Fraud news detection on social media by using online newspapers as a reference

A research proposal submitted by

Alam, Md. Iftakhar
Student Id: 17-35518-3
Section: D Group: G4
Email: iftekharifty722@gmail.com
Chowdhury, Ashfaq Afzal
Student Id: 17-35800-3
Section: D Group: G4
Email: aacfahim@gmail.com
Saha, Dip Kumar
Student Id: 17-35454-3
Section: D Group: G4
Email: dip9178@gmail.com

Course Supervisor:

Dr M.M. Mahbubul Syeed

Department of Computer Science

20/11/2020

Abstract

The research domain concerning fraud/fake news detection on social media is a huge field on its own but the scope of this research primarily focuses on the news that are misleading e.g. the mismatch between news title & the content in the body. News of this kind often creates riot among masses, defaming individuals or a particular establishment that sometimes claims the lives of many innocent people. Problems of this magnitude needs to be solved to ensure the tranquility of a sound society. The main objective of our research is to find the necessary solution to mitigate this particular issue. The study aims to find an insightful solution for this specific issue.

1 INTRODUCTION

Ever since the dawn of time mankind has tried to find a reliable yet easier way to communicate with the people around them. Through the ages news from different places relied upon the words of mouth to spread. But the invention of newspapers in the 17th century changed the way of how we consumed news around us [1]. The spread of the internet brought another revolution in the modern times. Ease of access & convenience has allowed us to share information within mere seconds. Because of the inherent nature of the internet a large number of people consume their news from the social media networks where news from different kinds of sources can be easily found in one place. A recent study from Pew Research claims that 62% of adults get their news from social media in the United States, with 29% among them doing so very often [2]. There are different types of fake news that propagates around social media. The most common of them are text, photo & video based fraud news. Our research focuses primarily on the propagation, detection & the mitigation of the fake news that are in text & photo format.

2 RESEARCH BACKGROUND

News plays a vital role in the way we perceive the world. Due to the limitation of mankind it is not possible for us to be present at more than one place at the same time. One of the most effective approaches to overcome this fact is to gather knowledge about the incidents that take place around the globe. A news portal can work as a time machine to delve into the past in a figurative sense. With the rise & spread of social media platforms like wildfire a sudden boom in the publications of fake news has taken place. A widely accepted definition of fake news is “fictitious articles deliberately fabricated to deceive readers” [3]. It is growing at an alarming rate & creating a severe problem to be dealt with due to the increasing number of

users [4]. Ease of access is solidifying the fact that a person can interact with the news that is published on the social media without almost any kind of non-existent restriction & the nature of the platforms makes them a breeding ground for the dissemination of fake news propagation [5]. However, detection of these kinds of news is not an easy task because of the dynamics of the news. They are not something that is created by a simple mistake or a sub conscious mind. They are made to fulfil a person's or an organization's motive [6]. One of the most common types of this kind of news is misleading titles or clickbait that are specifically designed to lure the users in to make profits off them. Using the user's curiosity & ingenuousness as a weapon the authors of the articles intentionally makes the titles or thumbnails look more alluring to make the users click which in turn increases the revenue through advertising to a wider demographic [7]. One of the main reasons why fake news is increasing rapidly can be traced back to one important thing. The authors themselves have said that creating news of this kind tends to entice more users to enter their website/platform that can significantly upturn their marketing reach & engage more consumers [8]. State of affairs such as these natures make it very difficult to detect fake news & a lot has to do with the fact that social media sites don't have a regulatory authority to dictate what is real and what is not [9]**.** This research offers to find a solution that can be implemented without much hassle & considerably decrease the rate of fake news on social media.

3 RESEARCH QUESTIONS

The main purpose of this research is to detect fraud news on social media using tools or algorithms which will be an effective solution to this problem. During the development of this project there will be questions about its impact on real life. Also there will be some questions for this project in future like maintainability, survivability, sustainability and quality issues.

We have to focus on some research question which is given below:

- What will be the effect of fake news on social media in social life?
- What aspects can be defined of this project with its long-term maintainability, survivability, sustainability and quality issue?
- What can be the most effective solution to detect the fake news?
- What kind of algorithm or tool will be sustainable for the evolution of this project?
- What correlations can be established between many social media like Facebook, Twitter, Instagram?
- How to analyze and design the solution to detect the fraud news?

4 RESEARCH METHOD

The dataset would find the Textual and Visual news content among the news sources and add it to the news detection dataset. As the system finds more news sources it would constantly add up the features onto the dataset table. It can also add dynamic information simultaneously. The news content would collect some textual and visual content to compare with the fake and legit news. Later it can be used to detect fake news.

