Media-Rich Fake News Detection: A Survey

Shivam B. Parikh and Pradeep K. Atrey
Albany Lab for Privacy and Security, College of Engineering and Applied Sciences
University at Albany, State University of New York, Albany, NY, USA
Email: {sparikh, patrey}@albany.edu

***Abstract*—Fake News has been around for decades and with the advent of social media and modern day journalism at its peak, detection of media-rich fake news has been a popular topic in the research community. Given the challenges associated with detecting fake news research problem, researchers around the globe are trying to understand the basic characteristics of the problem statement. This paper aims to present an insight on characterization of news story in the modern diaspora combined with the differential content types of news story and its impact on readers. Subsequently, we dive into existing fake news detection approaches that are heavily based on text-based analysis, and also describe popular fake news data-sets. We conclude the paper by identifying 4 key open research challenges that can guide future research.**

***Keywords*-Fake News; Media-rich; Social Media**

I. INTRODUCTION

Fake news detection topic has gained a great deal of interest from researchers around the world. There are numerous social science studies have been done on the impact of fake news and how humans react to them. Fake news can be any content that is not truthful and generated to convince its readers to believe in something that is not true. For instance, when a fake news story titled "Did Palestinians Recognize Texas as Part of Mexico?" broke out on social media (refer to left part in Figure 1), multiple news sites and blogs including Reddit.com picked up on this story, the tone of the title of this story convinces its readers to believe that Palestinians recognized Texas as part of Mexico.

Let us pay attention to two key factors of this story (see Figure 1): (i) Title of the news story and (ii) Cover image of the news story. Let us now understand underlying psychological factors that contribute in convincing readers that the story content is realistic - after reading following factors take a step back and analyze how you look at a news story:

- *Headline:* Headline of the story makes a compelling statement about Palestinians recognizing Texas as part of Mexico. An individual who does not know the side-story of this incident would not perform fact checking to find the truthfulness of this news story.
- *Image:* Image of a news story plays a huge role especially when it comes to fake news. In this news story, the image validates what the headline is screaming about. Image portraits, USA President Donald J. Trump and Palestinian President Mahmoud Abbas shaking hands while standing in the Oval office, standing in front of USA and Mexico map where Texas looks like a part of Mexico. In short, the image portraits that USA and Palestine are in agreement with Texas being recognized as part of Mexico.

Now let us have a look at how one of the post shared by Mobin Master (Musician in Newport Beach, California) Page appears to its followers. This post contains a story published on https://www.thebeaverton.com and statistics to prove its outreach is presented in Figure 2.

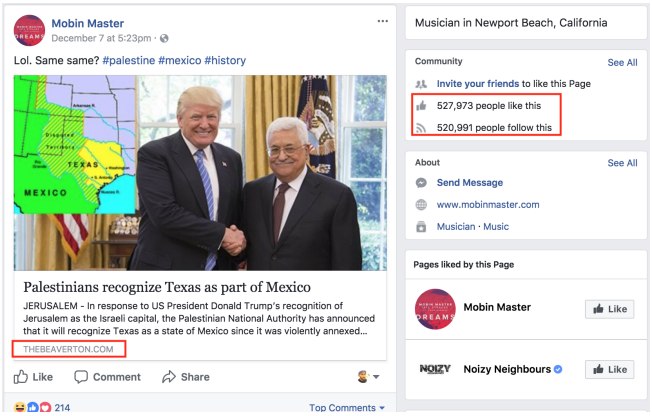

Figure 1: An illustration of how the story titled "Palestinians recognizes Texas as part of Mexico" appears on Facebook [Source: http://www.facebook.com/]

A short list of sources where this story (shown in Figure 1) was published: on snopes site[1] as "Did Palestinians Recognize Texas as Part of Mexico?", on topix site[2] as "Did Palestinians Recognize Texas as Part of Mexico?", on thebeaverton site[3] as "Palestinians recognize Texas as part of Mexico", and on Reddit site[4] as a discussion thread pointing to this story.

