

Counter Fake News Using AI

Pratima Chavan¹ and Prof. Divakar Jha²

Students, Department of MCA¹

Mentor, Department of MCA¹

Late Bhausaheb Hiray S S Trust's Hiray Institute of Computer Application, Mumbai, India

Abstract: *The COVID 19 pandemic is a humanitarian emergency that poses an enormous threat to society and has impacted various social media platforms and journalism. News and social media has become an immensely popular platform for consumption of information. The wide spread of fake news could inflict damages on social media platform. And hence, the need of Fake News Detection in the present scenario is inevitable. In this paper, we survey the recent literature about different approaches to detect fake news over the Internet. In particular, firstly I have discussed about fake news and the various terms related to it that have been considered in the literature. Further I have highlighted various publicly available datasets and various online tools that are available and can debunk Fake News in real time. Hence, I have described fake news detection methods based on two broader areas i.e., it's content and the social context. Finally, I have curated a comparison of various techniques that are used to detect fake news.*

Keywords: Fake news detection, Internet, content context, social context

I. INTRODUCTION

SOCIAL MEDIA for content and news consumption is a sword with two edges and can have serious implications and consequence if not used ethically. Therefore, to maintain social harmony it is highly crucial to detect fake news on these platforms and also regulate these to ensure that the users receive genuine information. Owing to its easily accessible nature, low cost and rapid relaying of information often lead people to pursue and consume news from social media. On the other hand, it permits the inescapable wide spread of fake news. Fake news may be spread intentionally or due to lack of awareness. The substantial spread of fake news has the capacity of having immensely negative influence on individuals and society. Fake news detection on social media has become a prevalent difficulty. Fake news is purposely written to mislead readers and to make them believe false information. Detection of origin of fake news is difficult since there is no one particular source. Anyone can be a peddler of false information on these gargantuan social media platforms. Misinformation is always a threat to humanity. Reception of fake news have caused several tragedies and have adversely affected people. A salient recent example of this is the COVID-19 pandemic, which leads to the circulation of false and unauthenticated news which is not only causing panic amongst citizens but also poses a threat to society. After reading fake news about Covid-19, many citizens have taken extreme actions which have affected their health too. The only way to eradicate the problem of fake news is to prevent oneself from consumption and further pass on of false information. I have discussed the definition of fake news and the various terms related to it.

An overview of existing fake news detection methods on the two broad categories i.e. 1) Content and 2) Social Context of the news.⁵⁹²

II. FAKE NEWS, RELATED TERMS AND VARIOUS TOOLS

Fake news is false or misleading information presented as news. Fake news often has the aim of damaging the reputation of a person or entity, or making money through advertising revenue.^{[1][2]} Although false news has always been spread throughout history, the term "fake news" was first used in the 1890s when sensational reports in newspapers were common.^{[3][4]} Nevertheless, the term does not have a fixed definition and has been applied broadly to any type of false information. It's also been used by high-profile people to apply to any news unfavourable to them. Further, disinformation involves spreading false information with harmful intent and is sometimes generated and propagated by hostile foreign actors, particularly during elections. In some definitions, fake news includes satirical articles misinterpreted as genuine, and articles that employ sensationalist or clickbait headlines that are not supported in the text.^[1] Because of this diversity of types of false news, researchers are beginning to favour information disorder as a more neutral and informative term.

Disinformation

Disinformation is fake or inaccurate information that is intentionally spread. The only difference between disinformation and misinformation is the intention of the spreaders. An example of disinformation is the fabricated false statements on Wikipedia made to masquerade as truth. However, in the context of social media, the intent is usually difficult to tell. Therefore, misinformation is usually used to denote fake and inaccurate information in social media, regardless of spreader intentions.

Rumor

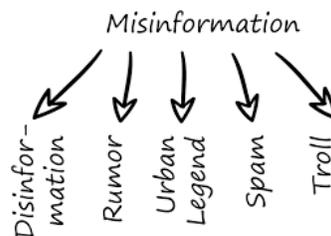
A rumor is a story circulating from person to person, of which the truth is unverified or doubtful. Rumors usually arise in the presence of ambiguous or threatening events. When its statement is proved to be false, a rumor is a type of misinformation. Therefore, among efforts in fighting against rumors, methods usually consist of two steps: rumor detection and truthfulness prediction.

