Detecting Fake News Through Machine Learning Techniques

Daemin Lee
College of Computing and Software Engineering
Kennesaw State University
Marietta, GA, United States of America
dlee80@students.kennesaw.edu

Joe Potchen
College of Computing and Software Engineering
Kennesaw State University
Atlanta, Georgia, United States
jpotchen@students.kennesaw.edu

Leland Burns
College of Computing and Software Engineering
Kennesaw State University
Marietta, GA, United States of America
email address

Jason Wein
College of Computing and Software Engineering
Kennesaw State University
Marietta, United States
Jwein2@students.kennesaw.edu

Kyle Hedden
College of Computing and Software Engineering
Kennesaw State University
Marietta, United States
Khedden1@students.kennesaw.edu

***Abstract*—Fake News is an widespread problem for users on social media and on the internet. The elusiveness to detect for "fake news" is difficult to detect and determine by an simple analysis.**

By using Machine Learning techniques it is possible to determine whether a given news article is "Fake News" or not by utilizing a data set that consists of web-scraped articles that have been classified into "fake-news" or not. One solution would be to use critical thinking and research on an suspected article by referencing against other news sources that may be covering the same story or debunking on a fake news article or social media post that would be spreading.

Though, through using an data set containing a mix of different types of news including fake news to predict for fake news identification through Artificial Neural Networks. Based on the results assessed, It is possible to determine fake news, but biases are likely to occur based on the researchers' biases.

1. Introduction

Fake News in its most raw form is an embodiment of yellow journalism that is used to falsify, fabricate, and exaggerate online news stories in some way.

The effects of Fake News has caused political and societal discourse that may have had effects to the government of the United States intentionally by foreign governments.

At the end of the 2016 Presidential Election, Donald J. Trump had won the election by the electoral college. Midway of the first year of his presidency, the FBI and CIA had alleged that there were possible meddling by the Russian government through the spread of "Fake News" on social media platforms [3].

The awareness of this problem has caused for concern in what the public is able to discern between faked and real news.

The research group believes that using Machine Learning to identify for fake news is needed to combat the expanding amount of misinformation that has found its place as an problem within modern journalism.

1.1. Research Problem

This paper proposes the question of whether it is possible to detect fake news through machine learning models. Specifically, the aim of this project is to determine the ideal model that is efficient in predicting fake news while also limiting the cost of memory and storage for computation. "Fake news" has been a very recent and prevalent problem within recent years.

A recent and direct consequence of fake news would be the "Pizzagate" controversy [2] where conspiracy theorist, Alex Jones, had claimed that a pizza shop basement in Washington D.C. was part of a child trafficking ring that Hillary Clinton and Democrats were allegedly part of. This theory riled many of the viewers to the point where one of Jones's viewers, Edgar Maddison Welch, had armed himself with an AR-15 assualt rifle. Drove out to that exact pizza shop surrounding the conspiracy and starting shooting at it in an attempt to save the children as purported by the claim. Later, he would realize that there were no enslaved children in the basement.

As a whole, fake news is a societal, and informational problem that can pose detrimental to democracy with the potential to cause physical and societal collateral damage.

As a consequence of the increase in cases of fake news in recent years, efforts have been made to crackdown on the spread of misinformation throughout social media platforms. All popular social media platforms (Facebook, Twitter, Spotify, and YouTube) have permanently banned Alex Jones from using their networks [8] following the events of "Pizza Gate" in addition to multiple questionable accusations made by Jones, including an accusation made by Jones claiming that the Sandy Hook shooting was "faked".

Despite efforts of many social media websites and governments cracking down on fake news, many young people today generally are not able to tell the difference between fake news and real news. According to a Stanford study it found that many students have a very strong inability in discerning between fake news. In the study, high school students high school students were were given two posts announcing the candidacy of Donald Trump's presidential Campaign. One post was given by an actual Fox News account another one posted by an account that "looked" like it was from Fox News. 25% of the could not tell the difference between real and fake news sources. With over 30% of students favoring that the fake news account was more trustworthy.

In fact, some politically charged or false articles that would be deemed untrue often have more views and shares on social media sites than actual news articles towards the last three months of the election. According to an analysis by *Buzzfeed*, posts and stories written from the top twenty highest-performing hoax sites and hyperpartisans had over 8.7 million shares, "reactions" and comments compared to the top twenty highest-performing major news organizations had nearly 7.4 million shares, "reactions" and comments on social media sites.

