Use of Implicit and Explicit Features in Fake News Detection

L.N.SAHU
Research Scholar
MATS University
Raipur(C.G.)
pro.lns711@gmail.com

Dr.Bhavna Narain
Professor(CS)
MATS University
Raipur(C.G.)
narainbhawna@gmail.com

Abstract: Because of the exponential growth of online information, it is becoming impossible to decipher the true from the false. Thus this causes to the problem of fake news. This research considers previous methods as well as current methods for fake news detection in textual formats while detailing why and how fake news exists in the first place. However On the other hand, social media provides an ideal place to the creation and spread of fake news. Fake news can become extremely influential and has the ability to spread exceedingly very fast. With the increment of people using social media, they are being exposed to new information and stories every day, automated classification of a text article as misinformation or disinformation is a challenging task. Even an expert in a particular domain has to explore multiple aspects before giving a verdict on the truthfulness of an article. Various machine learning and deep learning methods are available. This paper provides survey of various fake news detection methods.

Keywords: Fake news detection, deception detection, deep learning, GRU, RNN, LSTM

1 Introduction

The change to neoteric methods of news consumptions in recent era has brought various issues of fake news and misinformation to the forefront of discussion. Now a day's thousands of new news articles proliferate on social media networks every day, where each having neither a credibility nor a validation check, an ecosystem fuelled by misinformation (the inadvertent or careless sharing of false information) and disinformation(the deliberate or intestinally creation and sharing of information known to be false) has been established. Social media networks have enabled news articles to grow from traditional text-only news to news with images and videos which can provide a better storytelling experience and has the ability to engage and attract more readers. Recent fake news articles gets benefits from this very change to visual context aided news. Fake news article can now contain misrepresented, irrelevant and forged images for misleading readers. Peer-to-peer networks (primarily social media networks) assent fake news (propaganda) to be targeted at users who are more likely to accept and share a particular message. Fake news has been in the limelight in recent years for having an extensive negative effect on public events and activities. A major turning point was realized the 2016 U.S. presidential elections. It was believed that within the final three months leading up to the election, fake news favoring either of the two nominees was accepted and shared by more than 37 million times on Facebook [1]. This motivates towards the task of detecting fake news a crucial one.

Figure 1: shows three examples of fake news from the Twitter dataset [3].

Each tweet in figure 1 has certain textual content and an image associated with it. For the tweet on the left, both the image and text indicate that it is probably fake news. In the tweet on the right, the image does not add substantial information but the text indicates that it may be fake news. In the tweet in the middle, it is difficult to reach a conclusion from the text, but the morphed image suggests that it is possibly fake news. This example reflects the hypothesis that pairs of visual and textual information can give better insights into fake news detection. On a conceptual level, the task for detecting fake news has undergone a variety of labels from misinformation to rumor. The target of our study is to detect news content that is fabricated and can be verified to be false. Rumor and fake news detection techniques range from traditional learning methods to deep learning models. Initial approaches [2] tried to detect fake news using only linguistic features extracted from the text content of news stories. Ma et al. [3] explored the possibility of representing tweets with deep neural networks by capturing temporal-linguistic features followed by Chen et al. [4] who built upon this by introducing attention mechanism into the RNNs. Recent works [5] in the field of deep learning to detect fake news have shown performance improvements over traditional methods due to their enhanced ability to extract relevant features. Jin et al. [6] combine visual, textual and social context features, using an attention mechanism to make predictions about fake news. Wang et al. [5] use an additional event discriminator to learn common features shared among all the events with the goal of doing away with non-transferable event-specific features, and claim that they can handle novel and newly emerged events better. A shortcoming of the existing models is that they do not have any explicit objective function to discover correlations across the modalities. Kingma et al. [7] proposed that latent variable models act as a powerful approach to generatively model complicated distributions and brought forth the ideas of Variational Autoencoder (VAE). Variational Autoencoder can learn probabilistic latent variable models by optimizing a bound on the marginal likelihood of the observed data.

