Development of a 'Fake News' Machine Learning Classifier and a Dataset for its Testing

William Fleck, Nicholas Snell, Terry Traylor, Jeremy Straub

Department of Computer Science, North Dakota State University

ABSTRACT

Fabricated news stories that contain false information but are presented as factually accurate (commonly known as 'fake news') have generated substantial interest and media attention following the 2016 U.S. presidential election. While the full details of what transpired during the election are still not known, it appears that multiple groups used social media to spread false information packaged in fabricated news articles that were presented as truthful. Some have argued that this campaign had a material impact on the election. Moreover, the 2016 U.S. presidential election is far from the only campaign where fake news had an apparent role. In this paper, work on a counter-fake-news research effort is presented. In the long term, this project is focused on building an indications and warnings systems for potentially deceptive false content.

As part of this project, a dataset of manually classified legitimate and deceptive news articles was curated. The key criteria for classifying legitimate and deceptive articles, identified by the manual classification project, are identified and discussed. The identified criteria can be embodied in a natural language processing system to perform illegitimate content detection. The criteria include the document's source and origin, title, political perspective, and several key content characteristics. This paper presents and evaluates the efficacy of each of these characteristics and their suitability for legitimate versus illegitimate classification. The paper concludes by discussing the use of these characteristics as input to a customized naïve Bayesian probability classifier, the results of the use of this classifier and future work on its development.

Keywords: fake news detection, deep learning, naive Bayesian classifier, qualitative analysis

1. INTRODUCTION

During the final three months of the 2016 U.S. presidential campaign, the 20 top-performing false election stories from hoax sites and hyper partisan blogs generated more engagement than the 20 top-performing stories from 19 major news websites on Facebook.[1] A similar study conducted in France and Italy found equal or greater engagement with false news outlets than popular news brands on Facebook for the two nations.[2] A survey conducted shortly after the 2016 U.S. general election found that fabricated news was able to fool American adults 75 percent of the time.[3] Another survey conducted in December 2016 found that 32% of Americans saw completely made up news 32 percent of the time and not fully accurate news 51 percent of the time when they went online.[4] Further, a study conducted in 2017 by Yale suggested that the more exposure an individual had to fake news, the more they were likely to believe it.[5] The aforementioned findings suggest the significant level of impact and reach that fake news may have had during the 2016 election. This, of course, poses problems for democratic countries that rely on an informed electorate to elect leaders. When the electorate's decisions become based on fabricated information instead of facts, it voters' ability to make informed decisions is impaired and the integrity of the democratic process is damaged.

Because of this, there is a need for voters to have tools at their disposal to help identify whether or not the news they are consuming is fabricated or real. The work described herein aims to assist voters in that effort by creating an automated system to help classify news as either fabricated or real. It builds upon previous work on automated fake news detection. Choi, et al.[6], for example, found that there were certain characteristics that researchers could use to identify fake news. These included strongly and weakly subjective words, which can be used to dramatize or sensationalize a story, hedging, which can indicate vague or obscuring language, and intensifying lexicons from Wiktionary including comparatives, superlatives, action adverbs, manner adverbs, and modal adverbs, which can enliven stories to attract readers. They demonstrated that researchers could use various lexical features to differentiate between reliable and unreliable digital news sources.[6] Horne, et al., similarly, found characteristics in article titles including proper noun and verb phrase usage that differentiated real news from fake news. This

facilitated attaining between 71% and 91% accuracy in classifying articles.[7] Pérez-Rosas, et al.[8], similarly, found key linguistic differences between fake and legitimate news articles. Fake content included social and positive words, more certainty, and focused on present and future actions. The content also included more adverbs, verbs, and punctuation characters than legitimate news. Legitimate content included words associated with cognitive processes such as insight and differentiation and function words such as he, she, and negations.

The work presented in this paper focuses on using the aforementioned characteristics, and others, to determine whether or not an article was fake through qualitative analysis. The methodology of doing so is discussed and its accuracy is characterized. Supporting this paper, a unique database of 277 news articles was curated and manually classified to assist in developing and testing systems to determine the key characteristics of and classify presented articles as fake or real news.

This paper continues with a discussion of the differences between fake and legitimate news articles and establishing a definition for each. Then, key characteristics used to differentiate the two are presented. Following this, the use of these characteristics for a Bayesian classifier is discussed, before concluding.