The research problem was originally defined through the following use cases: Based on an single event/story, the system would determine if certain sources or articles are deemed to be fake news based on a given probability. Through the sources analyzed the machine learning agent would assign an level of bias and factuality of these articles by comparing them to each other and assign scores of the sources bias and factuality. By using binary classification, it would be possible to determine the probability of a given source to be fake news. Based on count of strong adjectives within the articles and sensationalist headline. This would be modeled using a Logistic Regression model.

As time progressed through the development of the project, it was decided that it would be subjectively cohesive if we had created the different machine learning models that were aimed to solve the same problem in discerning between fake news articles and non-fake news articles. As it would strictly determine if the news article was fake news or not and determine if the model was accurate

1.2. Purpose of the Study

The difficulty in cracking down on fake news is how subtle and invisible it can be to those who are unable to tell the differences between them. Hopefully through machine learning it could help users differentiate between them.

The main purposes of the study is to find the best machine learning technique that would be used to optimize the process of finding fake news. This project takes into account accuracy and speed. Accuracy refers to how accurate the machine learning model is able to predict if new articles are deemed to be fake news accurately based on the dataset's classification of the news article.

Speed is measured in how fast the machine learning model is at processing the data into the model

In understanding and critically analyzing these data models and its qualities we would be able to fully assess the efficiency of each of these models.

1.3. Audience

The findings of this project would appeal to organizations within social media, political journalism, and federal investigations. In the hopes that it would help on cracking down on the spread of misinformation. The model could also be implemented into an application that tells a reader whether and article they're about to read is fake news or not.

1.4. Contribution

The project's contributions are split between six members. Joe Potchen researched, cleansed and formalized the dataset used for the machine learning techniques and implemented the models. Leland Burns helped visualize the models' results through the SciKit library. Jason Wein implemented the machine learning models of the project. Kyle Hedden helped with research of the other related works done by external organizations. Daemin Lee served as the project lead in conceptualizing the idea and technical writing for the group.

1.5. Motivation

The motivation for research on this topic was that this is a relatively new area of research with many opinions but not many concrete solutions. Many implementations focus primarily on the host of the article, but even articles hosted on otherwise trustworthy websites can be classified as fake news. The primary motivation of this project was to bring awareness, propose a solution, and work towards minimizing the effects of fake news.

1.6. Paper Goals and Organization

The purpose of this paper is to design and implement a machine learning implementation that correctly predicts

if a given article would be considered as fake news. The contributions of this paper are as follows:

- Introduces the topic of fake news and the various machine learning algorithms that would be effective in classifying fake news.
- Provides an overview of the history and implications of fake news.
- Presents a possible solution and lays some ground work in further study in this area.

Section 2 introduces the techniques that were used in this study. Section 3 discusses the proposed method for dealing with fake news. Section 4 discusses the conclusions of the project and possible improvements for future research.

1.7. Related Works

Several groups and organization have also worked on similar ideas in their own implementations. These works highlight some of the challenges of fake news detection.

One implementation by Katharine Jarmul, founder of data analysis company Kjamistan, uses a Passive Aggressive Classifier to detect fake news [4]. The implementation is a tutorial on using different Bayesian models posted on DataCamp [4], which offers courses on a variety of data science topics including R and Python.

The set-up in extrapolating the data would first take in all the words found within an article content and places each of them in to a vector excluding the "stop-words" generally being words that are used very commonly in the english language. Which would be words consisting of articles, conjunctions, and pronouns as an example. Vectors are a data structure within programming that would hold a list of data, and information.

Once the articles are vectorized, a passive-aggressive classifier classifies each of the words through *word2vec* this classification model embeds the keywords from the vectors from each of the articles. By the classification, it would highlight the keywords of the articles. Sci-kit library holds an support vector machine model that the data is being fed into. Support Vector Machines is similar to using Linear Regression with the addition that it seeks to optimize the line of regression as found for linear regression such that spacing distances between boundary data within a graph's line of regression are formatted into a plane.

The findings within her tutorial details on how it is possible to create a machine learning model that can detect for fake news through machine learning models by embedding the words either through tf-df, or word2vec word embedding techniques.