We may summarize, the contributions of our work are as follows:

- We propose an approach for classifying social media posts using only the content of the post, i.e., the text and the attached image.

- The proposed model uses a text feature and image feature separately to detect if a post is fake or not.

1.1 Fake news in social media

Nowadays online fake news tends to be intrusive and diverse in terms of topics, styles and platforms [8]. And it is not easy to define a generally accepted definition for "fake news". Stanford University defines fake news as: "the news articles that are intentionally and verifiably false, and can mislead readers (Detecting fake news with nlp)". According to Wikipedia (Fake news), fake news is: "a type of journalism which contains deliberative misinformation spread via traditional print and broadcast news media or online social media." As the substantial development of the Internet, more fake news is distributed via social networks. In this paper, we proposed our definition as: "fake news are all types of false stories or news post which are created and mainly published and distributed on the Internet, in order to intentionally mislead, befool or lure readers for financial, political or other gains."

Figure2: Fake news and everything related to it. [9]

To clearly understand the scope and variety of online fake information, some important aspects for defining fake news are shown in Figure 2. In Figure 1.2, the term "Fake News" is in the core of the onion-shaped graph, and it contains four major components: Creator/ Spreader, Target Victims, News Content, and Social Context.

Figure 3: Example of Fake news Shared by a facebook user. [9]

All the components are in the first inner layer around "Fake News".

- **Creator or Spreader:** The creators of online fake news can be either real human or non-human. A real human fake news creators includes both benign authors and users who publish fake news unintentionally, and malicious users who create false information intentionally.

- **Target Victims:** The Victims are the main targets of the online fake news. They may be users from online social media or other online news platforms. Based on the purposes of the news, the targets may be students, voters, parents, senior people, and so on.

- **News Content:** News content refers to the body of the news post. It contains physical content (e.g., title, body text, and multimedia) and non-physical content (e.g., purpose, topics, sentiment).

- **Social Context:** Social context indicates how the news is distributed via the Internet. Social context analysis comprises user network analysis (how online users are involved in the news) and broadcast pattern analysis (temporal pattern of the dissemination).

Figure 3 illustrates an example of fake news shared by a Facebook user. In this Figure, A, B, C, and D represents creator or spreader, news content, social context, and target respectively. As we know that, www.dailypresser.com is the sources of the news, and the Facebook user Bob is obe of the news spreader. Both of them can be considered as A. B comprises the title of the news, the body of the news, the multimedia of the news if any available, and the comment from Bob. C includes all the interactions between other users and this news post (e.g., comments, likes or dislikes, the timestamp, and so on). And any user who involves with this news post by the above mentioned ways can be considered as potential targets, which is D.

1.2 Fake news creators and target users

It is significant to demonstrate who is behind the fake news, and why the fake news is written and shared via the social media. The creator or spreader of the fake news can be either real human beings or non-humans.

• **Non-humans:** Social cyborgs and bots are the most common non-human fake news creators. Social bots are computer algorithms that are designed to exhibit human-like behaviors, and automatically produce content and interact with humans on social media [10]. Although some social bots perform important roles in the spread of legitimate information [11], many bots are designed particularly to distribute rumors, spam, malware, misinformation, slander, or even just noise. For example, millions of social bots are created for supporting either Trump or Clinton in 2016 U.S. election, injecting thousands of tweets pointing to websites with kinds of fake news. Cyborgs refer to either bot-assisted humans or human-assisted bots [12]. After being registered by a human, the cyborg account can be able to post tweets and participate with the social community. Similar to social bots, malicious cyborg accounts can mislead and exploit online social users by disseminating fake information and messages that can result in damaging the social belief and trust. With the nature of bot, cyborg becomes an important and essential platform to spread fake news very fast and easily. In this paper, the cyborg is contemplated as one type of non-human fake news creators.

