Fake News Characterization Using Linguistic Features

XU Wentao[1] and SASAHARA Kazutoshi[1, 2, 3]

1. Graduate School of Informatics, Nagoya University
2. Center for Computational Social Science, Kobe University
3. JST PRESTO

Introduction

People began to understand the role of fake news in social network services during the 2016 U.S. presidential election, during which the most-frequently discussed fake news on Facebook helped Donald Trump win over Hillary Clinton. Some people claimed that Donald Trump would not have won the campaign without the fake news about his competitors [2] [3]. Fake news is an emerging serious concern in our society. The study of fake news detection is thus an important research topic. The definition of fake news varies by the viewer's perspective. Fake news generally consists of rumors and misinformation. Fake news occurs across history, medicine, science, politics, and real-time social events. Journalists from the European Association for Viewers Interests even classify fake news into 10 categories [1]. The objective of this research is to examine how much linguistic features are effective to characterize and distinguish fake news from mainstream news.

Data and Methods

To validate whether there is language pattern significance in fake news, we compared news from trustworthy mainstream news sources and online "fake news" and "true news" labeled by the recognized fact news-checking websites PolitiFact, BuzzFeed, and Snopes. To this end, we obtained the labeled online news from two well-recognized public benchmark datasets of fake news detection sources, FakeNewsNet [12] and MisInfoText [13]. FakeNewsNet consists of 240 stories from PolitiFact and 90 from Buzzfeed, as shown in [Table 1]. PolitiFact, founded in 2007, is a not-for-profit national news organization [14]. PolitiFact identifies online political fake and true news, in addition to other types of misinformation, including social history. Buzzfeed is a for-profit international media group [15]; their website includes a fact-checking column where the public is able to identify the truth of a piece of single news. In addition, we collected 16416 pieces of BBC news from [11].

Table 1 Statistics of FakeNewsNet, MisInfoText, and "Professional" Fake News.

	PolitiFact	BuzzFeed	MisInfoText PolitiFact	"Professional" Fake News	BBC News
True	**120**	**90**	**2922**	**0**	**16416**
Fake	**120**	**90**	**4736**	**700**	**0**
Total	**240**	**180**	**7658**	**700**	**16416**

MisInfoText is composed of fake news from Snopes, PolitiFact, and Emergent. Snopes, founded in 1994, is a misinformation-checking company [16]. Although political fact-checking makes up a

large portion of their website, checking urban legends, hoaxes, and folklore is also part of their website. Emergent is a real-time rumor tracker that is part of a research project with the Tow Center for Digital Journalism at Columbia University [17]. MisInfoText is under development [13]; therefore, the version varies. In this research, we only employed PolitiFact data from MisInfoText. Our MisInfoText version is described in [Table 1]. We then mixed the two datasets to get 4946 fake news samples and 3132 true news samples. Finally, we collected 700 news samples by web crawling recognized "professional" writer' s news websites [18].

Previous research showed that fake news from various sources shares common language patterns [4] [5] [6] [7]. Here we used 107 out of 116 language features from [8] (Model A); the nine features we excluded are "#%paragraph", "#%unique words", "Average # Sentences Per Paragraph", "# Common Clickbait Phrases", "# Common Clickbait Expressions Overall", "# Common Clickbait Patterns", and "#Long Words"(#: count, %: percentage). This research focused on the language features of fake news content instead of fake news titles; therefore, we excluded clickbait-related and paragraph number-related features. "Unique Word" is used to characterize an article's word diversity and we assume that this has little contribution to the news truth description. "Long word" is quite a vague concept and there is not a recognized standard of a "long word;" therefore, we excluded them from this study.

Furthermore, we used 140 language features from [9] (Model B). In addition, we used embedding representations by word2vec (w2c), fastText [10], and Doc2Vec to represent news as vectors. We trained these models and compared the results with the language pattern models. The experimental data were prepared as follows. We sampled both BBC and online news. The online news included fake and true news in terms of content. The volume of each sampling dataset ranged from 800 to 2400 with steps of 800. The volume of fake news to true news was 1:1. The proportion of the sub-dataset of training to testing samples was 4:1. XGboost and Random Forest (RF) were used as classifiers in each experiment. We report the 20-iterations average accuracy and average f-1 scores of fake news identification by the fine-tuned RF and XGboost classifiers.

Results

We found that online true news and fake news share much in common with each other in terms of language features and that there is a significant difference between BBC news and online news (both true and fake) in terms of linguistic features.

