Computational and Mathematical Organization Theory (2020) 26:350–364
https://doi.org/10.1007/s10588-020-09307-8

S.I. : SBP-BRIMS 2019

Analyzing and distinguishing fake and real news to mitigate the problem of disinformation

Alina Vereshchaka[1] · Seth Cosimini[2] · Wen Dong[1]

Published online: 21 March 2020

Abstract
Identifying fake news has become an important issue. Increasing usage of social media has led to an increase in the number of people who can be influenced, thus the spread of fake news can potentially impact important events. Fake news has become a major societal issue and a technical challenge for social media companies to identify and has led many to extreme measures, such as WhatsApp deleting two million of its users every month to prevent the spread of fake news. The current problem of fake news is rooted in the historical problem of disinformation, which is false information intentionally, and usually clandestinely, disseminated to manipulate public opinion or obfuscate the truth. Our work addresses the problem of identifying fake news by (i) detecting and analyzing fake news features (ii) identifying the textual and sociocultural characteristics fake news features.

Keywords Fake news · Real news · Fake news identification · Data analysis · Deep learning · Sociocultural textual analysis

1 Introduction

Identifying fake news has become an important issue both for the public and the academic communities. Increasing usage of social media has led to an increase in the number of people who can be influenced, thus the spread of fake news can potentially impact important events, such as elections or contribute to propaganda.

✉ Alina Vereshchaka
avereshc@buffalo.edu

Seth Cosimini
scosimini@unr.edu

Wen Dong
wendong@buffalo.edu

1 Department of Computer Science and Engineering, University at Buffalo, Buffalo, USA

2 Department of English, University of Nevada, Reno, Reno, USA

Recent examples included the controversy created during the 2016 U.S. presidential campaign (Grinberg et al. 2019; Allcott and Gentzkow 2017) and the rise of global propaganda campaigns (Khaldarova and Pantti 2016; Rashkin et al. 2017; Bisgin et al. 2019) by influencing mainstream media and internet discussions. Fake news has become a major societal issue and a technical challenge for social media companies to identify and has led many to extreme measures, such as WhatsApp deleting two million of its users every month to prevent the spread of fake news. One of the other challenges for researchers is the lack of data, since most of the time this kind of dataset has to be manually labeled and there can be difficulties in keeping the consistency in these classification processes (Wu et al. 2019). Our research was conducted using the datasets extracted using the FakeNewsNet tool, and the results may not be generalizable to other datasets since they may contain different features and different formulations of misinformation.

Therefore, there is a strong need for that tool that can early distinguishing fake news and help to stop the viral spread of such news. The current problem of fake news is rooted in the historical problem of disinformation, which is false information intentionally, and usually clandestinely, disseminated to manipulate public opinion or obfuscate the truth. Our work addresses the problem of identifying fake news by (i) detecting and analyzing fake news features (ii) identifying the textual and sociocultural characteristics fake news features.

2 Related work

While the cultural, rhetorical, and sociolinguistic research on fake news and misinformation often employs different but related methodologies to the problem of its identification, they generally align on the need to develop readerships' literacies with certain sensitives to certain linguistic constructions and cultural-historical significance. Linguists, for instance, have used news articles engaging with disinformation and "fake news" to study syntax and lexical cohesion (Jones 2018), but have not turned this attention to language construction as a tool for distinguishing disinformation from reputable reporting. While folklore scholars have examined of the unique vernacular uses of language fake news (Goldstein 2018) and Media Studies researchers have developed analyses of the media consumption patterns for fake news audiences (Nelson and Taneja 2018), the vast majority of language- and text-oriented research on disinformation from the humanities and social sciences comes from pedagogically-focused research on developing robust student literacies in the classroom. Emphasizing that the problem of disinformation is neither new nor unique to our highly-technologically-mediated moment, research focusing on teaching the evolving concept and practice of "information literacy" can be found in the disciplines of Library Science (Bluemle 2018; Lenker 2017) and Rhetoric and Composition (Fash 2017; Miller and Leon 2017; Carillo 2019).

In addition to sociocultural textual analysis we are also interested in applying deep neural networks that have demonstrated success with various text-related classification problems (Zhang et al. 2015; Joulin et al. 2016; Lee and Dernoncourt 2016; Ruchansky et al. 2017). Compared with traditional machine learning methods, deep learning models have achieved better results in fake news identification, as they can automatically learn informative features form the data. Particularly recurrent neural networks (RNN) are suitable for capturing text sequences and using information from preceding texts improves the classification accuracy. Usage of RNN based structures have been achieving promising results in fake news identification (Lai et al. 2015; Ma et al. 2016).