• **Real humans:** Real humans operating social media are crucial sources for fake news diffusion. Actually, social bots and cyborgs are only the carriers of fake news on social media that automate accounts are being programmed for spreading false messages by humans. There is no matter that, the fake news is spread manually or automatically, real humans, who target to disturb the credibility of online social community, are the ultimate creators for the untreatable information. The fake contents are generated intentionally by the malicious online users, so it is really difficult to distinguish between fake information and truth information only by content and linguistic analysis [8]. Furthermore, some beginner users can also contribute for the distribution of fake news. For example, the following news: “FBI agent found dead in apparent murder-suicide, suspected in Hillary E-mail leaks” is completely false, but this news is shared on Facebook over half a million times (Denver guardian). It is obvious that many valid users become the spreaders of the news. This news can be posted and shared in the certain community groups, where the friends and followers of the legitimate users can behave as the next-generation to spread as well. Therefore, an echo chamber comes in existence which makes the propagation of the false news widespread. On the social media, online users do not need to take responsibilities for what they post, share and comment. This became problematic since the unidentified messages may undergo far-reaching dissemination, and may have material impacts on the Internet.

The rest of the paper is organised as follow: Section II will contain some of the review on the fake news detection using multi-modal features.

2 Review of Literature

Fake news detection is not simple but rather is a complicated problem [13]. The detection task requires several steps to classify a given set of news articles (text documents) [14]. The preprocessing of the collected news documents vary depending on the type of data and the language being used in the documents. Figure 4 shows the typical types of data in news documents. Generally, most of news articles contain textual data. Therefore, the documents should be transformed to another representation, in order to be able to extract feature vectors that contain enough information to ensure accurate classification, and appropriate to be maintained by machines.

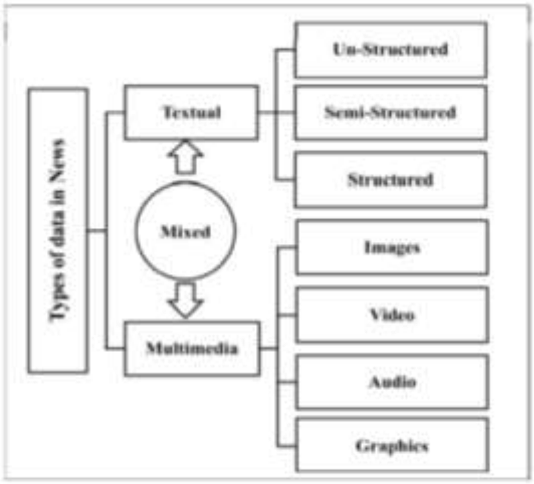

Figure 4: Types of data in News [15][16].

Kai Shu et. al. (2017) [15], and Ugur Kursuncu et. al. (2019) [16], all gave a categorization for the types of extracted features according to their perspective. From our point of view, the extracted features could be categorized into content-based features, context based features, and domain-based features, depending on the application and the type of data in the news document. The content dependent features are Meta information that represents the raw content of a news document including the author, editor, publisher, news title, the body of the news itself, and any attached multimedia. So, from the content that is available in our data, we can extract different types of features. We can extract Syntactic and Lexical features, such as number of URLS, word length, hashtag, retweet, and term frequency, from the title and the body of the news document. In a news post from any attached multimedia, we can extract visual and statistical features with loot of information such as clarity score, coherence score, similarity distribution histogram, image ratio, multi-image ratio, etc. The author/editor/publisher of the news document besides being a part of the news content, they are also highly related to the news context. A news context is the social engagements of the news article consumption on social media platform. This social engagement represents the news propagation over time and the group of users engaged with this news document. Hence, we can extract from individuals, group, and postings, social based features such as number of friends count, followers, registration age, number of authored posts/tweets, related social groups, demographic information, user stance, average credibility scores, etc. Moreover, we could extract features about the domain that the news article belongs to, by extracting propagation features that consider characteristics related to the propagation tree, which can be constructed from the retweets of a message in a certain domain. These comprise the depth of the re-tweet tree and the number of initial tweets of a topic. Some studies on text classification in general, used the weighted feature vectors to improve classification results. The weight of each feature pinpoints the importance of feature thus enhancing the classification results.