More specifically, we examined whether the reported words, biased words, and factive words mentioned in [8] contribute to the classification of BBC news and online news using the linguistic features mentioned above. Table [3]-[5] show the average classification accuracy with regard to the online news (or fake news in Table [3]) f1-score of each model. Table 3 shows the averaged results of online fake news and true news performance. We first trained models using the online fake news and online true news data and found that the models' accuracies were just above 50% (ranging from 51% to 55%), which was slightly larger than that of higher-representation models (which range from 0.48 to 0.54). Table 4 shows the averaged results of BBC news and online fake news performance. The accuracy of the language models rose dramatically, up to a range from 89% to 99%, and the accuracy of the embedding representations model increased to 82% to 88%, although the Doc2Vec model fluctuated around 50%. We asked whether this performance difference existed exclusively between BBC news and online fake news. Table 5 shows the averaged results of BBC news and online true news performance. We found that the higher

accuracy remained. The accuracy ranged from 91% to 98% by the language pattern models and the embedding representations scores ranged from 83% to 90%, except for the Doc2Vec model. Finally, we asked whether it is possible to distinguish between BBC news and professional fake news. Table 6 shows the averaged results of BBC news and the "professional" online fake news performance. We found that the accuracy of Model B with readability and quality reached its highest value at 98.7%.

We were not able to identify any accuracy improvement with the help of report verbs, factive words, or biased words across the three experiments. In addition, the quality module proposed by [9] contributed little to the improvement; however, the readability module contributed to the accuracy improvement in both of the BBC/online news experiments.

Table3 Online Fake News and Online True News Classifcation

classifier	XGBoost						Random Forest					
samlpe	800		1600		2400		800		1600		2400	
score	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score
model A	50.7	50.7	53.0	52.2	52.7	52.7	50.7	50.7	53.0	52.2	52.7	52.7
model A+report verbs	50.6	50.6	52.8	52.0	53.1	52.9	50.6	50.6	52.8	52.0	53.1	52.9
model A+factives	50.9	50.8	53.4	52.5	53.5	53.1	50.9	50.8	53.4	52.5	53.5	53.1
model A+biased words	50.7	50.9	53.0	52.1	53.0	53.1	50.7	50.9	53.0	52.1	53.0	53.1
model A+report verbs+factives	50.9	51.0	53.2	52.0	53.2	53.0	50.9	51.0	53.2	52.0	53.2	53.0
model A+report verbs+biased words	51.1	51.5	52.8	52.0	52.9	52.8	51.1	51.5	52.8	52.0	52.9	52.8
model A + factives + biased words	51.3	51.0	53.6	52.6	53.1	53.1	51.3	51.0	53.6	52.6	53.1	53.1
model A + factives + biased words+report verbs	51.0	50.9	52.7	52.0	52.6	52.4	51.0	50.9	52.7	52.0	52.6	52.4
model B	53.6	53.4	55.0	54.9	54.9	54.4	53.6	53.4	55.0	54.9	54.9	54.4
model B+readability	53.4	53.2	54.4	54.2	54.8	54.4	53.4	53.2	54.4	54.2	54.8	54.4
model B+quality	53.1	53.3	55.4	55.3	54.5	54.0	53.1	53.3	55.4	55.3	54.5	54.0
model B+readability&quality	53.5	53.9	55.3	55.1	54.8	54.2	53.5	53.9	55.3	55.1	54.8	54.2
w2c	51.4	51.6	51.7	51.4	53.6	53.3	51.4	51.6	51.7	51.4	53.6	53.3
d2v	49.0	48.0	49.3	49.0	49.7	49.2	49.0	48.0	49.3	49.0	49.7	49.2
fastText	50.0	49.8	51.6	51.2	52.0	51.8	50.0	49.8	51.6	51.2	52.0	51.8