There have been recent advances in applying deep learning methods to assess and explain the credibility of the news (Popat et al. 2018), as well as fake news identification based on graph convolutional neural network (Yao et al. 2019) and reinforcement learning approach (Wang et al. 2019).

3 Methodologies

For this challenge we addressed the problem of fake news identification using three approaches to make it more manageable and accurate. These methodologies include: sociocultural textual analysis, computational linguistics analysis, and textual classification using deep learning. The cultural and sociolinguistic approach allows us to identify the rhetorical and textual characteristics that distinguish real" or "fake" information. From the data science approach, we investigate the data analytics to build a concordance of word and phrase frequency. We also built binary classifiers that extract features from fake and real news using deep learning models, such as, long short term memory (LSTM), recurrent neural network (RNN), and gated recurrent unit (GRU).

3.1 Sociocultural textual analysis

Identifying the rhetorical, cultural, and textual characteristics that distinguish "real" or "fake" information provides a useful methodology for understanding the problem of recognizing disinformation. Accounting for the sociolinguistic, historical, cultural, and ideological meanings attached to particular words, phrases, and syntactical constructions, the methodologies of rhetorical and cultural analysis can show patterns in both the duplicitous forms used in "fake news" to present itself as true information and the ideological intent attached to certain language constructions.

This methodology of sociocultural analysis relies heavily on cultural theorist Stuart Hall's revision to the traditional sender/message/receiver model of communication (Hall 2001). While the sender/message/receiver model presents a linear understanding of communication that focuses on the message as it is transferred from its sender to its receiver, Hall's model of communication recognizes the discursive nature by which information is transmitted and specifically emphasizes the complex structure of relations in which communication takes place. He calls this process by which he charts how this discursive context in which all communication takes places "encoding" and "decoding." By this model, an act of communication is not simply the transference of an object, that is, a message, but a translation of a message into

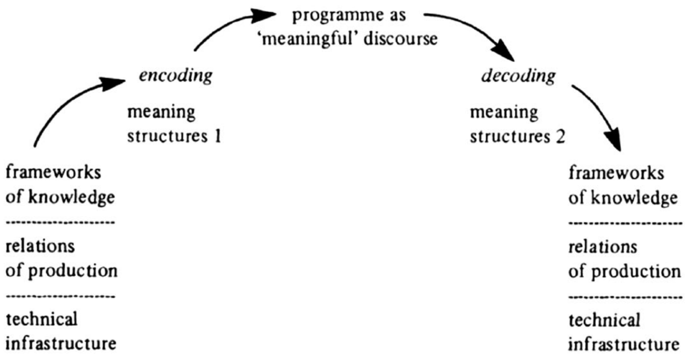


Fig. 1 S. Hall's visual explanation of the encoding/decoding process

a symbolic system structured by historically-established discursive rules that is then decoded by a receiver according to their own sets of discursive rules.

The first and most obvious symbolic structure of a message is often language, which is organized by its attendant and historically-specific rules. That is, even in a news broadcast a "raw" historical event is not transmitted as is, but translated into language, and, depending on the medium of communication, also into aural-visual forms as well for audience consumption. Consumers of this news media, whether they are simply reading it in a newspaper or watching a televised broadcast, must then again "translate," or "decode," the message in order for the communicative act about the "raw" historical event to have any meaning for them. Furthermore, the rules that structure these symbolic systems are organized by the historically-established social practices. Hall writes that because reality, though it does exist outside of language and other symbolic systems, is constantly mediated by language, "what we can know and say has to be produced in and through discourse." A receiver of the encoded message will then decode the message based upon their own relations to the discursive forms through which communication happens, which Hall notes may not always match up with the encoder's preferred meaning (Hall 2001) (Fig. 1).

The discursive forms into which a message is encoded are historically and ideologically structured by social relations, practices, and epistemological frameworks that Hall identifies as "not given but constructed" (Hall 2001). These constructions are the result the complex relations between historical processes and specific, technologically-bound modes of communication, such as news media.

In the context of this study, Hall's work establishes a methodology for considering the discursive forms encoded into the news article by the message's producers, whether they be established journalists or fake news content creators intent on deceiving audiences. As all information is transmitted via the codes of discursive forms, understanding the rules by which these forms work—such as established discursive conventions of journalism—becomes paramount to understanding how encoders established dominant or preferred meanings to their messages and, by this process, what frameworks of knowledge, ideological investments, and social practices are being privileged and re-inscribed through the encoding process. Especially as dominant discursive understandings of news media would establish it as objective or naturally occurring, the problem of identifying fake news begins with Hall's claim that there is "no degree zero in language" and the task of identifying the rules by which the discursive forms of news, both real and fake, operate (Hall 2001).