[1]https://www.snopes.com/palestinians-texas-mexico/
[2]http://www.topix.com/world/mexico/2017/12/did-palestinians-recognize-texas-as-part-of-mexico
[3]https://www.thebeaverton.com/2017/12/palestinians-recognize-texas-part-mexico/
[4]https://redd.it/7i2prk

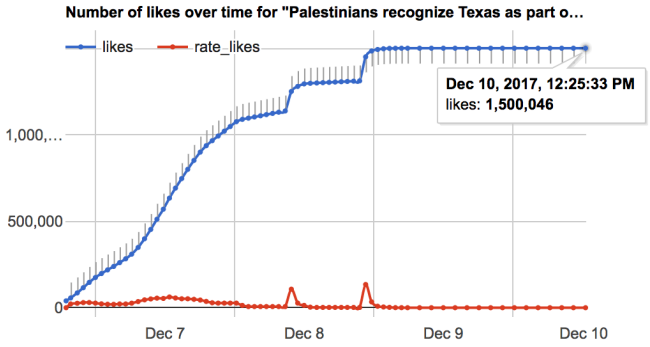

Figure 2: Social Media Trend Report on "Palestinians recognizes Texas as part of Mexico" [Source: http://www.trendolizer.com/]

Here is a look at its social media impact and followed by statistics validating users' reaction on social media. As displayed in Figure 2, a trend chart provided by www.trendolizer.com, this particular story has accumulated over 1.5 million "likes" on social media in just 4 days of being published. That number represents how this news story is trending over social media. What does that number mean? - readers in 21st century mainly focus on headline and multimedia content of the story. It is noteworthy that a news story is generally made of the following three things: headline (usually text in bold), multimedia (image, video, audio, etc.), and body (actual story - usually text content); the first two being more prominent and effective than the third one. A study shows that 70% of Facebook users only read the headline of science stories before commenting or sharing [1].

Why solving this problem is not an easy thing to do: as Chen et al. [2] pointed out that automatic detection of fake news is not an easy problem to solve since these days a news article generally comprises of images and videos (as compared to only text), which is easy to fake. Moreover, with the social media on the rise, fake news stories are very reachable and have a very high impact factor. Also, fake news detection is difficult mainly because there is no governance in-place to control over what citizens can read and what carrier they are using to get that particular news nor who is behind that particular news story. It is safe to say that traditional printed media is slowly dying and every social media account has power to be the news writer/journalist. The challenge we have is, how do we as researchers produce a tool that can help readers of any type of content (i.e. news story) to detect if what they are viewing is fake or real. Before we start coming up with new solutions, it is necessary to survey state of the art techniques for learning purposes.

In the rest of this paper, we cover multiple aspects of research problem of fake news detection. In Section II, we describe various platforms that can be used to disseminate the news content effectively and widely. Next, Section III discusses the types of data a news article can contain and what is the impact of each type of data to readers, and then Section IV provides an understanding of different categories of fake news. Following that, in Section V, we present an overview of existing fake news detection methods and compare them from different perspectives. Further, Section VI describes the existing data sets that are available for fake news detection researchers. Finally, in Section VII, we conclude the paper by highlighting open research challenges in the area of fake news detection.

II. News Carrier Platforms

Carrier platforms are the provider of any news content to the end users. In 2017, two-thirds of U.S. adults get news from social media, that is 5% jump from 2016 numbers [3]. In this section, we breakdown carrier platforms by categories and analyze the underline source of media outlet. Here is a list of specific platforms that are popular among the readers and major source of news to the majority of the audience, broken down by categories.

1) Standalone website: Any sites that produce news stories and each story have a dedicated URL. Generally, these URLs can be used to create a social media post or a share.
   a) Popular news sites: Popular news sites are slightly similar to standalone websites, popular sites also have their own social media presence and tend to generate authentic content.
   b) Blog sites: Blog sites are big on user-generated content and heavily rely on unsupervised content, also considered the best place to get wrong information.
   c) Media sites: These sites are run by content media companies (i.e. The Vox Media), these sites focus on wide range of media-rich content and design their site to drive users by style-based and user-based content creation.
2) Social media: Sharing is the most common way of circulating the content on these sites. More than 70% of its users use them for their daily news source [3].
   a) Facebook (Status — Wall Posts — Dedicated Page — Ad): Users can make a Facebook wall post and/or create Facebook pages and produce/share content using these pages. Better yet, it is concerning to see Facebook allowing users to create paid ads for pretty much any post, which could very well be fake news and can reach larger audience.
   b) Twitter (Tweet — Re-Tweet): Twitter is also a social media site, allows you to create a tweet (limited character) and retweet (share a popular tweet with other).
3) Emails: Emails (news) are also a great way for consumers to receive news, it is really challenging to validate the authenticity of news emails.