Spam

Spam is unsolicited messages sent to a large number of recipients, containing irrelevant or inappropriate information, which is unwanted. Spamming messages are usually involved with spreading ads, malware, and even leading to scams. Spam is not necessarily misinformation, since some piece of information may be true and unsolicited simultaneously. However, spam that conveys misinformation can directly leads to information and financial loss.

Troll

A troll is a user who posts messages that are deliberately offensive or provocative, with the aim of upsetting other people. Trolling content is often happening in conversations, disrupting normal on-topic discussions and leading to emotional responses. The content trolls post is usually nonsense, and the troll spreaders post them for their own amusement.



III. FAKE NEWS DETECTION METHODS

The wide usage of social media platforms worldwide has provided a fertile ground for the widespread dissemination of online fake news in an unprecedented way. The social network is flooded with massive, diverse, and heterogeneous information (both real and fake), and spreads rapidly on these platforms causing severe impact to the whole society. Therefore, many researchers and technical giants are working together to detect fake news on online media. The traditional automatic rumour detection methods were based on hand crafted feature but with the advent of big data and a huge base of user generated data we have seen a shift to deep-level features. In this section, we discuss various state-of-the-art studies on fake news detection under the broader umbrella of content and social context of the news article

3.1. Content Based

The content-based methods [25–28] use various types of information from the news, such as article content, news. source, headline, image/video, to build fake news detection classifiers. Most content-based methods use stylometry features (e.g. sentence segmentation, tokenization, and POS tagging) and linguistic features (e.g. lexical features, bag-of words, frequency of words, case schemes) of the news articles to capture deceptive cues or writing styles

3.1.1. Knowledge Based

Knowledge-based approaches utilize fact checking method in which the given claim is compared with the external sources to verify the authenticity of the given claim. The existing fact checking methods can be categorized as manual and automatic fact checking.

In the state-of-the-art, the fake news detection methods are categorized into two types: (1) manual fact-checking; (2) automatic detection methods.

Expert Based

The expert-based methods use expert-oriented approach and rely on human experts working in specific domains for decision making. The manual fact-checking methods have some limitations: 1) it is time-consuming to detect and report every fake news produced on the internet; 2) it is challenging to scale well with the bulks of newly created news, especially on social media; 3) it is quite possible that the fact-checkers' biases (such as gender, race, prejudices) may affect the ground truth label.

The automatic detection methods are alternative to the manual fact-checking ones, which are widely used to detect the veracity of the news. In the previous research, the characteristics of fake news are usually extracted from the news-related features (e.g. news content) [21] or from the social contexts (social engagements of the users) [4, 22, 24] using automatic detection methods. Instead of relying on human intelligence these methods heavily rely on Natural Language Processing (NLP), Data Mining, Machine Learning (ML) techniques and network/graph theory [13]. The automatic fact-checking process can be divided into two stages: (1) fact extraction which is related to collection of facts and construction of a Knowledge Base and (2) fact-checking which is used to assess the authenticity of news articles by comparing that with the facts in the knowledge base.

3.1.2. Style Based

Style-based fake news detection follows the same approach like knowledge-based fake news detection of analysing the news content. However, instead of evaluating the authenticity of news content this method assesses the intention of writer to mislead the public [13].

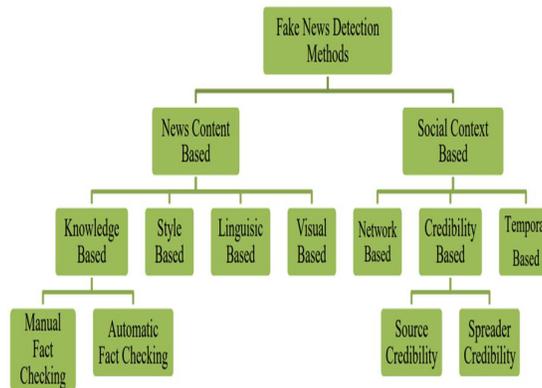


Figure 2. Fake News Detection Methods

3.1.3. Linguistic Based

Twenty-six linguistic based textual features were proposed in [19]. In [20] authors proposed an enhanced set of linguistic features to discriminate between fake and real news. [21] used network account features in addition to the linguistic features whereas [22] proposed Social Article Fusion (SAF) model that uses social engagements features along with linguistic

3.1.4. Visual Based

Visual content is often viewed as evidence that can increase the credibility of the news article [2] and hence the fake news publishers tend to utilize provocative visual content to attract and mislead readers. In [10] various visual and statistical image features are extracted for news authentication. Verifying Multimedia Use task [26] under the MediaEval-16 benchmark addresses the problem of detecting digitally manipulated (tampered) images.