2. DEFINITIONS

This section defines what "fake" and "real" news are for the purposes of this paper

2.1 Fake News Definition

For the purposes of this paper, fake news is defined as news that could be not be verified as true through organizations such as Snopes, PolitiFact, and FactCheck.org. Here are additional components to this definition:

- News that uses fabricated information to mislead readers.
- News that misconstrues legitimate evidence or reports to deceive readers.
- News that uses evidence that has not been verified (i.e., using discredited journal articles and papers or other fake news articles as evidence).
- Hoax news.
- Satire news.
- Speculation presented as truth.

2.2 Real News Definition

For the purposes of this paper, real news is defined as news that could be verified as true through organizations such as Snopes, PolitiFact, and FactCheck.org. Here are additional components to this definition:

- News that cites evidence from credible, peer reviewed academic journals.
- News that uses primary sources and reporters that have gone through the necessary education to accurately cover content.
- News that identifies speculation as speculation versus presenting it as truth.
- News that comes from organizations which fact check all articles they publish.

3. KEY CHARACTERISTICS

This section presents the primary characteristics used to manually differentiate legitimate news stories from fake news. The characteristics that were examined were article title, author, evidence / source, publisher / origin, and political perspective.

3.1 Article Titles

Through a qualitative analysis of articles in the manually scored and classified database created for this project, there was a clear difference between the real and fake articles in titles. In the examples below, it can be seen that the fake articles use emotional and extreme language not present in the title of the legitimate articles. This characteristic has also been found in the previously discussed work of others.

Table 1. This table compares the titles of fake and real articles to demonstrate the significant differences between the two.

Title Topic	Fake News Article Title	Real News Article Title
Google Political Bias	"If Google is not stopped, the rogue search engine will STEAL every election for Democrats from this day forward" Source: News Target	"Google CEO Says He Leads 'Without Political Bias' In Congressional Testimony" Source: NPR
James Comey FBI Testimony	"Sarah Huckabee Sanders torches criminal Comey after he slams the president, Fox News" Source: The National Sentinel	"Read James Comey's full testimony before House committees" Source: PBS News Hour
Jerome Corsi's Involvement in Robert Mueller Probe	"Mueller proves his own probe is BOGUS by focusing on Jerome Corsi while ignoring blatant lawbreaking by Hillary Clinton" Source: Deepstate.news	"Jerome Corsi, Friend of Roger Stone, Is in Plea Talks With Mueller" Source: New York Times

3.2   Author

One characteristic of fake news is that one author will often write many fake articles. For example, in the database of 277 manually sourced and classified articles, the following authors had written multiple fake news articles:

- Stefan Stanford - 13
- Mike Adams - 11
- Ethan Huff - 11
- John Velisek - 5
- S.D. Wells - 5
- William B. Stoecker – 3

This analysis demonstrates that author name is one factor that can be used to predict article legitimacy in an automated system. The reasoning behind this is that there is a clear trend of authors writing multiple fake articles under the same pen name. The longevity of this type of detection technique is likely limited, however, as fake names (or perhaps alternate fake names, if the current names are themselves fake) could be used to thwart this input.

3.3   Evidence / Source

A qualitative analysis process was utilized and identified key differences in the evidence presented in fake and real articles as well as the sources that they obtained that evidence from. Real articles tended to have authors with verified credentials that included secondary education, writing awards, and previous experience at reputable, national news institutions. Real articles also tended to cite other real news sources as well as primary sources of information such as press releases, public officials, and publicly known and verifiable information. These articles typically also have to go through a rigorous review by their organization (i.e., review by editors before being published).

Fake articles, on the other hand, often cited other fake sources and often did not reference publicly available and verifiable information. There was also a notable difference in how information and evidence was presented in real and fake articles. Real articles always omitted a personal opinion of the author while fake articles frequently

included it. Fake articles also often made unsubstantiated claims or claims that had been proven to be false by fact checkers. Examples of demonstrative quotations from real and fake articles are presented in Table 2.

Table 2. This table compares the evidence and sources in fake and real articles side to side to point out the differences between the two.