4) Broadcast networks (PodCast): Podcasts are an audio multimedia category, very small number of users still use this service and consume their news.
5) Radio service: Radio talk shows are a popular source of news and it is challenging to validate the truthfulness of the audio.

III. Type of Data in News

In this section, we discuss the types of data that the news stories are made of, there are 4 major formats in which users consume their news. Some might be more popular than other types, but they are all major types.

1) Text: Text/string content is generally analyzed by text-linguistic and it is a branch of linguistics, which mainly focuses on the text as a communication system. It is much more than just sentence and words, it has characteristics like tone, grammar, and has pragmatics that allows discourse analysis.
2) Multimedia: Just like name defines, it is an integration of multiple forms of media. This includes images, video, audio, and graphics. This is very visual and catches viewers attention at very first.
3) Hyperlinks or Embedded Content: Hyperlinks enable writers to link off to different sources and gains readers trust by proving the hypothesis of the news story. With advent of social media, writers tend to embed a snapshot of relevant social media post (e.g., Facebook post, tweet, YouTube video, sound cloud clip, Ingram post, etc.)
4) Audio: Audio is a part of multimedia category, but it has standalone medium to be a news source. This category includes podcast, broadcast network, radio service and this medium reaches out to the greater audience to deliver the news.

IV. Fake News Types

Social science researchers have studied fake news from different perspectives and provided a general categorization of different types of fake news, e.g. by Rubin et al. in their recent paper [4]. We summarize this categorization below.

- *Visual-based:* Visual-based type of fake news uses graphical representation a lot more in as content, this includes use of photoshopped images, video, and/or combination of both [5].
- *User-based:* User-based type is oriented towards certain audience by fake accounts and their target audience could represent certain age groups, gender, culture, etc.
- *Post-based:* Post-based fake news are mainly concentrated to be appeared on social media platforms. Post can be a Facebook post along with image or video and caption, a tweet, meme, etc.
- *Network-based:* Network-based news are oriented towards certain members of a particular organization that are connected in one way or the other, this ideology is also applied to group of friends on Facebook and group of mutually connected individuals on LinkedIn.
- *Knowledge-based:* Knowledge-based fake news contain scientific or reasonable explanation to an unresolved issues, these type of news stories are designed to spread false information, e.g. false article on how to cure asthma.
- *Style-based* Style-based focuses on the way of presenting to its readers, fake news are written by majority of people who are not journalists - that being said the style of writing can be different.
- *Stance-based:* Stance-based type in-lines with above mentioned style-based type, stance is different in a sense that it focuses on how statements are being made in an article. Truthful news articles are written in a way to give sufficient information about the subject matter and it is on readers to take way the meaning of the story. Stance-based stories are written to provide very little information about the subject matter and to make a lot of statements (fake arguments).

V. Fake News Detection Methods

Many of existing fake news detection methods heavily rely on *feature extraction*. In [6], [7], [8], [9], [10], [11], authors have proposed approaches that are based on feature extraction. Here, we present a categorization of these approaches their key characteristics and analyze their advantages and limitations.

A. Linguistic Features based Methods

Linguistic based approaches are about using/extracting key linguistic features from fake news, which are described below:

- *Ngrams:* Unigrams and bigrams are extracted from the collection of words in a story. These are preferably stored as TFIDF (Term FrequencyInverse Document Frequency) values for information retrieval. TFIDF refers to a numerical statistic that is intended to reflect how important a word is to the document that it is used in.
- *Punctuation:* Use of punctuation can help the fake news detection algorithm to differentiate between deceptive and truthful texts. Punctuation feature collects eleven types of punctuation, which is implemented through this detection.
- *Psycho-linguistic features*: In order to extract psycho-linguistic features, some researchers recommended the use of LIWC lexicon (Linguistic Inquiry and Word Count) to pick out appropriate proportions of words. This allows the system to determine the tone of the language (e.g., positive emotions, perceptual processes), statistics of the text (e.g., word counts), part-of-speech category (e.g., articles, verbs). LIWC proves to be a valuable tool, it can cluster single LIWC categories into

multiple feature sets: summary categories (e.g., analytical thinking, emotional tone), linguistic processes (e.g., function words, pronouns), and psychological processes (e.g., effective processes, social processes) [12].