One paper titled 'Exploiting Network Structure to Detect Fake News', written by three Stanford University students, also implements a Neural Network for classifying fake news [11]. Their implementation also takes into account the social context in addition to article-specific features such as the title and content in an article in an attempt to improve prediction accuracy. This is one of the few possible ways to improve prediction accuracy without improving upon the natural language processing.

Another paper titled 'Fake News Detection: A Deep Learning Approach', implemented three different neural network models to compare with the only difference between them being how they took in the article content and title [10]. This indicates that the way one goes about processing text in an article makes a huge difference in the performance of a model. This makes sense considering that an article's content is generally the only thing that can be analyzed to truly determine its authenticity.

2. Techniques used in the study

This research paper was implemented using the Support Vector Machines, Passive Aggressive Classification, Logistic Regression, and Recurrent Neural Network machine learning models.

3. Proposed Method

Before the implementation process, it was necessary to understand the process of determining if certain articles are deemed to be fake news. After conducting interviews with Kelsey Richardson and Shannon Casas, Business Journalist, and Editor-in-Chief of the *Gainesville Times* respectively. [15][16]. We established that qualities of fake news can be found based on the some key factors found on proven fake news posts or articles. Some of these key factors are if there are no authors associated with the article, if the article are from sources that are deemed to have a reputation of being strongly biased or false in the stories that they spread, if the Article is "too good" or "too bad" to be true, if the article does not have an editorial process before publishing a story, and if there is a one sided source with articles not having multiple sources used to show the story at hand.

From this understanding gathered, it was determined necessary to focus on analyzing the article's headlines and title. There were considerations to take into account, such as sources known to put out fake news articles. After some discussion, it was decided that it would create significant biases from the engineers apply to the machine learning models, additionally losing the effect of the function and definition of machine learning by predetermined add those labels for articles associated with certain sources. Therefore, to fairly judge and assess for fake news, the models used will only focus on those two aforementioned features alone.

3.1. Dataset

This project incorporated *Fake News Corpus*, *All The News* as the data set used in training, and the dataset, *FakeNewsNet* for a validation set. Fake News Corpus is an open source repository that contains over nine million news articles of varying "fake" types, All The News is a data set hosted on Kaggle that hosts over one-hundred thousand articles from legitimate news sources like CNN, and the

Specifications	Details
Processor	4.3 GHz Intel I7-8700k
Memory	32GB DDR4-3200MHz
Graphics	GTX 1070
Operating System	Windows 10
Programming Language	Python
Tools	Tensorflow version 2.0.0a

TABLE 1. METHODS USED

New York Times, and FakeNewsNet is an open source fake news repository that gets its content from *Politifact* and *Gossipcop*.

The article contents as well as the type of news it was (real or fake) was stripped out of the datasets for training and validation.

From the news types enlisted on the Fake News Corpus, it allows us to be able to validate and verify for fake news against the results deemed to be fake news through the machine learning models. Ideally, the model used should be efficient at discriminating against fake news sources specifically.

3.2. Machine Learning Models

The primary model used was an Artificial Neural Network (ANN). ANNs are considered to be loosely based upon how the brain functions. Similar to how neurons transmit messages within our brains, the ANN model contains 'nodes' which pass along data after multiplying them by 'weights', which are the values the model must adjust to get an accurate prediction. In each node, a function must be used to account for non-linearity in the data; this function is referred to as an activation function as it gives the probability a node/neuron will 'activate', or fire off.

With regards to predicting fake news, the only feature used is the content; other features such as the date were not used, because they were deemed to have little to no effect on prediction accuracy. The activation functions used include ReLU for the hidden layers, and sigmoid for the output layer.

4. Experiments and Results

4.1. Methods Used

The hardware used in the implementation of this project was an Intel I7-8700k with 4.2GHz, 32Gb RAM, a GTX 1070 video card running on Windows 10. The project was written using Tensorflow version 2.0.0a in Python.

4.2. Data

The training data was extracted from the Fake News Corpus and All The News datasets. 74,000 rows of the Fake News Corpus and 53,291 rows from the All The News were used. These two datasets were cleaned and then combined and shuffled before being vectorized using *Scikit-learn's TfidfVectorizer* so that it could be run through this model.

The evaluation data was first scraped from the news websites it was hosted on using FakeNewsNet's web scraping script. It was then taken from the json files the script saved in and processed, with the real and fake news articles being shuffled randomly. It was then processed to match the other two datasets.