Topic	Article Evidence / Source Comparison
Jerome Corsi's Involvement in Robert Mueller Prove	"Jerome Corsi, Friend of Roger Stone, Is in Plea Talks With Mueller" Real article source: New York Times "Mueller proves his own probe is BOGUS by focusing on Jerome Corsi while ignoring blatant lawbreaking by Hillary Clinton" Fake article source: Deepstate.news Author Legitimacy The real article provided biography links for both reporters that contained their credentials while the fake article provides no author credentials, only links to other articles written by them. These credentials give claims made by authors credibility. Article Sourcing / Reference Legitimacy The real article references The Washington Post while the fake article cited TheGatewayPundit.com and TheNationalSentinel.com as its sources. Politifact's Fake News Almanac classifies TheGatewayPundit.com as having some fake stories[9] and TheNationalSentinel.com has been manually classified as having fake stories. Presentation of Evidence and Claim Legitimacy The real article made claims based on verifiable evidence while the fake article made unsubstantiated claims repeatedly.
Generic Comparison (*Compares the characteristics of a fake and real article regardless of their topic*)	"Republican Martha McSally Picked To Fill Senate Seat Formerly Held By John McCain" Real article source: NPR "Lunatic Joy Behar demonstrates tyranny of the Left: Anyone who merely SUPPORTS Trump should be thrown in jail, she says" Fake article source: Liberal Mob Author Credibility The real article provided a link to the author's credentials while the fake article provided no author credentials and only links to other articles written by them. Article Sourcing / Reference Legitimacy The information presented in the real article was written by a reporter with verifiable credentials who reported the facts of the situation and statements of relevant officials without any personal opinions. The fake article used NewsTarget.com as one of its references which has been manually classified as a site that contains fake news. Presentation of Evidence and Claim Legitimacy The author of the fake article repeatedly made unsubstantiated claims such as "If Behar and the rest of the authoritarian Left had their way, the nearly 65 million American voters who cast ballots for POTUS Trump are

	co-conspirators and thus equally guilty for those pretend crimes he has been accused of committing" while the real article only presented verifiable facts and only public, and not private, speculation.
Harvard University Admission Practices Relating to Asians Students Being Denied Entry	"The Justice Department threw its hat into a bitter fight over Harvard's admission practices, and now the ACLU is getting involved" Real article source: Business Insider "The ACLU declares war on Asian-Americans, says universities should punish people for being born Asian" Fake article source: Liberal Mob Author Legitimacy The real article provided a link to the author's credentials while the fake article provided no author credentials and only links to other articles written by them. Article Sourcing / Reference Legitimacy The fake article listed NaturalNews.com as one of its sources which is considered propaganda[10] by U.S. News and manually classified to contain fake news. The real article referenced another Business Insider article and a press release from the ACLU. Evidence The author of the real article with verifiable journalism credentials presented the facts of the case without any personal opinions while the fake article contained phrases such as "If you're white, the American Civil Liberties Union (ACLU) wants to see you wiped off the face of the planet" and other opinions of the author who had no journalistic credentials listed.
Jerome Corsi's Involvement in Robert Mueller Prove	"Jerome Corsi, Friend of Roger Stone, Is in Plea Talks With Mueller" Real article source: New York Times "Mueller proves his own probe is BOGUS by focusing on Jerome Corsi while ignoring blatant lawbreaking by Hillary Clinton" Fake article source: Deepstate.news Author Legitimacy The real article provided biography links for both reporters that contained their credentials while the fake article provides no author credentials, only links to other articles written by them. These credentials give claims made by authors credibility. Article Sourcing / Reference Legitimacy The real article references The Washington Post while the fake article cited TheGatewayPundit.com and TheNationalSentinel.com as its sources. Politifact's Fake News Almanac classifies TheGatewayPundit.com as having some fake stories[9] and TheNationalSentinel.com has been manually classified as having fake stories. Presentation of Evidence and Claim Legitimacy The real article made claims based on verifiable evidence while the fake article made unsubstantiated claims repeatedly.

3.4   Origin / Publisher

Another prominent and frequent characteristic of fake news articles is that the same articles were often published across multiple platforms. This is demonstrated in Table 3. The first column contains a news source that was manually identified as having a large number of fake articles. The second column indicates whether or not fake articles on that site appeared on other sites listed in the first column as well. Finding the same fake articles on multiple sites can be a mechanism that is used to identified other questionable sites and to aid the classification of those sites as containing fake news. With a set of questionable sites identified, the site origin of an article can serve as a valuable input in predicting whether or not an article is fake. In all of the cases listed in Table 3, the site had at multiple fake articles.

Table 3. This tables demonstrates that there is a high frequency of fake articles on one site being found on other sites containing fake articles as well.

Origin / Publisher	Contained Fake News Articles	Same Articles Found on Multiple Sites
https://www.naturalnews.com/	Yes	Yes
https://censorship.news/	Yes	Yes
https://www.food.news/	Yes	Yes
https://www.health.news/	Yes	Yes
https://www.pollution.news/	Yes	Yes
https://www.cancer.news/	Yes	Yes
https://www.climate.news/	Yes	Yes
https://www.science.news/	Yes	Yes
https://www.wsws.org/	Yes	No
https://worldtruth.tv/	Yes	Yes
https://libtards.news/	Yes	Yes
https://liberalmob.com/	Yes	Yes
http://thoughtpolice.news/	Yes	Yes
http://identitypolitics.news/	Yes	Yes
http://allnewspipeline.com/	Yes	No
https://deepstate.news	Yes	Yes