- *Readability:* This includes extraction of content features such as the number of characters, complex words, long words, number of syllables, word types, and number of paragraphs [12]. Having these content features allow us to perform readability metrics, such as Flesch-Kincaid, Flesch Reading Ease, Gunning Fog, and the Automatic Readability Index (ARI).
- *Syntax:* This technique extracts a set of features based on CFG (Context-free grammar). These features are heavily dependent on lexicalized production rules combined with their parent and grandparent nodes. Functions in this set are also encoded in TFIDF for information retrieval purposes.

B. Deception Modeling based Methods

The process of clustering deceptive vs. truthful stories relies on theoretical approaches: Rhetorical Structure Theory (RST) and Vector Space Modeling (VSM) [13]. This process involves applying RST, which results in each analyzed text that is converted to a set of rhetorical relations connected in a hierarchical tree, VSM is then utilized to identify the results of rhetorical structure relations. Both of these techniques are briefly described below:

- *RST:* RST procedural analysis captures the logic of a story in terms of functional relations among different meaningful text units and describes a hierarchical structure for each story [13]. According to [6], In past couple of decades, empirical research confirms that writers tend to emphasize certain parts of a text in order to express their most essential idea. RST theory uses rhetorical connections to systematically identify emphasized parts of text.
- *VSM:* VSM is used to identify rhetorical structure relations in RST resulted sets. VSM interprets every news text as vectors in high dimensional space, this requires the extracted text to be modeled in a suitable manner for the application of various computational algorithms [6]. Here, each dimension of the vector space refers to the number of rhetorical relations in a complete set of news reports. This representation makes the vector space very simple and applicable to perform further analysis [14], [15], [16].

RST-VSM methodology gives us an edge of curating data much better than similarity cluster analysis, which is simply based on distances between samples in the original vector space.

C. Clustering based Methods

Clustering is a known method to compare and contrast a large amount of data, in [6], authors have used gCLUTO (Graphical CLUstering TOolkit) clustering package to help differentiate news reports based on their similarity based on chosen clustering algorithm. This method involves running a large number of data set and forming/sorting a small number of clusters using agglomerative clustering with the k-nearest neighbor approach, clustering similar news reports based on the normalized frequency of relations.

The ability of this model to detect the deceptive value of a new story is measured based on the principle of coordinate distances. As proven in [6], after deceptive and non-deceptive cluster centers were computed, new incoming stories were assessed of their deceptive values based on the Euclidean distances to these centers. According to what author claims based on achieving 63% of success using this method, seems to be very useful on large data sets. One obvious challenge could be that this approach might not be able to provide the accurate result, if it is applied on a very recent fake news story, because the similar news story sets might not be available.

D. Predictive Modeling based Methods

In [6], authors proposed a logistic regression model based on training data set of 100 out of 132 news reports. According to this approach, positive coefficients increase the probability of truth and negative one increase the probability of deception. This method gives 70% of accuracy on training data-set and 56% of accuracy on test data-set [6].

Authors claimed that regression indicators like, Disjunction, Purpose, Restatement, and Solutionhood points to truth, and Condition regression indicator pointed to deception [6].

It is very important to note that both Clustering and Predictive Modeling has success rate of 63% and 70% respectively. However, Predictive Modeling approach shows real promise to perform instant fake detection, machine learning techniques can be used to improve the coefficients in ongoing way.

E. Content Cues based Methods

In [7], authors Chen, et al. explained Content Cues Method, this method is based on the ideology of what journalists like to write for users and what users like to read (choice gap). Having certain content on the news story lures users to read more. Contaminated news stories tend to promote interactivity and encouragement that actually attracts users. These news stories are produced by more than one sources delivering the same message, but written in multiple ways. This method leverages two different analyses:

i Lexical and Semantic Levels of Analysis:
Choice of vocabulary plays an important role in convincing readers to believe in the story. Automated methods can be used to extract stylometric features of the text (i.e., part of speech, word length and subjective terms) that can be used to discriminate between two journalistic formats.