3.2 N-gram features

The bag-of-words method has a wide range of application: from topic discovery (Wang et al. 2007), to image description (Li et al. 2011). When performing machine learning tasks related to natural language processing, generating n-grams from input sentences can be used as features that represents the text. It has been applied in various domains, such as authorship attribution (Sapkota et al. 2015).

The n-gram model assumes that words are conditionally dependent up to a window-length of n, where n can be 1, 2 or any other positive integers. The choice of n can be challenging, and can be motivated by the computational limitations in practice, as the size of the joint distribution grows exponentially with n. n-gram statistics can be used to capture lexical and syntactic content.

In n-gram models, the probability of any given sequence $w_1, \ldots, w_m$ is approximated by Eq. 1.

$$p(w_1, \ldots, w_m) = \prod_{i=1}^{m} p(w_i | w_{i-(n-1)}, \ldots, w_{i-1}) \tag{1}$$

3.3 Text classification

Text classification is the process of assigning tags or categories to text according to its content. It is one of the fundamental tasks in Natural Language Processing (NLP) with broad applications, such as, sentiment analysis (Go et al. 2009), topic labeling (Wang et al. 2011), spam detection (Wang 2010), and web search (Dumais and Chen 2000).

By utilizing syntax structures of sentences, deep learning models have been proved to be effective for many NLP tasks. Most common neural network structures include recurrent neural network (RNN) (Mikolov et al. 2010), long short term memory (LSTM) (Hochreiter and Schmidhuber 1997) and gated recurrent unit (GRU) (Cho et al. 2014).

LSTM were introduced by Hochreiter and Schmidhuber (1997) and were adapted and applied to a large variety of domains. LSTMs are designed to avoid the long-term dependency problem by remembering information.

LSTM model maps an input sequence $x = (x_1, \ldots, x_T)$ to an output sequence $y = (y_1, \ldots, y_T)$ by calculating the network unit activations using Eqs. 2–7 iteratively from $t = 1$ to T.

$$i_t = \sigma(W^{(i)} x_t + U^{(i)} h_{t-1}) \tag{2}$$

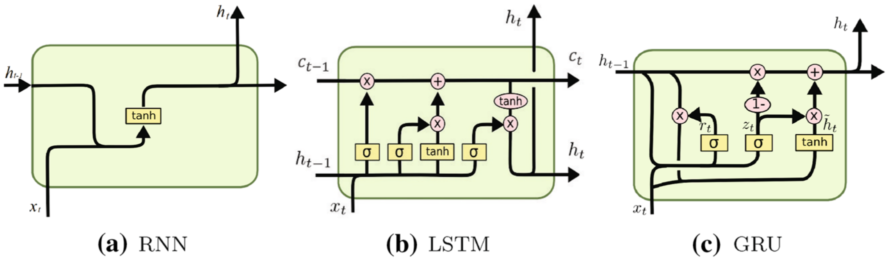


Fig. 2 Structure of **a** RNN unit, **b** LSTM and **c** GRU (Olah 2015)

Table 1 Dataset overview

Total number	Fake	Real
News articles	432	624
News articles with text	420	528
News articles with images	336	447

$$f_t = \sigma(W^{(f)}x_t + U^{(f)}h_{t-1}) \tag{3}$$

$$o_t = \sigma(W^{(o)}x_t + U^{(o)}h_{t-1}) \tag{4}$$

$$\hat{c}_t = \tanh(W^{(c)}x_t + U^{(c)}h_{t-1}) \tag{5}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \hat{c}_t \tag{6}$$

$$h_t = o_t \cdot \tanh(c_t) \tag{7}$$

where W denotes weight matrices (e.g. $W^{(i)}$ is the matrix of weights from the input gate to the input) (Fig. 2).

4 Dataset

To approach this challenge of identifying fake news, we worked with datasets extracted using the FakeNewsNet[1] tool. The final dataset contains both fake and real news in the political domain. The main characteristics of the datasets are shown in Table 1 (Shu et al. 2017a, b, 2018).

[1] https://github.com/KaiDMML/FakeNewsNet.

(a) **(b)**

Fig. 3 Example of the use of all capital letters in both fake (**a**) and real (**b**) news story headlines to communicate urgency

5 Results

5.1 Sociocultural textual analysis

From a textual and rhetorical analysis perspective, the devious tactics of disinformation is frequently located in assumptions made about a readership's literacy. Such "fake news" articles often rely on the having a basic understanding of the signifiers of journalistic and investigative integrity, but are deployed under the assumption that these signifiers will not be given close scrutiny over the dubious content itself. Many of the features that distinguish fake news from real news, then, are not unique to fake news, but rather are unique and dubious uses of features generally found in reputable news sources.