3.5   Political Perspective

Another major input into the process of determining whether or not an article was fake was its political perspective. In many cases, articles with conservative, centrist, and liberal perspectives had an apparent bias; however, the facts presented in them were almost always true. However, articles containing far-left and far-right perspectives frequently contained fabricated and misleading information and unsubstantiated or blatantly false claims. The author's opinion also appeared throughout the article. While the presence of a political perspective is not necessarily indicative of false content, articles with strong political perspectives should be viewed with higher levels of scrutiny. While the political perspective of an article could perhaps be used as an effective predictor to whether or not it is fake or real, this has implications beyond the purely technical. Care needs to be taken that any fake news identification system does not, itself, develop a bias by explicitly or implicitly associating fake content with certain viewpoints. Table 4 presents language that could be used to identify an article as being written by a far left or far right source versus a source with more moderate liberal or conservative bias or a source that was neutral.

Table 4. This table compares the language revealing political perspectives of fake and real articles to demonstrate the significant differences between the two.

Political Topic	Alt-Left / Alt-Right Language	Conservation, Liberal, or Neutral Language
New York Abortion Laws	"We have New York Democrats jumping up and down in fits of joy, celebrating the outright murder of live-birth babies, who could be completely healthy, but still suffocated to death, or their spinal cord severed, or getting their jugular slit with a scalpel." Source: Natural News	"New York, which last week passed a controversial law loosening the rules on abortion, is just the latest state to do so, as advocates say they fear the newly cemented U.S. Supreme Court could overturn Roe v. Wade." Source: Fox News
Slavery Reparations	"The liberal mob sees absolutely nothing wrong with, and even actively encourages, treating white people as second-class citizens because doing so is their way of creating "equality" throughout society – even though judging people for the way they were born doesn't apply to brown and black people, nor does it apply to the LGBT mafia, which we're all told were "born that way." Twenty-twenty presidential wannabes Kamala Harris and Alexandria Ocasio-Cortez (AOC) are two prominent examples of how modern-day Leftists have somehow convinced themselves that being racist against white people is a "moral virtue," as they push for reparations to "right the wrongs of the past."" Source: Natural News	"Several Democratic presidential candidates are embracing reparations for the descendants of slaves — but not in the traditional sense. Over the past week, Sen. Kamala Harris of California, Sen. Elizabeth Warren of Massachusetts and former Housing and Urban Development Secretary Julián Castro spoke of the need for the U.S. government to reckon with and make up for centuries of stolen labor and legal oppression. But instead of backing the direct compensation of African-Americans for the legacy of slavery, the Democratic candidates are talking about using tax credits and other subsidies." Source: NBC News
Green New Deal	"Bottom line: AOC's "Green New Deal" is a systemically corrupt plan to disembowel the entire functioning U.S. infrastructure, debilitate all farming capabilities and initiate a catastrophic food collapse, wipe out the entire middle class, confiscate all guns, private property and private businesses, and put in place a tyrannical police-state government that will never be able to be removed without a full-scale revolution."	"A draft text of the plan circulating Congress includes a framework aimed at eliminating greenhouse gas emission from manufacturing and agriculture and "dramatically" expanding energy sources to meet 100 percent of power demand through renewable sources. The proposal describes this effort as a "historic opportunity to virtually eliminate poverty in the United States and to make prosperity, wealth and economic security available to everyone participating in the transformation." The proposal also

	Source: Natural News	calls for a job-guarantee program offering a "living wage job to every person who wants one," a plan to aid workers affected by climate change, universal health care and basic income programs, among other things." Source: Fox News

4. USING CHARACTERISTICS AS WEIGHTS FOR NAIVES BAYES CLASSIFIER

This section discusses how each of the aforementioned characteristics could be used as an input into a custom naïve Bayesian classifier system.

4.1 Title

Benjamin Horne and his colleagues found characteristics in article titles including proper noun and verb phrase usage that differentiated real news from fake news and facilitated attaining between 71 percent to 91 percent accuracy in classifying articles[7]. The article title could be used to provide a similar benefit to the naïve Bayesian classifier. An input based on the identification of words in all caps in article titles, which was a characteristic that was only found in fake news article titles, could also be added.

4.2 Author

An article's author could be used as an input by cross referencing the name of an author against the names of authors of fake news articles which are in a database that has been manually (or otherwise reliably) classified. If an author name was associated with real articles, it could be used to support an article being true and if a name was associated with fake articles, it could be used to support an article being false.