3.2. Social Context Based

The existing social contexts-based approaches are categorized into two types:

- 1 Stance-based methods
- 2 Propagation-based methods.

The stance-based approaches exploit the users' viewpoints from social media posts to determine the truth. The users express the stances either explicitly or implicitly. The explicit stances are the direct expressions of users' opinions usually available from their reactions on social media. Previous works [4, 5, 22] mostly use upvotes/downvotes, thumbs up/down to extract explicit stances. The implicit stance-based methods [5, 31], on the other hand, are usually based on extracting linguistic features from social media posts.

The propagation-based methods [36–39] utilize information related to fake news, e.g. how users spread it. In general, the input to a propagation-based method can be either a news cascade (direct representation of news propagation) or self-defined graph (indirect representation capturing information on news propagation)

IV. REVIEW OF LITERATURE OF TECHNIQUES USED

Traditionally, the majority of approaches for detecting fake news focus on analysing the textual content only and utilized hand crafted textual features for the same. But, with an increasing number of articles which are attached with images over the Internet and the extensive use of social media networks, the multimodal features and social-context play a very vital role in better understanding the overall heuristics of the content. The traditional machine learning and rule-based algorithms are inefficient to detect the patterns in today's information age. Hence, to take advantage of big data Deep learning techniques are investigated for fake news detection. A general challenge of content-based methods is that fake news's style, platform, and topics keep changing. Models that are trained on one dataset may perform poorly on a new dataset with different content, style, or language. Furthermore, the target variables in fake news change over time, and some labels become obsolete, while others need to be re-labelled. Most content-based methods are not adaptable to these changes, which necessitates re-extracting news features and re-labelling data based on new features. These methods also require a large amount of training data to detect fake news. By the time these methods collect enough data, fake news has spread too far. Because the linguistic features used in content-based methods are mostly language-specific, their generality is also limited. To address the shortcomings of content-based methods, a significant body of research has begun to focus on social contexts to detect fake news. The social context-based detection methods examine users' social interactions and extract relevant features representing the users' posts (review/post, comments, replies) and network aspects (followers–following(relationships) from social media.

V. CONCLUSION

With an increase in the popularity and usage of social media over the past few years, a huge population of readers prefer to consume news from social media instead of traditional news media. Keeping this in mind, many publishers use social media and Internet in general as a breeding ground for spreading propaganda and rumours rapidly which has strong negative impacts on the society. In this text we have mentioned several freely available Fake News Detection tools that should be used so that we forward only credible and genuine news. In this paper, I have explored the present fake news detection methods by reviewing existing literature under two categories: The Content Based and The Social Context Based Fake news detection. In the content-based method, the article/post is considered that may contain the textual or visual content or both. In the social context-based method, the propagation structure and the credibility of the publisher is considered. While the content-based methods can be used for early detection of fake news the context-based methods fail to do so because of the absence of the propagation details in the very beginning of the proliferation of misinformation. Additionally, despite many researchers are focusing on this area but still there are only a few publicly available benchmark datasets.