Table I: Fake news methods for different fake news types

Fake News Types	Fake News Detection Methods					
	Linguistic	Modeling Deceptive	Clustering	Predictive Modeling	Content Cues	Non-Text Cues
Visual-based	NO	NO	NO	NO	NO	YES
User-based	NO	NO	NO	YES	YES	YES
Post-based	YES	YES	YES	YES	NO	YES**
Network-based	NO	NO	NO	NO	YES	NO
Knowledge-based	NO	NO	YES*	NO	NO	NO
Style-based	YES	NO	NO	YES	YES	NO
Stance-based	NO	NO	NO	NO	NO	NO

*With Limitations: For using **Clustering** Method for **Knowledge-based** Fake News Type: It's success relies on the size of the available data-set.
With Limitations: For using **Non-Text Cues Method for **Post-based** Fake News Type: Post may contain image and/or text (i.e. caption, comments), Non-Text Cues method can apply Image-Analysis techniques to detect tampering, but **Post** without image maybe limitation for using such method.

ii Syntactic and Pragmatic Levels of Analysis:
Pragmatic function of headlines invokes reference to forthcoming parts in the discourse [7]. This is done by making reference to forthcoming parts in the news story. Headlines are written to fill empty thoughts with leveraging ensuing text. This analysis also covers measuring news sites which have more share activity compared to sites that substantially produces more news content.

F. Non-Text Cues based Methods

Authors Chen, et al. [7] explained Non-Text Cues, this mainly focused on the non-text content of the news content. The non-text content of the news story is highly valuable in terms of convincing it's readers to believe in contaminated news. As seen in Fig. 1, the image plays a huge role and it is usually the most eye-catching content of the news story. This method leverages two different analyses:

i Image Analysis:
Strategic use of images is a known key method to manipulate emotion in observers. As shown in fig. 1, a number of readers react to a news story by looking at just headline and an image, therefore image (multimedia) plays a huge role in convincing readers to believe in the subject matter.

ii User Behavior Analysis:
User Behavior Analysis is content-independent method largely useful to assess how readers engage with news once they are lured into the story. News produces have to drive traffic to their original site from multiple avenues, such as, click-ads, social media presence, promotions, etc.. Understanding user behavior and use of teasing images is the key to gain more traction on social media.

Thus far we have understood the basic characteristics of what type of content there are on a news story and existing fake news detection methods. Table I shows that what type of fake news detection method is successful to detect deception particular type of Fake News Content Type.

VI. Fake News Data Sets

The following are popular data-sets that have been used for fake news detection:

- BuzzFeedNews [17]: BuzzFeedNews is a collection of title and links to an actual story or a post that is considered fake news. This data-set is useful for testing Linguistic methods, however, multimedia content is not part of this data-set, therefore certain analysis are not possible on text-only data-set.
- LIAR [18]: LIAR is a bench-marking framework made available by University of California, Santa Barbara researchers. This data-set is also linguistic-based data-set and only contains text only data and has similar limitations like BuzzFeedNews data-set.
- PHEME [19]: This data-set includes rumor tweets, collected and annotated within the journalism use case of the project [19]. It contains Twitter conversations which are initiated by a rumor tweet. Also, it is linguistic based data-set. It contains about 330 conversations (297 in English and 33 Germany).
- CREDBANK [20]: The only data-set has contained social media data and allows users to perform analysis on Twitter data. This data-set signs off on all the categories except the visual data. It misses out on having multimedia data, but still makes it a very compelling choice for researchers who are also focused on fake news detection on social media.

Table II: A comparison of fake news data-sets [11]

Features	Data-sets			
	BuzzFeedNews	LIAR	PHEME	CREDBANK
Linguistic	YES	YES	YES	YES
Visual	NO	NO	NO	NO
User	NO	NO	YES	YES
Post	NO	NO	YES	YES
Network	NO	NO	NO	YES

Table II provides a summary of different data-sets from the perspective of what types of features they are based on. Linguistic and visual features are content-based whereas user, post, and network features are context-based.

VII. Open Research Challenges

We reckon that the following are the key research challenges that can guide future research on fake news detection:

1) Multi-modal Data-set: As Table II demonstrates, out of all 4 popular fake news data-sets, none of the data-set provide a complete multi-modal collection of fake news. This opens up an opportunity for researchers to create a multi-modal data-set that covers all the fake news data types.
2) Multi-modal Verification Method: Number of methods are designed to detect fake news using linguistic approach, and it is very effective in many cases, however visual presentation plays a huge role in people believing in fake news content. This calls for verification of not just language, but images, audio, embedded content (i.e. embedded video, tweet, Facebook post) and hyperlinks (i.e. links to different URLs).
3) Source Verification: Source of the news story has not been done in proposed existing methods, this calls for a new fake news detection method that can perform source verification and considers the source in evaluating fake news stories.
4) Author Credibility Check: One of the method proposes that detecting tone of a news story to detect fake news, research challenge could be that author credibility check allows system to detect chain of news written same author or same group of authors to detect fake news.