2.1 Multi-modal Fake News Detection

In social media, with the textual features, news often comprises various kinds of data, which provides more comprehensive features for detecting fake news. Thus, investigating multi-modal data for fake news detection is attracting increasing attention [17]-[19]. A survey on different content kinds of news and their impacts on readers can be seen in [19]. In general, multi-modal based fake news detection specialize in extracting features from news content, including news publisher, textual contents and image/video. By using these three features mentioned, different varieties of news representations are often built, which capture discriminative aspects of news stories. The deep networks are used to capture both visual and textual information of news in multimedia based methods. Researchers usually apply classification models to distinguish fake news from the real ones. In literature [20], the authors propose an attention based Recurrent Neural Network to fuse the multi-modal data from tweets for rumor detection. An attention mechanism is added to find the correlations between images and texts of a news post. The architecture of Event Adversarial Neural Networks (EANN) is proposed in literature [21]. Both text and image in a news article are taken into consideration. The authors form and train an event discriminator mechanism in order to remove the effects of the event-specific features and retain the common features among all these studied events. Despite the success of multi-modal based fake news detection methods, only few of them explicitly model user sentiments towards news for fake news detection; while sentiments are very strong signal that have great potential to improve fake news detection performance. Therefore, in this work, they investigate a novel problem of exploring user sentiments for fake news detection with multi-modal data. Fake-news images are more visually striking and emotional provocative than real-news images. In the fake news detection task, there exist two important requirements that arise from the unique characteristics of news data.

Table below shows some existing work done in the field of Fake news detection using image and text features.

Table 1: Comparison of Some Existing Methods.

Ref.	Year	Methods Used	Key Contribution	Feature
[22]	2017	Factors affecting fake news creation and spreading.	Highlighted the different characteristics of fake news detection.	Only Text
[23]	2017	A new n-gram model is proposed by incorporating two feature extraction techniques and result of six machine learned classifier is compared.	The increase of n-gram size resulted in decrease of accuracy. High accuracy is obtained using 5000 and 10000 features.	Only Text
[24]	2017	Classification models based on linguistic differences are developed. Also features representing text readability properties were proposed.	The best performing models achieved accuracy which is comparable to human ability to detect fake news.	Only Text
[25]	2017	Statistical and visual features were proposed for images.	Work done on images can be combined with other relevant work for better results.	Only Images
[26]	2018	Fake news game has been developed which helped in reducing the persuasiveness of such articles	First to develop a multi-player fake news game that can manage the impact that fake news can have on society	Only Text
[27]	2018	Analysis of implicit and explicit profile features.	Correlation between fake/real news and user profiles has been highlighted.	Only Text
[28]	2018	Collections of news articles as multi-dimensional tensors, leverage	A semi-supervised content-based method for detecting fake	Text and Image

		tensor decomposition to derive concise article embeddings that capture spatial/contextual information about each news article.	news articles, leveraging tensor-based article embeddings and guilt-by-association.	
[29]	2018	objective is to build a classifier that can predict whether a piece of news is fake or not based only its content, thereby approaching the problem from a purely deep learning perspective by RNN technique models (vanilla, GRU) and LSTMs.	Implementation of RNN technique models (Vanilla, GRU) and LSTMs that have been proposed for the detection of online fake news.	Only Text Articles
[30]	2018	This paper faces the specific problem of recognizing images in online news which have been modified or mis-contextualized, i.e. images taken in a different place and/or time with respect to the event to which they are associated.	To identify image tampering a number of image forensic techniques were exploited and combined. Also textual analysis approach is proposed based on the extraction of features from the news the image is associated with, and from textual information retrieved online using the image at stake as pivot.	Only Image Articles.
[31]	2019	Author had proposed a novel framework Multi-domain Visual Neural Network (MVNN) to fuse the visual information of frequency and pixel domains for detecting fake news.	Design a CNN-based network to automatically capture the complex patterns of fake-news images in the frequency domain.	Only Image Articles.