Table 4 BBC News and Online Fake News Classifcation

classifier	XGBoost						Random Forest					
samlpe	800		1600		2000	2000	800		1600		2000	
score	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score
model A	97.3	97.3	97.3	97.1	97.3	97.2	97.2	97.2	96.9	96.7	96.9	96.7
model A+report verbs	97.4	97.3	97.4	97.2	97.3	97.2	97.2	97.1	96.9	96.7	96.8	96.7
model A+factives	97.4	97.3	97.3	97.1	97.2	97.1	97.2	97.2	96.9	96.7	96.8	96.7
model A+biased words	97.4	97.4	97.4	97.2	97.3	97.2	97.3	97.3	96.9	96.7	96.8	96.7
model A+report verbs+factives	97.3	97.3	97.4	97.2	97.3	97.2	97.2	97.2	96.9	96.6	96.9	96.8
model A+report verbs+biased words	97.4	97.3	97.4	97.2	97.3	97.2	97.3	97.3	96.9	96.6	96.9	96.7
model A + factives + biased words	97.4	97.3	97.4	97.2	97.3	97.2	97.3	97.2	96.9	96.7	96.8	96.7
model A + factives + biased words+report verbs	97.3	97.3	97.3	97.1	97.2	97.1	97.3	97.3	96.9	96.7	96.9	96.7
model B	89.7	89.4	91.7	90.9	92.4	92.0	89.5	89.3	91.1	90.2	91.3	90.8
model B+readability	98.1	98.1	98.2	98.0	98.6	98.5	98.0	98.0	97.4	97.3	97.6	97.5
model B+quality	89.9	89.7	92.1	91.5	92.5	92.1	89.8	89.6	91.2	90.4	91.6	91.1
model B+readability&quality	98.1	98.1	98.3	98.2	98.5	98.5	98.3	98.3	97.9	97.7	98.1	98.0
w2c	85.2	85.2	86.9	86.1	88.4	88.0	84.4	84.4	86.4	85.6	87.9	87.5
d2v	50.3	50.3	49.5	49.0	51.1	51.6	51.1	50.4	51.1	49.9	51.0	49.4
fastText	81.8	81.8	83.5	82.5	85.1	84.6	81.5	81.7	83.3	82.4	84.0	83.6

Table 5 BBC News and Online True News Classifcation

classifier	XGBoost						Random Forest					
samlpe	800		1600		2000	2000	800		1600		2000	
score	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score
model A	96.6	96.6	96.3	96.0	97.1	97.0	95.9	95.9	95.7	95.4	96.4	96.3
model A+report verbs	96.6	96.6	96.4	96.1	97.2	97.0	96.1	96.1	95.8	95.5	96.5	96.4
model A+factives	96.5	96.5	96.5	96.2	97.0	96.8	96.2	96.2	95.7	95.4	96.5	96.3
model A+biased words	96.6	96.5	96.3	96.0	97.1	97.0	96.1	96.1	95.7	95.4	96.5	96.4
model A+report verbs+factives	96.7	96.7	96.5	96.2	97.1	97.0	96.2	96.1	95.8	95.5	96.5	96.3
model A+report verbs+biased words	96.6	96.6	96.5	96.2	97.2	97.1	96.1	96.1	95.8	95.5	96.5	96.4
model A + factives + biased words	96.5	96.5	96.4	96.1	97.0	96.9	96.0	96.0	95.8	95.4	96.5	96.3
model A + factives + biased words+report verbs	96.6	96.6	96.4	96.1	97.1	97.0	96.1	96.1	95.8	95.5	96.5	96.4
model B	91.3	91.2	91.7	91.0	92.4	92.1	90.5	90.3	90.8	90.0	91.3	90.8
model B+readability	97.8	97.8	97.7	97.5	98.2	98.2	97.1	97.1	96.6	96.4	97.3	97.2
model B+quality	91.8	91.7	91.9	91.2	92.6	92.3	91.1	91.0	91.3	90.6	91.8	91.4
model B+readability&quality	97.8	97.8	97.8	97.6	98.4	98.3	97.7	97.7	97.1	96.9	97.8	97.7
w2c	87.2	87.0	88.6	87.9	89.6	89.3	87.2	87.1	88.1	87.4	89.7	89.4
d2v	51.2	49.5	49.3	49.3	50.5	50.6	50.5	48.7	50.5	49.4	50.6	50.1
fastText	84.1	83.9	85.8	85.0	86.4	86.0	82.9	82.9	84.6	83.8	85.7	85.4

Table 6 BBC news and "Professional" Fake News Classification

classifier	XGBoost		Random Forest	
score	Accuracy	F1-Score	Accuracy	F1-Score
model A	96.6	96.6	96.2	96.2
model A+report verbs	96.6	96.6	96.1	96.1
model A+factives	96.6	96.6	96.0	95.9
model A+biased words	96.6	96.5	96.0	95.9
model A+report verbs+factives	96.7	96.7	95.9	95.8
model A+report verbs+biased words	96.6	96.5	96.0	95.9
model A + factives + biased words	96.6	96.5	96.0	95.9
model A + factives + biased words+report verbs	96.6	96.6	96.0	95.9
model B	98.3	98.3	98.5	98.5
model B+readability	98.7	98.7	98.3	98.2
model B+quality	98.4	98.4	98.1	98.1
model B+readability&quality	98.7	98.7	98.4	98.4
w2c	96.3	96.3	96.0	96.0
d2v	47.8	46.9	48.7	48.0
fastText	94.3	94.4	93.8	93.8

We also suggested that well trained word2vec and fastText models may also play roles in the identification of main stream news and online news although it is more difficult to be explained

than language feature patterns. For online ture and fake news classification, we could add more patterns to classify them such as diffusion network patterns [19] [20].