4.3 Evidence / Source

While the work presented herein has demonstrated the efficacy of evaluating the evidence and sources used by an article in its classification, there is no way to immediately implement this as an input to a classifier. Determining how to best incorporate the evidence and sources used by an article as an input to a naïve Bayesian classifier will serve as a topic for prospective future work.

Approaches to this may include incorporating items such as links to an author's biography and lexical analysis. It is key that any implemented approach be effective at aiding classification, reliable and automated. The Lexical analysis approach could be particularly valuable as it would prospectively analyze evidence presented directly within articles which should be helpful in detecting fake news. This would take current practices a step further by specifically looking at evidence versus the words in the article as a whole.

4.4 Origin / Publisher

There are two related pieces of information, related to article origin, that can serve as inputs to a classifier. First, the source of an article can be cross referenced with an existing classified database to see if that site is associated with fake news. Second, an article's text can be cross referenced with the text of articles from other sites to see if the same text is appearing on multiple sites (including those known to host fake news articles). The presence of content on multiple sites was indicative of fake news; however, this is not absolute as some legitimate sources (such as the Associated Press) produce content for syndication. However, typically, this legitimate syndicated content discloses it syndicated nature, while fake articles do not.

4.5 Political Perspective

The analysis performed showed that articles offering an alt-left or alt-right political perspective often had distinct language that differed significantly from articles that had less pronounced conservative or liberal biases or articles of a more neutral or non-political nature. Because of this, another input that can be used is a lexical analysis that looks for language that conveys a certain political bias. This remains a topic for future work. However, if developed, it could add additional information that would make the classifier even more accurate.

4.6 Overall Use of Inputs

By combining all of the aforementioned as inputs into one naïve Bayesian classifier, automated detection of fake news could become more effective than current methods. The classifier could be used to determine if an article is fake or real by having a threshold that would be used to infer an article's legitimacy. This remains a topic for future work.

5. CONCLUSIONS & FUTURE WORK

By using key characteristics such as an article's title, author, evidence and sources, origin and publisher, and political perspective, it is possible to determine reliability, through a qualitative analysis, whether or not an article is deliberatively deceptive (fake) or real news. These characteristics also have the potential of being used as inputs to an automated system as well to determine article legitimacy.

Further research is needed for how to better identify fake articles, in particular. Real articles, coming from sources like the New York Times or NPR, were very easy to identify; however, the risk of an overreliance on article source is that an attacker could potentially compromise a site and place deceptive content on it. Thus, source must be only one of several metrics used.

Fake articles, however, were much more difficult to determine with certainty. Therefore, future work in this area should focus on identifying additional characteristics that differentiate deliberately deceptive (fake) articles from real ones. Additional work is also needed on the Bayesian classifier system.

ACKNOWLEDGEMENTS

REFERENCES

[1] Silverman, C., "This Analysis Shows How Viral Fake Election News Stories Outperformed Real News On Facebook," Buzzfeed News, (2016).

[2] Fletcher, R., Cornia, A., Graves, L., Nielsen, R., "Measuring the reach of "fake news" and online disinformation in Europe," Reuters Institute/University of Oxford, (2018).

[3] Silverman, C., Singer-Vine, J., "Most Americans Who See Fake News Believe It, New Survey Says," Buzzfeed News/Ipsos Public Affairs, (2016).

[4] Barthel, M., Mitchell, A., Holcomb, J., "Many Americans Believe Fake News Is Sowing Confusion," Pew Research Center, (2016).

[5] Pennycook, G., Cannon, T., Rand, D., "Prior Exposure Increases Perceived Accuracy of Fake News," Forthcoming in Journal of Experimental Psychology: General, (2018).

[6] Choi, E., Choi, E., Jang, J., Svitlana, V., Rashkin, H., "Truth of Varying Shades: Analyzing Language in Fake News and Political Fact-Checking," Paul G. Allen School of Computer Science & Engineering, University of Washington/Data Sciences and Analytics, Pacific Northwest National Laboratory, (2017).

[7] Horne, B., Adah, S., "This Just In: Fake News Packs a Lot in Title, Uses Simpler, Repetitive Content in Text Body, More Similar to Satire than Real News," Rensselaer Polytechnic Institute, (2017).

[8] Pérez-Rosas, V., Kleinberg, B., Lefevre, A., Mihalcea, R., "Automatic Detection of Fake News," Computer Science and Engineering, University of Michigan/Department of Psychology, University of Amsterdam, (2017).

[9] Gillin, J., Waychoff, S., "POLITIFACT'S FAKE NEWS ALMANAC," PolitiFact, (2017).

[10] Dicker, R., "Avoid These Fake News Sites at All Costs," U.S. News, (2016).