3Conclusion

Fake news is defined as news articles that are intentionally and verifiably false and could mislead readers, which has been widely adopted in recent studies. Different from the traditional definition, in the context of micro blogs, fake news aims at news posts that are published on social media by users which are usually less than 140 characters instead of news articles. Real images are consistent with the time and location of the event and are original and not doctored. Fake images are either tampered or mis-contextualized, where mis-contextualization is defined as a temporal and/or geographical misplacement of the image with respect to the temporal and geographical event described in the news to which the image is associated. This paper has surveyed many existing solution and found various issues and challenges in them. Also found that multimodal solution provides good accuracy.

References

[1] Hunt Allcott and Matthew Gentzkow. 2017. Social Media and Fake News in the 2016 Election. 31 (05 2017), 211–236.

[2] Carlos Castillo, Marcelo Mendoza, Barbara Poblete, Information credibility on twitter. In Proceedings of the 20th international conference on World Wide Web. 2011. ACM, 675–684.

[3] Jing Ma, Wei Gao, PrasenjitMitra, Sejeong Kwon, Bernard J Jansen, Kam-Fai Wong, and Meeyoung Cha. 2016, “Detecting Rumors from Microblogs with Recurrent Neural Networks”, In IJCAI. 3818–3824.

[4] Tong Chen, Lin Wu, Xue Li, Jun Zhang, Hongzhi Yin, and Yang Wang. 2017. Call Attention to Rumors: Deep Attention Based Recurrent Neural Networks for Early Rumor Detection. ArXiv preprint arXiv: 1704.05973 (2017).

[5] Yaqing Wang, Fenglong Ma, Zhiwei Jin, Ye Yuan, GuangxuXun, KishlayJha, Lu Su, and Jing Gao. 2018. EANN: Event Adversarial Neural Networks for Multimodal Fake News Detection. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. ACM, 849–857.

[6] Zhiwei Jin, Juan Cao, Yongdong Zhang, Jianshe Zhou, and Qi Tian. 2017. Novel visual and statistical image features for