These 2 data files were then saved to files, along with the trained vectorizer, using joblib so that they could be loaded easily when training, without having to reimport and format all the data again.

4.3. Model

The model for this research project was built using an ANN. It consisted of three dense layers, the first 2 layers use RELU as their activation function and have 6 nodes, while the third and final layer uses sigmoid and contains 1 node. An l2 regularization of .01 was used on both the input and middle layers.

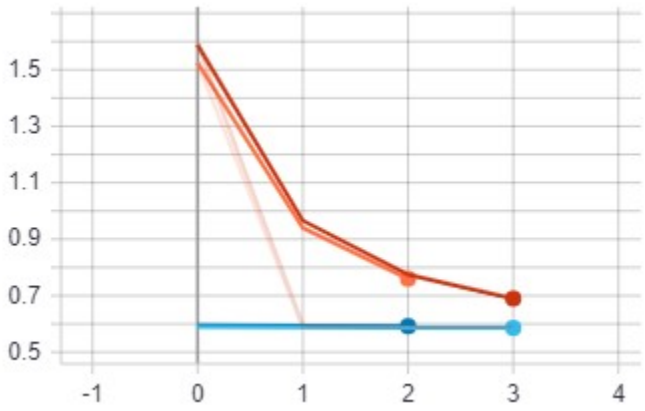

Figure 1. Loss per epoch with 3 epochs vs 4 epochs.

The optimization function used was adam and the loss was computed using binary cross-entropy. The model was fit in 3 epochs with a batch size of 256 samples. 3 epochs were used here because of the size of the dataset. Within the first few thousand samples being run through the dataset the loss was already nearing 1, and by the end of the first epoch it had hit 1. More than 3 epochs had diminishing returns on the models loss and it was decided that to prevent overfitting only 3 epochs would be run.

4.4. Results

TABLE 2. RESULTS

Model	Training Accuracy	Validation Accuracy	Loss	Testing Accuracy
ANN	73%	72%	0.59	60%

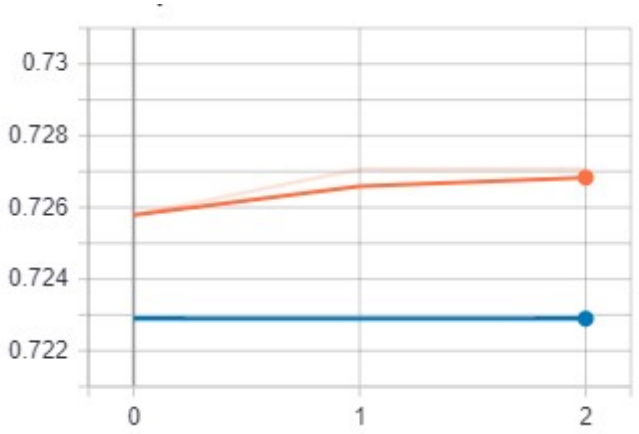

Figure 2. Average accuracy per epoch of final model. (Orange is training, blue is validation)

In training, the model had an accuracy of roughly 72% and a final loss of 0.59. The model was able to predict whether an article was fake news 60% of the time using the FakeNewsNet Politifact dataset. This dataset was chosen because it was formatted differently than the Fake News Net dataset. A benefit of the Politifact portion of the dataset is that it had an almost equal amount of real vs vake news samples, This ensured that the model was not biased towards real or fake in testing.

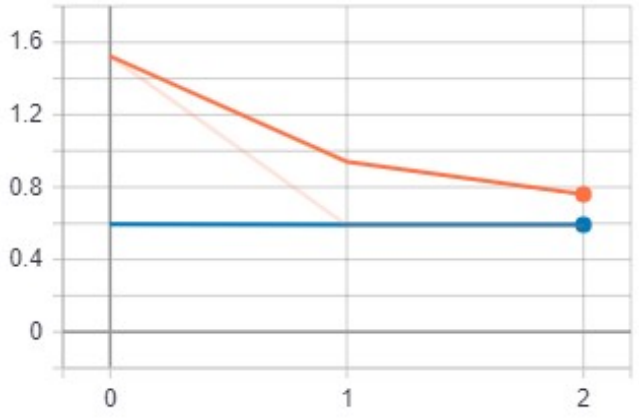

Figure 3. Average loss per epoch. (Orange is training, blue is validation)

5. Conclusions and Future Works

It is certainly possible to classify news content into two types: fake news and not fake news, however there will always be an inherent bias to this classification based on the researcher's own personal beliefs. Even though this is true, with tools like this research it could be possible to at least cut down on the amount of objectively fake news that exists in the world today. With a preliminary result of 60% this project could potentially contribute to accurately finding fake news and publicizing it, without the need for humans to have to do that work themselves.