Some of these indicators are easier to see than others. Oftentimes the text of such disinformation takes journalistic conventions to unusual, almost parodic, extremes. Some of these conventional features that appear across numerous categories of fake news sources include: excessive signifiers of urgency or extreme overuse of quotation marks.

In both real and fake news articles, headlines will often begin with a word like "breaking" in all capital letters to catch a reader's attention and communicate an urgency to the information contained within the article (Fig. 3).

Both headlines and the articles' content may redeploy such words and phrases, like "explosive new report" or "warning issued," that may not in themselves indicate the truth or falsity of an article, but the repetition and excessive use of such phrases will often be more prevalent in a piece of disinformation because its aim is: to keep readers' attention; influence readers through pathos, that is, through an appeal to emotions; and conceal the lack of verifiable information behind an authoritative tone. Such urgency can also be noted in the inappropriate overuse of capital letters, which a reputable news source would not include.

Quotation mark usage also shows how punctuation is duplicitously used to represent disinformation as truth. In a "real news" article, quotation marks signify that information is coming directly from a source and, assumedly, has not been

It is being reported that Australia is becoming the first nation in the world to begin

(a)

As Alice Slater, the New York director of the Nuclear Age Peace Foundation, wrote for The Nation, U.S. military bases are not only responsible for such massive amounts of greenhouse emissions but also devastating impacts caused by pollutants and toxic weapons.

(b)

Fig. 4 Examples of in-text reference and sourcing in fake and real news pages. News page stating "it is being reported" rather than citing a reputable source for this information is an example of fake news (**a**), while directly referencing source, even if no direct hyperlink is provided can be an example of real news article (**b**)

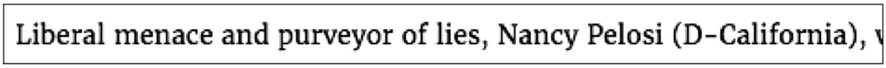
deliberately misrepresented to deceive readers. A common characteristic of "fake news" articles, however, is an overuse of quotation marks, sometimes placing each paragraph in quotation marks, to signify truth simply through a symbolic gesture (Fig. 4).

Beyond the use of quotation marks in order to signify information as factual, understanding sourcing is especially problematic for distinguishing real news from disinformation. A first layer of making this distinction is to note whether or not an online article offers a source at all. Fake news will often report information without even including a source, simply using phrases such as "told reporters" or "is being reported" to signify credibility rather than offering actual sourcing.

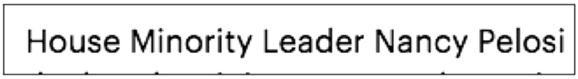
Some articles, however, will include sources, but will either only reference them without directly linking to the source of the information or, potentially even more suspicious, only referencing and linking to information internal to the website to suggest that the author has researched the particular issue being reported but actually not presenting information that would corroborate their analysis. That said, only internal linking may signify something to be suspicious of, but is also dictated by the economics of web-based content: you do not want site visitors to leave, and, in fact, you want to increase clicks on your website. Such an example presents an important problem to keep in mind: profit-driven news organizations will *always* have interests outside of journalistic and investigative integrity.

Furthermore, these characteristics and methods of spreading disinformation across media can also have more complex structures, as in a far-right website *Breitbart*'s now infamous piece "Birth Control Makes Women Unattractive and Crazy" that makes various appeals to authority through quotations and links to outside sources but is, in fact, ideologically-driven disinformation that deliberately misrepresents information with a misogynist intent.

Taking this example of linking to and quoting outside sources but, in fact, presenting disinformation, readers can observe the inclusion of words like "unattractive" and "crazy" in the headline. The deployment of such words represents perhaps one of the most significant differences can be seen in the use of adjectives and adjectival phrases, especially when referencing specific public figures (Fig. 5). For example, consider these two introductions to Congresswoman Nancy Pelosi: "Liberal



(a)



(b)

Fig. 5 Examples of use of adjectives and adjectival phrases for prominent figures referenced in the text. **a** is an example of adjectival phrases providing non-factual, politically-motivated description on fake news page and **b** is an example of adjectival phrase used to provide contextual information on real news page

menace and purveyor of lies, Nancy Pelosi (D-California)" versus "House Minority Leader Nancy Pelosi."