REFERENCES

[1] "Study: 70% of facebook users only read the headline of science stories before commenting the science post, 2017," http://thesciencepost.com/study-70-of-facebook-commenters-only-read-the-headline/.

[2] Y. Chen, N. J. Conroy, and V. L. Rubin, "News in an online world: The need for an automatic crap detector," *Proceedings of the Association for Information Science and Technology*, vol. 52, no. 1, pp. 1–4, 2015.

[3] "Two-thirds of american adults get news from social media: survey," https://www.reuters.com/article/us-usa-internet-socialmedia/two-thirds-of-american-adults-get-news-from-social-media-survey-idUSKCN1BJ2A8.

[4] V. L. Rubin, Y. Chen, and N. J. Conroy, "Deception detection for news: three types of fakes," *Proceedings of the Association for Information Science and Technology*, vol. 52, no. 1, pp. 1–4, 2015.

[5] Y. Li, "Image copy-move forgery detection based on polar cosine transform and approximate nearest neighbor searching," *Forensic science international*, vol. 224, no. 1-3, pp. 59–67, 2013.

[6] V. L. Rubin, N. J. Conroy, and Y. Chen, "Towards news verification: Deception detection methods for news discourse," in *Hawaii International Conference on System Sciences*, 2015.

[7] Y. Chen, N. J. Conroy, and V. L. Rubin, "Misleading online content: Recognizing clickbait as false news," in *Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection*, pp. 15–19.

[8] B. Markines, C. Cattuto, and F. Menczer, "Social spam detection," in *Proceedings of the 5th International Workshop on Adversarial Information Retrieval on the Web*. ACM, 2009, pp. 41–48.

[9] N. J. Conroy, V. L. Rubin, and Y. Chen, "Automatic deception detection: Methods for finding fake news," *Proceedings of the Association for Information Science and Technology*, vol. 52, no. 1, pp. 1–4, 2015.

[10] D. S. K. R. Vivek Singh, Rupanjal Dasgupta and I. Ghosh, "Automated fake news detection using linguistic analysis and machine learning," in *International Conference on Social Computing, Behavioral-Cultural Modeling, & Prediction and Behavior Representation in Modeling and Simulation (SBP-BRiMS)*, 2017, pp. 1–3.

[11] 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.

[12] V. Pérez-Rosas, B. Kleinberg, A. Lefevre, and R. Mihalcea, "Automatic detection of fake news," *arXiv preprint arXiv:1708.07104*, 2017.

[13] W. C. Mann and S. A. Thompson, "Rhetorical structure theory: Toward a functional theory of text organization," *Text-Interdisciplinary Journal for the Study of Discourse*, vol. 8, no. 3, pp. 243–281, 1988.

[14] R. Baeza-Yates, B. Ribeiro-Neto *et al.*, *Modern information retrieval*. ACM press New York, 1999, vol. 463.

[15] V. L. Rubin and T. Lukoianova, "Truth and deception at the rhetorical structure level," *Journal of the Association for Information Science and Technology*, vol. 66, no. 5, pp. 905–917, 2015.

[16] V. L. Rubin and T. Vashchilko, "Identification of truth and deception in text: Application of vector space model to rhetorical structure theory," in *Proceedings of the Workshop on Computational Approaches to Deception Detection*. Association for Computational Linguistics, 2012, pp. 97–106.

[17] "Buzzfeednews: 2017-12-fake-news-top-50," https://github.com/BuzzFeedNews/2017-12-fake-news-top-50.

[18] W. Y. Wang, "" liar, liar pants on fire": A new benchmark dataset for fake news detection," *arXiv preprint arXiv:1705.00648*, 2017.

[19] A. Zubiaga, M. Liakata, R. Procter, G. W. S. Hoi, and P. Tolmie, "Analysing how people orient to and spread rumours in social media by looking at conversational threads," *PloS one*, vol. 11, no. 3, p. e0150989, 2016.

[20] T. Mitra and E. Gilbert, "Credbank: A large-scale social media corpus with associated credibility annotations." in *ICWSM*, 2015, pp. 258–267.