REFERENCES

- [1]. FAKE NEWS DETECTION TOOLS AND METHODS – A REVIEW Sakshini Hangloo¹, Bhavna Arora² Ph.D Scholar¹, Assistant Professor² Email: sakshini.hangloo@gmail.com¹, bhavna.aroramakin@gmail.com² Department of Computer Science & Information Technology Central University of Jammu, Bagla (Rahya Suchani), District-Samba, Pin Code 181143, Jammu, J&K, India
- [2]. Fake News Detection: A Deep Learning Approach Aswini Thota Southern Methodist University, athota@smu.edu Priyanka Tilak Southern Methodist University, ptalak@smu.edu Simrat Ahluwalia Southern Methodist University, simeratjeeta@smu.edu Nibrat Lohia Southern Methodist University, nlohia@smu.edu
- [3]. Identifying Fake News on Social Networks Based on Natural Language Processing: Trends and Challenges Nicollas R. de Oliveira 1,†, Pedro S. Pisa 2,†, Martin Andreoni Lopez 3,†, Dianne Scherly V. de Medeiros 1,† and Diogo M. F. Mattos 1,
- [4]. Combating Misinformation Dissemination through Verification and Content Driven Recommendation 1Sarah Hawa, 2Lanita Lobo, 3Unnati Dogra, 4Prof Vijaya Kamble 1,2,3,4,5Sardar Patel Institute of Technology, Bhartiya Vidya Bhavan's Campus, Andheri, West Mumbai, India 1 sarah.hawa@spit.ac.in, 2 lanita.lobo@spit.ac.in, 3unnati.dogra@spit.ac.in, 4 vijaya.kamble@spit.ac.in
- [5]. A Deep Learning Framework for Detection of COVID-19 Fake News on Social Media Platforms Yahya Tashtoush 1,* , Balqis Alrababah 1 , Omar Darwish 2,* , Majdi Maabreh 3 and Nasser Alsaedi 4
- [6]. Fake news detection based on news content and social contexts: a transformer-based approach Shaina Raza1 · Chen Ding Sentiment Analysis for Fake News Detection Miguel A. Alonso * , David Vilares , Carlos Gómez-Rodríguez and Jesús Vilares
- [7]. Analyzing Machine Learning Enabled Fake News Detection Techniques for Diversified Datasets Shubha Mishra , 1 Piyush Shukla , 2 and Ratish Agarwal
- [8]. .Evaluating Deep Learning Approaches for Covid19 Fake News Detection Apurva Wani 1 ? , Isha Joshi 1 ? , Snehal Khandve 1 ? , Vedangi Wagh 1 ? , and Raviraj Joshi 2
- [9]. A Deep Learning Approach for Automatic Detection of Fake News Tanik Saikh, Arkadipta De, Asif Ekbal, Pushpak Bhattacharyya
- [10]. Al-Ahmad, B.; Al-Zoubi, A.M.; Abu Khurma, R.; Aljarah, I. An Evolutionary Fake News Detection Method for COVID-19 Pandemic Information. *Symmetry* 2021, 13, 1091. [CrossRef]
- [11]. COVID-19 Pandemic—Wikipedia. Available online: https://en.wikipedia.org/wiki/COVID-19_pandemic (accessed on 20 December 2021).
- [12]. Coronavirus: Hundreds Dead in Iran from Drinking Methanol Amid Fake Reports It Cures Disease. Available online: <https://www.independent.co.uk/news/world/middle-east/iran-coronavirus-methanol-drink-cure-deaths-fake-a9429956.html> (accessed on 15 April 2022).
- [13]. Kaliyar, R.K. Fake news detection using a deep neural network. In Proceedings of the 2018 4th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India, 14–15 December 2018; pp. 1–7.
- [14]. Gupta, A.; Sukumaran, R.; John, K.; Teki, S. Hostility detection and COVID-19 fake news detection in social media. *arXiv* 2021, arXiv:2101.05953.
- [15]. Kaliyar, R.K.; Goswami, A.; Narang, P. FakeBERT: Fake news detection in social media with a BERT-based deep learning approach. *Multimed. Tools Appl.* 2021, 80, 11765–11788. [CrossRef] [PubMed] 8. Elhadad, M.K.; Li, K.F.; Gebali, F. Detecting misleading information on COVID-19. *IEEE Access* 2020, 8, 165201–165215. [CrossRef]
- [16]. Raza, S. Automatic Fake News Detection in Political Platforms-A Transformer-based Approach. In Proceedings of the 4th Workshop on Challenges and Applications of Automated Extraction of Socio-Political Events from Text (CASE 2021), Online, 5–6 August 2021; pp. 68–78.
- [17]. Zhang, X.; Ghorbani, A.A. An overview of online fake news: Characterization, detection, and discussion. *Inf. Process. Manag.* 2020, 57, 102025. [CrossRef]
- [18]. Shu, K.; Sliva, A.; Wang, S.; Tang, J.; Liu, H. Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explor. Newsl.* 2017, 19, 22–36. [CrossRef]