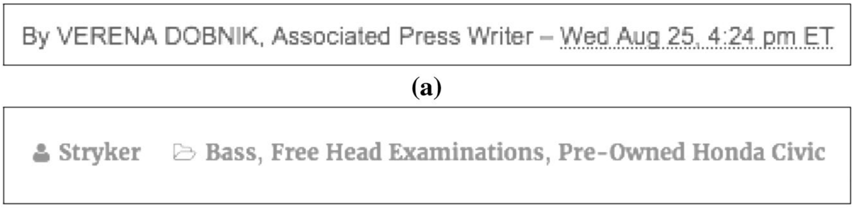
The latter example is from a reputable news source and provides readers with useful contextual information by marking Congresswoman Pelosi's significant position within the House of Representatives. The former introduces her as a "menace" and a "purveyor of lies," neither of which provide contextual information but instead offer value judgments on her person to elicit an emotional response from readers. It is also important to observe, however, that the former also includes the common journalistic practice of identifying a politician's party and state affiliation—"(D-California)"—such that, while the "fake news" article is employing rhetorical tactics that signal its disinformation agenda, it is also signifying reputable journalistic style. Such an example represents the need to assess the entire semiotic structure of a piece of online news media. That is, rather than isolating individual textual or rhetorical aspects, these elements must be assessed in their systematic relationship to one another based on the cultural and ideological contexts in which they are working.

Furthermore, it must also be noted that a clear signaling of particular political orientation within a published piece of writing does not in itself indicate that the information in the piece is fabricated or deliberately misrepresented. The ability to discern true information from such sources requires readers to have a literacy strong enough to understand the difference between information and analysis of the information, and the ideological investments in the textual and rhetorical methods of organizing and distributing information (Fig. 6).

That said, there are some basic clues beyond the rhetorical organization of the written information that are a part of the visual rhetoric of such Web articles. Irrelevant or content-less photos, missing publication date, missing author biography or an author biography that provides no information about their journalistic affiliation, and, perhaps one of the most obvious indicators, erroneous metadata.

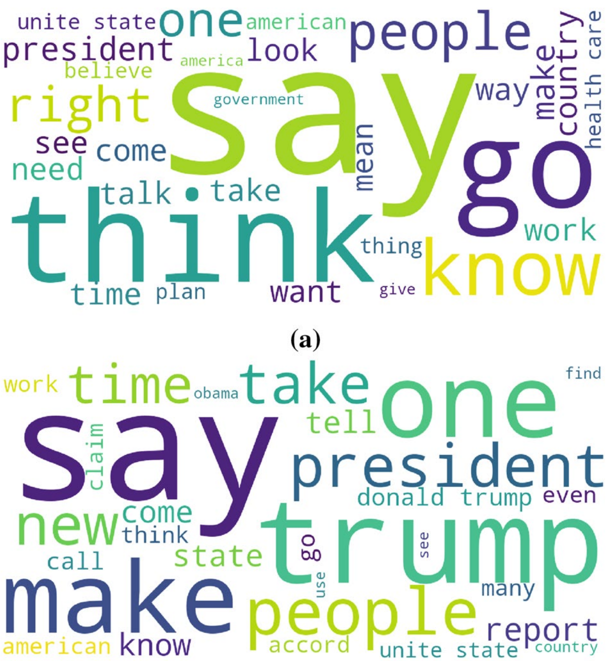
5.2 Computational linguistics analysis

We parse though the real and fake news headlines and text in order to identify the top-20 words (Fig. 7).



(b)

Fig. 6 Examples of differences in publication information for real and fake news pages, where **a** is an example from real news article containing full author name, affiliation, and publication date of article and **b** is an example from fake news tagging the page with erroneous metadata and no identifying information about the writer



(b)

Fig. 7 Wordclouds of top-20 words for real (**a**) and fake news (**b**)