microblogs news verification. IEEE transactions on multimedia 19, 3 (2017), 598–608.
[7] Diederik P Kingma and Max Welling. 2013. Auto-encoding variationalbayes. ArXiv preprint arXiv: 1312.6114 (2013).
[8] Shu, K., Wang, S., Tang, J., Zafarani, R., & Liu, H. (2017). User identity linkage across online social networks: A review. ACM SIGKDD Explorations Newsletter, 18(2), 5–17.
[9] Xichen Zhang and Ali A. Ghorbani," An overview of online fake news: Characterization, detection, and discussion" Information Processing and Management.
[10] Ferrara E., Varol, O., Davis, C., Menczer, F., &Flammini, A. (2016). The rise of social bots. Communications of the ACM, 59(7), 96–104.
[11] Ratkiewicz, J., Conover, M., Meiss, M. R., Goncalves, B., Flammini, A., &Menczer, F. (2011). Detecting and tracking political abuse in social media. ICWSM, 11,297–304.
[12] Chu, Z., Gianvecchio, S., Wang, H., &Jajodia, S. (2012). Detecting automation of twitter accounts: Are you a human, bot, or cyborg? IEEE Transactions on Dependable and Secure Computing, 9(6), 811–824.
[13] Zhiwei Jin, Juan Cao, Yu-Gang Jiang, and Yongdong Zhang. 2014. News credibility evaluation on micro-blog with a hierarchical propagation model. In Data Mining (ICDM), 2014 IEEE International Conference on. IEEE, 230–239.
[14] HuaShen, Fenglong Ma, Xianchao Zhang, LinlinZong, Xinyue Liu, and Wenxin Liang. 2017. Discovering social spammers from multiple views. Neurocomputing 225 (2017), 49–57.
[15] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," ACM SIGKDD Explorations Newsletter, vol. 19, no. 1, pp. 22–36, 2017.
[16] U. Kursuncu, M. Gaur, U. Lokala, K. Thirunarayan, A. Sheth, and I. B. Arpinar, "Predictive analysis on twitter: Techniques and applications," in Emerging ResearchChallenges and Opportunities in Computational SNA andMining. Springer, 2019, pp. 67–104.
[17] F. Qian, C. Gong, K. Sharma, and Y. Liu, "Neural user response generator: Fake news detection with collective user intelligence." In IJCAI, 2018, pp. 3834–3840.
[18] S. Tschiatschek, A. Singla, M. Gomez Rodriguez, A. Merchant, and A. Krause, "Fake news detection in social networks via crowd signals," in Companion of the The Web Conference 2018 on The Web Conference 2018. International World Wide Web Conferences Steering Committee, 2018, pp. 517–524.
[19] L. Wu and H. Liu, "Tracing fake-news footprints: Characterizing social media messages by how they propagate," in WSDM. ACM, 2018, pp. 637–645.
[20] Stanislaw Antol, AishwaryaAgrawal, Jiasen Lu, Margaret Mitchell, DhruvBatra, C Lawrence Zitnick, and Devi Parikh. 2015. Vqa: Visual question answering. In Proceedings of the IEEE International Conference on Computer Vision. 2425–2433.
[21] Andrej Karpathy and Li Fei-Fei. 2015. Deep visual-semantic alignments for generating image descriptions, In Proceedings of the IEEE conference on computer vision and pattern recognition. 3128–3137.
[22] A. Campan, A. Cuzzocrea and T. M. Truta, "Fighting fake news spread in onlinesocial networks: Actual trends and future research directions," in IEEE International Conference on Big Data (Big Data), Boston, MA, 2017.
[23] H. Ahmed, I. Traore and S. Saad, "Detecting opinion spams and fake news using textclassification," Security and Privacy, vol. 1, no. 1, 2017.
[24] V. P´erez-Rosas, B. Kleinberg, A. Lefevre and R. Mihalcea, "Automatic Detection of Fake News," 2017.
[25] Z. Jin, J. Cao, Y. Zhang, J. Zhou and Q. Tian, "Novel visual and statistical image features for microblogs news verification," IEEE, Transactions on Multimedia, vol. 19, no. 3, pp. 598-608, 2017.
[26] J. Roozenbeek and S. van der Linden, "The fake news game: activelyinoculating against the risk of misinformation," Journal of Risk, 2018.
[27] S. Kai, S. Wang and H. Liu, "Understanding User Profiles on Social Media for Fake News Detection," in IEEE Conference on MultimediaInformation Processing and Retrieval (MIPR), Miami, FL, 2018.
[28] GiselBastidasGuacho, Sara Abdali, Neil Shah, Evangelos E. Papalexakis," Semi supervised Content-based Detection of Misinformation via Tensor Embeddings", IEEE/ACM ASONAM 2018.
[29] Sherry Girgis, EslamAmer, Mahmoud Gadallah, " Deep Learning Algorithm for Detecting Fake News in Online Text", IEEE-2018.
[30] Federica Lago, Quoc-Tin Phan, Giulia Boato," Image Forensics in Online News", IEEE-2018.
[31] Peng Qi1,2 , Juan Cao1,2 , Tianyun Yang1,2 , Junbo Guo1 and Jintao Li," Exploiting Multi-domain Visual Information for Fake News Detection", 2019 IEEE International Conference on Data Mining (ICDM).