Above all, mainstream media news and online news have significant differences in terms of linguistic usages. Linguistic features can be used either in identifying main stream news with fake news or in identifying main stream news with "professional" fake news. This study has limitations as well. More mainstream news samples from different sources are required to develop more concrete conclusions. We also require fake news that is fabricated by well-trained professional journalists in order to identify the potential different language feature patterns in such fake news.

Acknowledgement

This research was supported by JSPS/MEXT KAKENHI Grant Number JP19H04217.

Reference

[1] "Infographic: Beyond Fake News – 10 Types of Misleading News – Sixteen Languages | EAVI." [Online]. Available: https://eavi.eu/beyond-fake-news-10-types-misleading-info/. [Accessed: 27-Jan-2020].

[2] "Donald Trump Won Because of Facebook." [Online]. Available: http://nymag.com/intelligencer/2016/11/donald-trump-won-because-of-facebook.html. [Accessed: 27-Jan-2020].

[3] "Click and elect: how fake news helped Donald Trump win a real election | Hannah Jane Parkinson | Opinion | The Guardian." [Online]. Available: https://www.theguardian.com/commentisfree/2016/nov/14/fake-news-donald-trump-election-alt-right-social-media-tech-companies. [Accessed: 27-Jan-2020].

[4] M. Recasens, C. Danescu-Niculescu-Mizil, and D. Jurafsky, "Linguistic models for analyzing and detecting biased language," *ACL 2013 - 51st Annu. Meet. Assoc. Comput. Linguist. Proc. Conf.*, vol. 1, pp. 1650–1659, 2013.

[5] R. Baly, G. Karadzhov, D. Alexandrov, J. Glass, and P. Nakov, "Predicting Factuality of Reporting and Bias of News Media Sources," 2018.

[6] H. Rashkin, E. Choi, J. Y. Jang, S. Volkova, Y. Choi, and P. G. Allen, "Truth of Varying Shades: Analyzing Language in Fake News and Political Fact-Checking."

[7] "View of Linguistic Cues to Deception: Identifying Political Trolls on Social Media." [Online]. Available: https://www.aaai.org/ojs/index.php/ICWSM/article/view/3205/3073. [Accessed: 13-Jan-2020].

[8] X. Zhou, A. Jain, V. V Phoha, and R. Zafarani, "arXiv:1904.11679," 2019.

[9] B. D. Horne, W. Dron, S. Khedr, and S. Adalı, "Assessing the News Landscape: A Multi-Module Toolkit for Evaluating the Credibility of News," in *WWW*, 2018.

[10] A. Joulin, E. Grave, P. Bojanowski, and T. Mikolov, "arXiv:1607.01759," 2017.

[11] J. Norregaard, B. D. Horne, and S. Adali, "arXiv:1904.01546 [cs.CY]," Apr. 2019.

[12] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, "FakeNewsNet: A Data Repository with News Content, Social Context and Spatiotemporal Information for Studying Fake News on Social Media."

[13] "Fatemeh Torabi Asr and Maite Taboada (2019) MisInfoText. A collection of news articles, with false and true labels. Dataset." [Online]. Available: https://github.com/sfu-discourse-lab/Misinformation_detection. [Accessed: 27-Jan-2020].

[14] "The Principles of the Truth-O-Meter: PolitiFact's methodology for independent fact-checking | PolitiFact." [Online]. Available: https://www.politifact.com/truth-o-meter/article/2018/feb/12/principles-truth-o-meter-politifacts-methodology-i/. [Accessed: 27-Jan-2020].

[15] "About BuzzFeed News." [Online]. Available: https://www.buzzfeednews.com/article/buzzfeednews/about-buzzfeed-news. [Accessed: 27-Jan-2020].

[16] "Snopes is the internet's definitive fact-checking resource." [Online]. Available: https://www.snopes.com/about-snopes/. [Accessed: 27-Jan-2020].

[17] "Emergent." [Online]. Available: http://www.emergent.info/about. [Accessed: 27-Jan-2020].

[18] "The godfather of fake news - BBC News." [Online]. Available: https://www.bbc.co.uk/news/resources/idt-sh/the_godfather_of_fake_news. [Accessed: 12-Feb-2020].

[19] K. Shu, H. R. Bernard, and H. Liu, "Studying Fake News via Network Analysis: Detection and Mitigation," no. January, pp. 43–65, 2019.

[20] X. Zhou and R. Zafarani, "Network-based Fake News Detection: A Pattern-driven Approach," no. June, Jun. 2019.