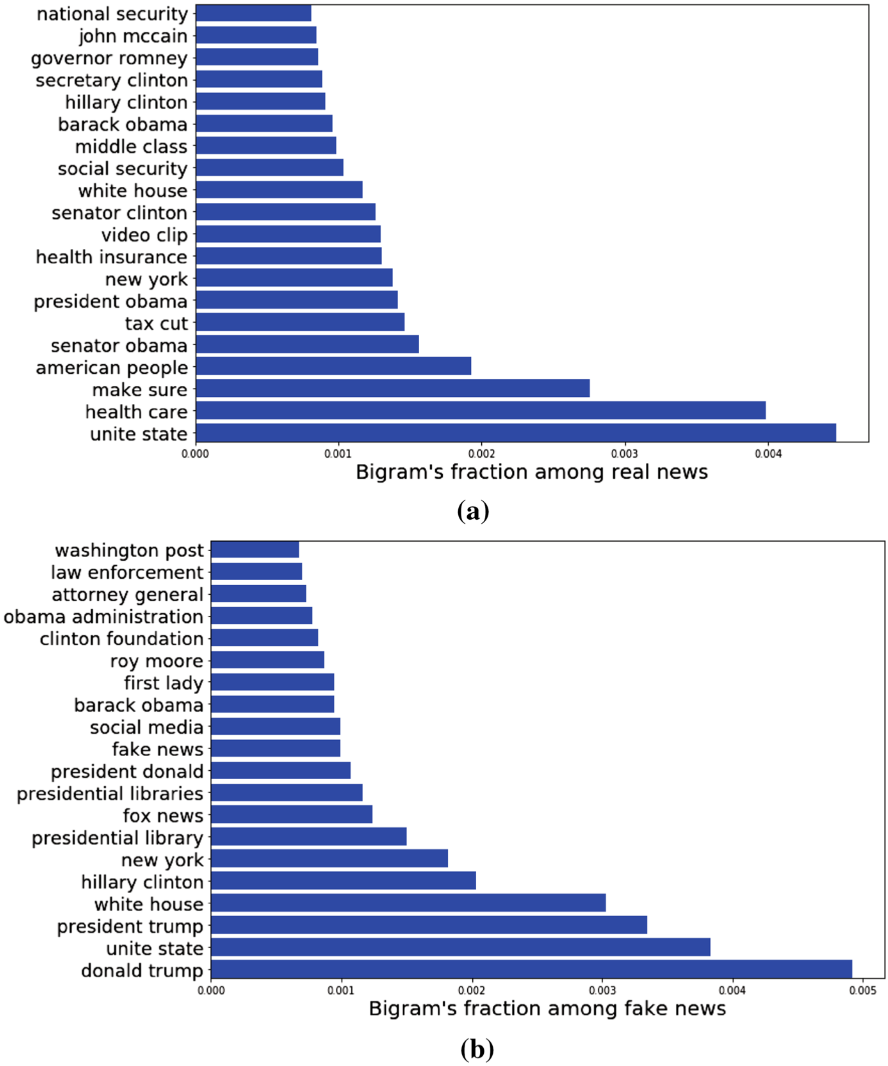


Fig. 8 Top 20 bigrams for real (**a**) and fake news (**b**)

Our bigrams represent top-20 phrase combination among real (Fig. 8a) and fake news (Fig. 8b).

5.3 Deep learning approach

To prepare the data for applying deep learning model, we perform data preprocessing: clean the text of punctuation (Fig. 9), we apply multiple deep learning models to perform binary classification (Fig. 10, Table 2).

```
A Japanese whaling crew has fallen vic
tim to a dramatic full on assault by a
school of killer whales, killing no le
ss than 16 crew members and injuring 1
2, has reported the Japanese Governmen
t this morning.

The crew of the MV Nisshin Maru (日新
丸), Japan's primary whaling vessel and
the world's only whaler factory ship,
was forced to leave the deck temporari
ly as a gas leak was detected within t
he ship's processing factory that resu
lted in the ship being temporarily dis
abled all while continuing to carry ap
proximately 1,000 tons of oil.
```

(a) Original text.

```
A Japanese whaling crew has fallen vic
tim to a dramatic full on assault by a
school of killer whales killing no les
s than  crew members and injuring  has
reported the Japanese Government this
morning  The crew of the MV Nisshin Ma
ru  Japans primary whaling vessel and
the worlds only whaler factory ship wa
s forced to leave the deck temporarily
as a gas leak was detected within the
ships processing factory that resulted
in the ship being temporarily disabled
all while continuing to carry approxim
ately  tons of oil
```

(b) Cleaned text.

```
[6, 6227, 6554, 6838, 615, 8, 2146, 65
53, 5926, 39, 3065, 3930, 3, 6, 3121,
707, 17, 1564, 38, 6, 361, 5, 6554, 67
37, 6615, 6573, 6553, 6554, 6838, 615,
6580, 6553, 6554, 1, 6553, 1684, 76, 4
05, 86, 6554, 1, 1, 6553, 5926, 468,
4, 6554, 8, 6848, 6594, 6581, 6553, 65
54, 1, 1, 6553, 6554, 1, 6553, 39, 86
6, 2, 6227, 104, 16, 509, 6554, 1, 655
3, 2, 5926, 5, 2, 6554, 7111, 6553, 65
54, 6641, 6638, 3494, 1634, 6553, 655
4, 5588, 6711, 6553, 6554, 1, 6553, 65
54, 1, 1, 1, 6553, 6554, 1, 1, 6553, 2
321, 6554, 1, 6553, 595, 795, 6554, 68
38, 615, 8, 2146, 6553, 6554, 6613, 66
38, 4895, 6553, 4, 2, 159, 6554, 1, 65
```

(c) Encoded text.

