzFeedNews:** A sample of News published on Facebook from 9 news agencies from 2016 is comprised in this dataset. All the posts were fact-checked by BuzzFeed journalists.
- **LIAR:** This dataset includes 12,836 short statements which are sampled from various contexts. These are the labels for news truthfulness.
- **BS Detector:** A browser extension called BS detector is used to collect this dataset.
- **CREDBANK:** It is a dataset for evaluating the social network events automatically on Twitter. It is used to detect potential rumors on Twitter.

Dataset	Features	News Content		Social Context		
		Linguistic	Visual	User	Post	Network
BuzzFeedNews		✓				
LIAR		✓				
BS Detector		✓				
CREDBANK		✓		✓	✓	✓

Figure 1

V. STOP WORDS, HYPERBOLIC AND COMMON PHRASES

Figure 2 shows the number of words used in both clickbait and non-clickbait headlines. According to the graph, we can indicate that the clickbait headlines are much longer than the non-clickbait headlines[2]

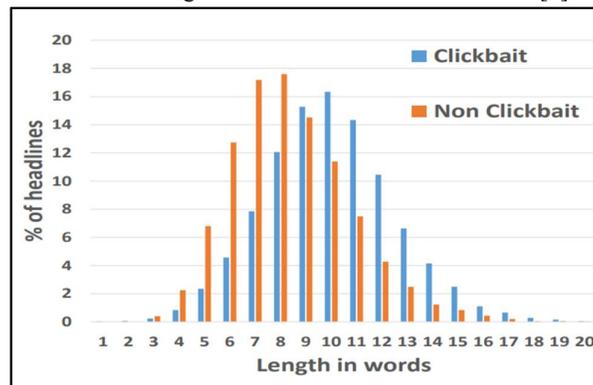


Figure 2

Figure 3 shows the percentage of clickbait and non-clickbait headlines which includes, word Contractions, Hyperbolic Words, and Determiners[2,3]

- **Stop Words:** In clickbait, stop words are used frequently.
- **Hyperbolic Words:** Awe-inspiring, breathtakingly, gut-wrenching, soul-stirring, etc. are some of the examples of hyperbolic words used in a clickbait headline.
- **Internet Slang:** WOW, LOL, LMAO, AMA are internet slang words that can be never found in a non-clickbait headline.
- **Length of Words:** Average word length of a non-clickbait headline is 6 and that of a clickbait headline is 4.5 characters.

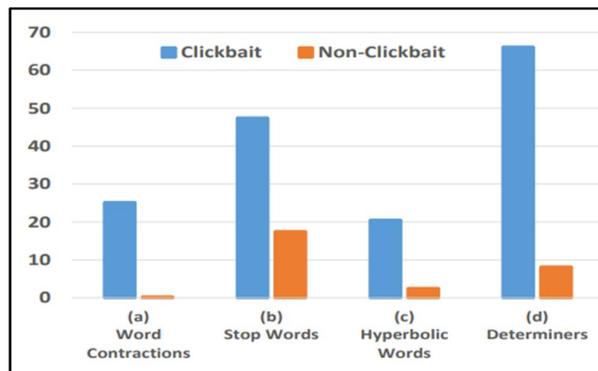


Figure 3

VI. PRECAUTIONARY METHODS TO AVOID FAKE NEWS AND CLICK BAITS

Browser extensions can be installed on the browsers which can help us to decipher between genuine news and fake news by analyzing the news using different datasets[11]

- Media Bias/Fact Check
- TrustServista
- TrustedNews
- Stop Clickbait

VII. CONCLUSION

Nowadays, people consume more news from social media than the traditional news media. Everything comes with its pros and cons, with its increasing popularity, social media is now playing a major role in the spread of the fake news across the globe. In this review paper, we explored the problem of fake news and discussed on how we can detect if a news is a genuine news or, a fake news or a clickbait.

In detection phase we used some data mining approaches to detect the fake news with the help of feature extraction. We also discussed different types of datasets to evaluate the news. Lastly, as a precautionary method, we discussed a few browser extensions that may help us to decipher between a real or fake news.

ACKNOWLEDGMENT

I would like to express my gratitude to my research supervisor Mr. Sujal Shah, Patkar Varde College for providing me with guidance throughout.

I would like to thank him for giving me the opportunity to do my research work on this topic. His positive attitude and helpful nature have always motivated and inspired me. I am always grateful to my parents for all their love and care towards me making me capable to pursue my higher education.

Finally, I owe deep gratitude to all the people who were directly or indirectly involved in my research work and supported me throughout.

REFERENCES

- [1]. Alessandro Bessi and Emilio Ferrara. Social bots distort the 2016 us presidential election online discussion. First Monday, 21(11), 2016.
- [2]. Prakhar Biyani, Kostas Tsioutsoulis, and John Blackmer. "8 amazing secrets for getting more clicks": Detecting clickbaits in news streams using article informality. In AAAI'16.
- [3]. Jonas Nygaard Blom and Kenneth Reinecke Hansen. Click bait: Forward-reference as lure in online news headlines. Journal of Pragmatics, 76:87–100, 2015.
- [4]. Paul R Brewer, Dannagal Goldthwaite Young, and Michelle Morreale. The impact of real news about news: Intertextual processes and political satire. International Journal of Public Opinion Research, 25(3):323–343, 2013.
- [5]. Carlos Castillo, Mohammed El-Haddad, Jürgen Pfeffer, and Matt Stempeck. Characterizing the life cycle of online news stories using social media reactions. In CSCW'14.
- [6]. Carlos Castillo, Marcelo Mendoza, and Barbara Poblete. Information credibility on twitter. In WWW'11.
- [7]. Abhijnan Chakraborty, Bhargavi Paranjape, Sourya Kakarla, and Niloy Ganguly. Stop clickbait: Detecting and preventing clickbaits in online news media. In ASONAM'16.
- [8]. Shu, Kai & Sliva, Amy & Wang, Suhang & Tang, Jiliang & Liu, Huan. (2017). Fake News Detection on Social Media: A Data Mining Perspective. ACM SIGKDD Explorations Newsletter. 19. 10.1145/3137597.3137600.
- [9]. 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.
- [10]. Conroy, N., Rubin, V., & Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. Proceedings of the Association for Information Science and Technology, 52(1), 1-4. Fan, C. (2017). Classifying fake news. Retrieved February 18, 2018, from <http://www.conniefan.com/2017/03/classifying-fake-news>
- [11]. <https://thetrustedweb.org/browser-extensions-to-detect-and-avoid-fake-news/>