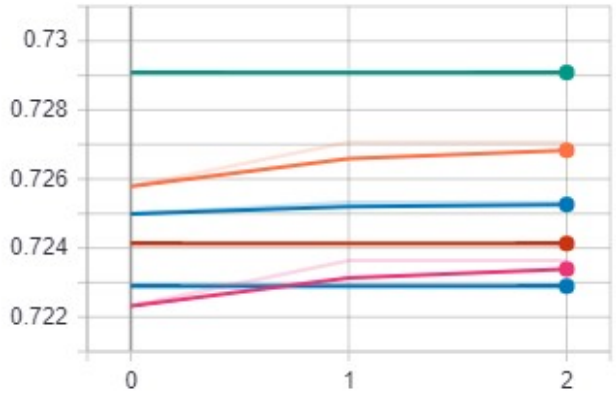

Figure 4. Average accuracy per epoch of multiple training runs.

The preliminary 60% accuracy is still not ideal, however with more time to tune hyperparameters it is possible to greatly improve this score. If this is done this model could potentially be able to accurately identify fake news with very little doubt.

The group had intended to implement more models to evaluate for fake news detection.Implementing an recurrent neural network for the project would have been of interest as it would have given better accuracy and may help with dealing with any fitting issues found in the data. Lastly, there were intents of testing other datasets to ensure that there is no underfitting or overfitting in any of the models used, and if there was an need to make changes to deal with it.

6. Acknowledgement

The research team would like to thank Kelsey Richardson and Shannon Casas of the *Gainesville Times* for providing advice on variables that could help determine whether a news source has aspects of "fake news" or not. As well as Dr. Mohammed Aledhari who is a Professor at Kennesaw State University for College of Software Engineering in advising us in our research.

References

[1] M. Szpakowski, "Fake news corpus," 2018, data retrieved from https://github.com/several27/FakeNewsCorpus.

[2] M. Haag and M. Salam, "Gunman in 'pizzagate' shooting is sentenced to 4 years in prison," Jun 2017. [Online]. Available: https://www.nytimes.com/2017/06/22/us/pizzagate-attack-sentence.html

[3] H. Allcott and M. Gentzkow, "Social media and fake news in the 2016 election," 2017.

[4] K. Jarmul, "Detecting fake news with scikit-learn," *DataCamp Community*, 2017. [Online]. Available: https://www.datacamp.com/community/tutorials/scikit-learn-fake-news

[5] S. Abeywardana, "Fake news classifier (using lstms)," Sep 2017. [Online]. Available: https://towardsdatascience.com/fake-news-classifier-e061b339ad6c