Fig. 9 Encoding the text

Fig. 10 Accuracy results after applying LSTM, GRU and RNN

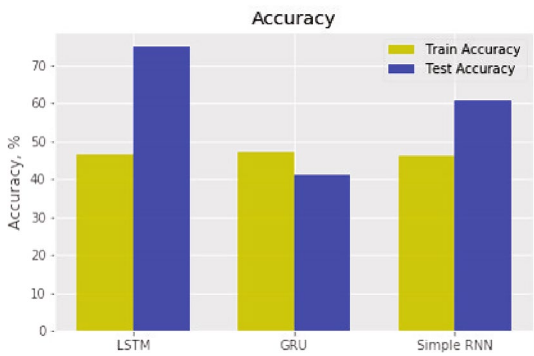


Table 2 Hyperparameters overview

Parameter	LSTM	GRU	RNN
Layers	1	2	1
Activation	–	SeLu	SeLu
Activation output	Sigmoid	Sigmoid	Sigmoid
Optimizer	Adam	Adam	SGD
Epochs	3	1	5
Nodes 1-layer	15	15	15
Nodes 2-layer	–	1	5

6 Conclusion and discussion

After getting the preliminary results, we can notice one of characteristics of disinformation is its ideological context. We are also in the process of synthesizing more in-depth sociocultural linguistic analysis and automatic classifiers using deep learning tools.

To our knowledge, this is the first time this kind of sociocultural textual analysis has been conducted using this dataset.

While these approaches were applied to the dataset, this dataset has its own characteristics of distinguishing information from disinformation that make the results of this research less generalizable than they may otherwise be. In this dataset, rather than distinctions between real and fake news, it provides distinctions from primary sources (such as, transcripts of political speeches and interviews, and statistical reports from governmental agencies and non-profits) from secondary sources (such as, news and Web articles reporting on this data). Regardless of whether or not the secondary sources were in fact real or fake news sources, they were marked as fake news in this dataset for reporting on, rather than being, primary document information. The approaches presented in this research would be better developed and explored using a variety of different datasets distinguishing between real and fake news.

Acknowledgements We want to thank Kathleen M. Carley and Nitin Agarwal for organizing the Disinformation Challenge as part of the International Conference on Social Computing, Behavioral-Cultural Modeling, & Prediction and Behavior Representation in Modeling and Simulation, 2019, Washington, DC, USA. Our winning work for the challenge became the basis for this paper.

References

Allcott H, Gentzkow M (2017) Social media and fake news in the 2016 election. J Econ Perspect 31(2):211–36

Bisgin H, Arslan H, Korkmaz Y (2019) Analyzing the Dabiq magazine: the language and the propaganda structure of ISIS. In: International conference on social computing, behavioral-cultural modeling and prediction and behavior representation in modeling and simulation. Springer, Berlin, pp 1–11

Bluemle SR (2018) Post-facts: information literacy and authority after the 2016 election. Portal Libr Acad 18(2):265–282

Carillo EC (2019) Navigating this perfect storm: teaching critical reading in the face of the common core state standards, fake news, and google. Pedag Crit Approaches Teach Lit Lang Compos Cult 19(1):135–159

Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y (2014) Learning phrase representations using rnn encoder-decoder for statistical machine translation. arXiv preprint. arXiv:1406.1078

Dumais S, Chen H (2000) Hierarchical classification of web content. In Proceedings of the 23rd annual international ACM SIGIR conference on research and development in information retrieval. ACM, New York, pp 256–263

Fash LG (2017) Information literacy in the American literature classroom. Transformations 27(2):195–201

Go A, Bhayani R, Huang L (2009) Twitter sentiment classification using distant supervision. CS224N Project Report, Stanford 1(12)

Goldstein DE (2018) Never remember: fake news turning points and vernacular critiques of bad faith communication. J Am Folklore 131(522):471–481

Grinberg N, Joseph K, Friedland L, Swire-Thompson B, Lazer D (2019) Fake news on twitter during the 2016 us presidential election. Science 363(6425):374–378

Hall S (2001) Encoding/decoding in culture, media, language. In: Reprinted from Media and cultural studies. Blackwell, Malden, pp 166–176

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780

Jones K (2018) ‘A man who revels in his own ignorance, racism and misogyny’: identifiable referents trump indefinite grammar. Funct Linguist 5(1):11

Joulin A, Grave E, Bojanowski P, Mikolov T (2016) Bag of tricks for efficient text classification. arXiv preprint. arXiv:1607.01759