But the problem is it is very difficult to differentiate the fake news and verified news during the early stage of publishing news. The system cannot identify the truthfulness of the news so that it can push the data inside the dataset. [10]

During the very initial stage of propagating news through any newspaper, it would distinct the newspaper sectors into two segments. Verified ones and the non-verified or less popular ones. [11]

Generally, the sources which are less familiar or unverified, seem to be more suspicious rather than the popular ones. All together the dataset will collect the news data from those sources who are less authentic and rate it by that order. Moreover, the dataset will also collect reliable news according to its rate and truthfulness. [12]

There will be another filtering feature that will come from user sentiment opinion that acts as a reliable ratio to identify the news authenticity.

In terms of sharing the news on social media websites, people who are sharing directly can be rated by the viewers as well. During propagating the news around social media sites like Facebook and Twitter people can show their own feature which can collect user opinion. Later on, all the data will be dynamically inserted into our dataset. The more it has data the more it can identify fake and reliable news. Our dataset can be also used as a checking database in order to check whether the news is completely fake or legit.

5 SIGNIFICANCE OF RESEARCH

Study of the research pattern to detect the fake news can be classified into three facts. There are predictions, tool or algorithm evolution and community evolution. The prediction can be defined with the simulation of the project. The tool or algorithm evolution explores the evolutionary behavior and design pattern of this evolution to evaluate the tool or algorithm. The study of community evaluation explores the social effect in the user communities and

also explores the interdependency between the tool or algorithm during the evolution of the project. Also there are some evolutionary parts which can be affected after the evolution of the project which clarified the maintainability, survivability, sustainability and quality issue.

Our main focus on this research to evolve a tool or algorithm. Also the community evolution needed. A project cannot survive in the community without the community's evolution. The evolution of the tool or algorithm can be defined for various kinds of social media like Facebook, Twitter, Instagram. This can be achieved by the prediction and also with the community evolution of the tool or algorithm evolution.

Our current research is going with the three kinds of evolution which have already been discussed above. We think all of them are needed for a long-term effective project in this community.

6 REFERENCES

1. **https://www.britannica.com/topic/publishing/The-first-newspapers**

2. News Use Across Social Media Platforms 2016 | Pew Research Center (journalism.org)

3. Monther Aldwairi, Ali Alwahedi. (2018). Detecting Fake News in Social Media Networks. Page 215. (http://www.sciencedirect.com/science/article/pii/S1877050918318210)

4. Shu, Kai & Mahudeswaran, Deepak & Wang, Suhang & Lee, Dongwon & Liu, Huan. (2018). FakeNewsNet: A Data Repository with News Content, Social Context and Dynamic Information for Studying Fake News on Social Media. Page 1. https://www.researchgate.net/publication/327464821

5. Yang, Shuo & Shu, Kai & Wang, Suhang & Gu, Renjie & Wu, Fan & Liu, Huan. (2019). Unsupervised Fake News Detection on Social Media: A

Generative Approach. Proceedings of the AAAI Conference on Artificial Intelligence. 33. 5644-5651. 10.1609/aaai.v33i01.33015644.

6. Liang Wu and Huan Liu. 2018. Tracing Fake-News Footprints: Characterizing Social Media Messages by How They Propagate. Page 637. DOI:https://doi.org/10.1145/3159652.3159677

7. Monther Aldwairi, Ali Alwahedi.(2018). Detecting Fake News in Social Media Networks, Page 215. (http://www.sciencedirect.com/science/article/pii/S1877050918318210)

8. Monther Aldwairi, Ali Alwahedi.(2018). Detecting Fake News in Social Media Networks, Pages 217. (http://www.sciencedirect.com/science/article/pii/S1877050918318210)

9. Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017. Fake News Detection on Social Media: A Data Mining Perspective. Pages 22-36. DOI:https://doi.org/10.1145/3137597.3137600

10. Shu, Kai & Mahudeswaran, Deepak & Wang, Suhang & Lee, Dongwon & Liu, Huan. (2018). FakeNewsNet: A Data Repository with News Content, Social Context and Dynamic Information for Studying Fake News on Social Media. Page 4-5, https://www.researchgate.net/publication/327464821

11. Wu, K.; Yang, S.; and Zhu, K. Q. 2015. False rumors detection on sina weibo by propagation structures, Page 5-7, DOI: 10.1109/ICDE.2015.7113322

12. Shu, Kai & Mahudeswaran, Deepak & Wang, Suhang & Lee, Dongwon & Liu, Huan. (2018). FakeNewsNet: A Data Repository with News Content, Social Context and Dynamic Information for Studying Fake News on Social Media. Page 5-8, https://www.researchgate.net/publication/327464821