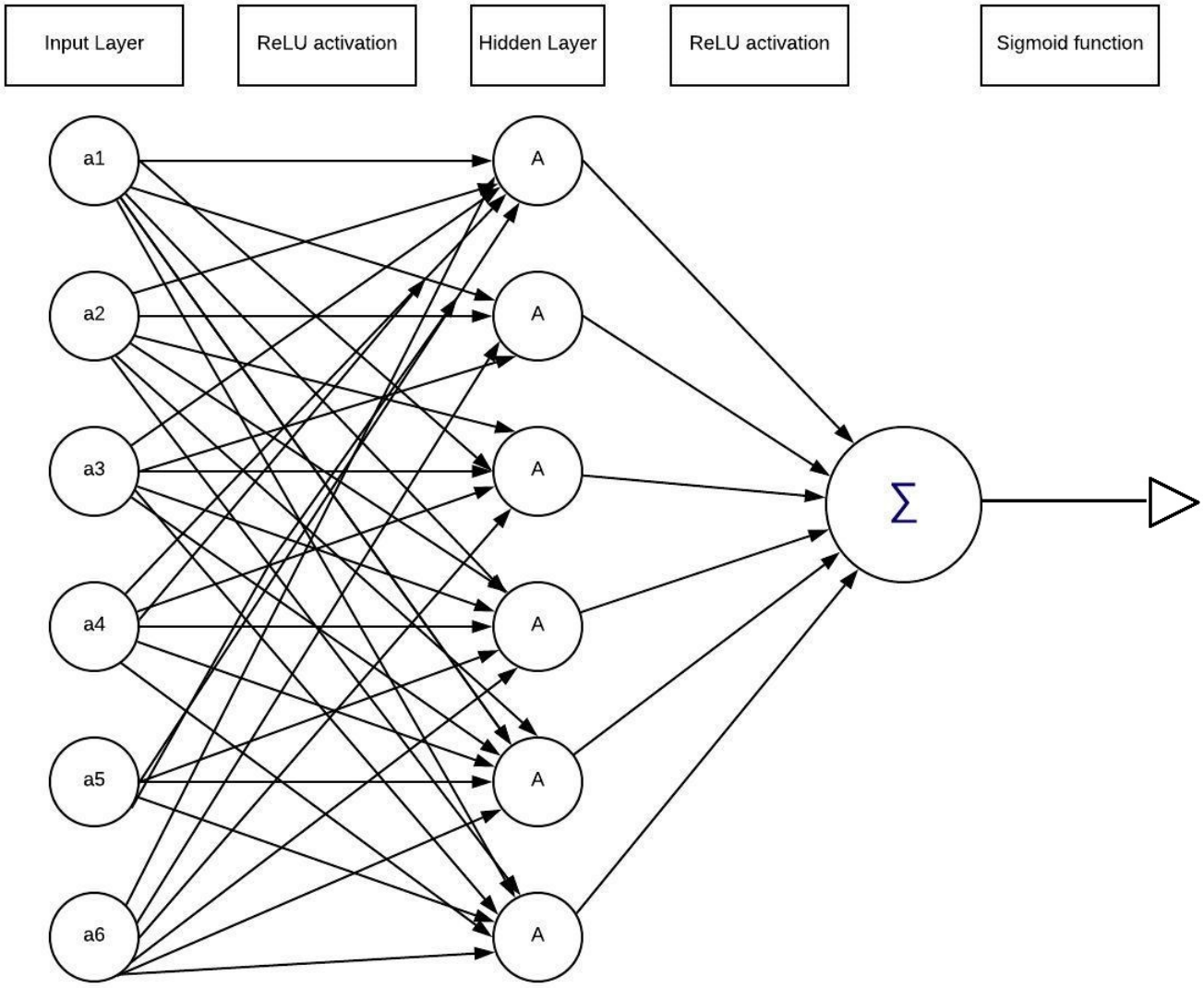

Figure 5. Diagram of the Neural Network with ReLU layers and sigmoid layer.

[6] Z. Tufekci, "It's the (democracy-poisoning) golden age of free speech," Feb 2018. [Online]. Available: https://www.wired.com/story/free-speech-issue-tech-turmoil-new-censorship/

[7] S. Shane and S. Frenkel, "Russian 2016 influence operation targeted african-americans on social media," Dec 2018. [Online]. Available: https://www.nytimes.com/2018/12/17/us/politics/russia-2016-influence-campaign.html

[8] "Why alex jones was banned from apple, facebook, youtube." [Online]. Available: https://www.politifact.com/truth-o-meter/article/2018/aug/07/why-infowars-alex-jones-was-banned-apple-facebook-/

[9] B. Molina, "Report: Fake election news performed better than real news on facebook," Nov 2016. [Online]. Available: https://www.usatoday.com/story/tech/news/2016/11/17/report-fake-election-news-performed-better-than-real-news-facebook/94028370/

[10] A. Thota, "Fake news detection: A deep learning approach," 2018. [Online]. Available: https://scholar.smu.edu/datasciencereview/vol1/iss3/10

[11] M. Rao, "Exploiting network structure to detect fake news," 2018. [Online]. Available: http://cs229.stanford.edu/proj2018/report/125.pdf

[12] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, "Fakenewsnet: A data repository with news content, social context and dynamic information for studying fake news on social media," *arXiv preprint arXiv:1809.01286*, 2018.

[13] 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.

[14] K. Shu, S. Wang, and H. Liu, "Exploiting tri-relationship for fake news detection," *arXiv preprint arXiv:1712.07709*, 2017.

[15] D. Lee and K. Richardson, Feb 2019.

[16] D. Lee and S. Casas, Feb 2019.