Khaldarova I, Pantti M (2016) Fake news: the narrative battle over the ukrainian conflict. Journal Pract 10(7):891–901
Lai S, Xu L, Liu K, Zhao J (2015) Recurrent convolutional neural networks for text classification. In: Twenty-ninth AAAI conference on artificial intelligence
Lee JY, Dernoncourt F (2016) Sequential short-text classification with recurrent and convolutional neural networks. arXiv preprint. arXiv:1603.03827
Lenker M (2017) Developmentalism: learning as the basis for evaluating information. Portal Libr Acad 17(4):721–737
Li S, Kulkarni G, Berg TL, Berg AC, Choi Y (2011) Composing simple image descriptions using web-scale n-grams. In: Proceedings of the fifteenth conference on computational natural language learning. Association for Computational Linguistics, Portland, pp 220–228
Ma J, Gao W, Mitra P, Kwon S, Jansen BJ, Wong KF, Cha M (2016) Detecting rumors from microblogs with recurrent neural networks. In: IJCAI, pp 3818–3824
Mikolov T, Karafiát M, Burget L, Černockỳ J, Khudanpur S (2010) Recurrent neural network based language model. In: Eleventh annual conference of the international speech communication association
Miller TP, Leon A (2017) Introduction to special issue on literacy, democracy, and fake news: making it right in the era of fast and slow literacies. Lit Compos Stud 5(2):10–23
Nelson JL, Taneja H (2018) The small, disloyal fake news audience: the role of audience availability in fake news consumption. New Media Soc 20(10):3720–3737
Olah C (2015) Understanding LSTM networks. http://colah.github.io/posts/2015-08-Understanding-LSTMs/. Accessed 19 Mar 2020
Popat K, Mukherjee S, Yates A, Weikum G (2018) Declare: debunking fake news and false claims using evidence-aware deep learning. arXiv preprint. arXiv:1809.06416
Rashkin H, Choi E, Jang JY, Volkova S, Choi Y (2017) Truth of varying shades: analyzing language in fake news and political fact-checking. In: Proceedings of the 2017 conference on empirical methods in natural language processing, pp 2931–2937
Ruchansky N, Seo S, Liu Y (2017) Csi: A hybrid deep model for fake news detection. In: Proceedings of the 2017 ACM on conference on information and knowledge management, pp 797–806
Sapkota U, Bethard S, Montes M, Solorio T (2015) Not all character n-grams are created equal: a study in authorship attribution. In: Proceedings of the 2015 conference of the North American chapter of the association for computational linguistics: human language technologies, pp 93–102
Shu K, Sliva A, Wang S, Tang J, Liu H (2017a) Fake news detection on social media: A data mining perspective. ACM SIGKDD Explor Newslett 19(1):22–36
Shu K, Wang S, Liu H (2017b) Exploiting tri-relationship for fake news detection. arXiv preprint. arXiv:1809.01286
Shu K, Mahudeswaran D, Wang S, Lee D, Liu H (2018) FakeNewsNet: a data repository with news content, social context and dynamic information for studying fake news on social media. arXiv preprint. arXiv:1809.01286
Wang AH (2010) Don't follow me: spam detection in Twitter. In: 2010 international conference on security and cryptography (SECRYPT). IEEE, pp 1–10
Wang X, McCallum A, Wei X (2007) Topical n-grams: phrase and topic discovery, with an application to information retrieval. In: Seventh IEEE international conference on data mining (ICDM 2007). IEEE, pp 697–702
Wang X, Wei F, Liu X, Zhou M, Zhang M (2011) Topic sentiment analysis in Twitter: a graph-based hashtag sentiment classification approach. In: Proceedings of the 20th ACM international conference on Information and knowledge management. ACM, New York, , pp 1031–1040
Wang Y, Yang W, Ma F, Xu J, Zhong B, Deng Q, Gao J (2019) Weak supervision for fake news detection via reinforcement learning. arXiv preprint. arXiv:1912.12520
Wu L, Morstatter F, Carley KM, Liu H (2019) Misinformation in social media: definition, manipulation, and detection. ACM SIGKDD Explor Newslett 21(2):80–90
Yao L, Mao C, Luo Y (2019) Graph convolutional networks for text classification. In: Proceedings of the AAAI conference on artificial intelligence, vol 33, pp 7370–7377
Zhang X, Zhao J, LeCun Y (2015) Character-level convolutional networks for text classification. In: Advances in neural information processing systems, pp 649–657

Alina Vereshchaka is a Ph.D. candidate in the Department of Computer Science and Engineering at the State University of New York at Buffalo. Her current research interests include deep reinforcement learning, optimization and multi-agent modeling in stochastic environments. She has conducted studies in the application areas of optimization, transportation, and social science.

Seth Cosimini is a Teaching Assistant Professor of English at the University of Nevada, Reno. His research engages the literature, culture and thought of the African diaspora; American literary and cultural history; and Cultural Studies. He holds a Ph.D. in English from the State University of New York at Buffalo.

Wen Dong is an Assistant Professor of Computer Science and Engineering with a joint appointment at the Institute of Sustainable Transportation and Logistics at the State University of New York at Buffalo. His research focuses on developing machine learning and signal processing tools to study the dynamics of large social systems. He has a Ph.D. degree from the M.I.T. Media Laboratory.