# Fake or Real:

AN ALGORITHM TO FIND FAKE NEWS

Jessica Everett Bellevue University October 24, 2020



- Confirmation bias
- Pizzagate
- Can fake news be stopped?





# • Fake and Real News Dataset (Bisaillon, 2020)

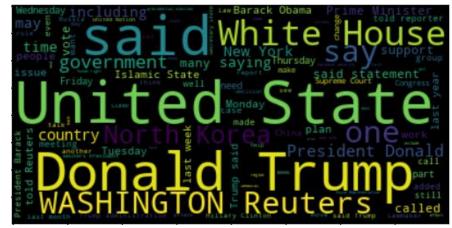
- 2 data sets
  - Fake News23,502 lines
  - Real News21,417 lines
- 4 columns
  - Title
  - o Text
  - Subject
  - Date



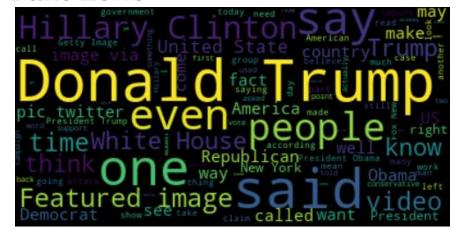
Methodology

# Data Exploration

o Real news

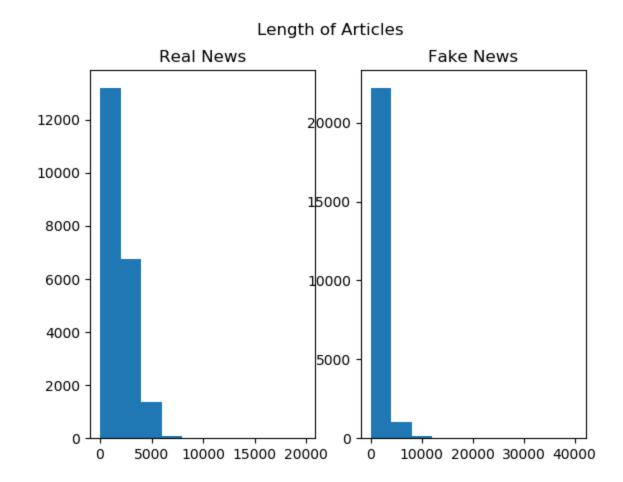


Fake news



#### Methodology

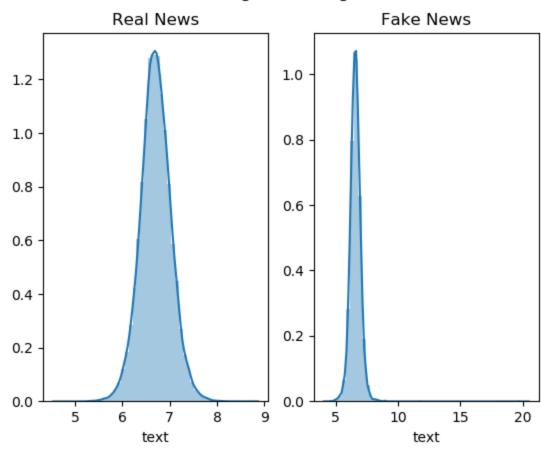
#### Data Exploration



#### Methodology

## Data Exploration

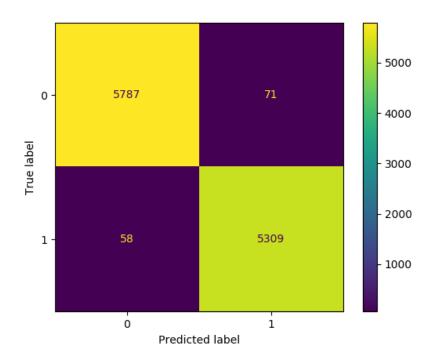






#### Models tested

- o NLP
- Logistic regression
  - × 98.85% Accuracy





- Social media
- Protect elections
- Protect individuals and groups



## References



- Ahmed H, Traore I, Saad S. (2017) "Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques. In: Traore I., Woungang I., Awad A. (eds) Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments. ISDDC 2017. Lecture Notes in Computer Science, vol 10618. Springer, Cham (pp. 127- 138).
- Ahmed H, Traore I, Saad S. (2018) "Detecting opinion spams and fake news using text classification", Journal of Security and Privacy, Volume 1, Issue 1, Wiley, January/February 2018. 2.
- Akpan, N. (2016). The very real consequences of fake news stories and why your brain can't ignore them. Retrieved October 3, 2020, from <a href="https://www.pbs.org/newshour/science/real-consequences-fakenews-stories-brain-cant-ignore">https://www.pbs.org/newshour/science/real-consequences-fakenews-stories-brain-cant-ignore</a>
- Bisaillon, C. (2020). Fake and Real News Dataset. Retrieved October 3, 2020, from <a href="https://www.kaggle.com/clmentbisaillon/fake-and-real-news-dataset">https://www.kaggle.com/clmentbisaillon/fake-and-real-news-dataset</a>
- Casad, B. J. (n.d.) Confirmation bias. Retrieved October 18, 2020, from <a href="https://www.britannica.com/science/confirmation-bias">https://www.britannica.com/science/confirmation-bias</a>
- "Fake News," Lies and Propaganda: How to Sort Fact from Fiction. (2020). Retrieved October 3, 2020, from <a href="https://guides.lib.umich.edu/fakenews">https://guides.lib.umich.edu/fakenews</a>
- Hand, J. (2020). 'Fake news' laws, privacy & free speech on trial: Government overreach in the infodemic?. Retrieved October 3, 2020, from <a href="https://firstdraftnews.org/latest/fake-news-lawsprivacy-free-speech-on-trial-government-overreach-in-the-infodemic/">https://firstdraftnews.org/latest/fake-news-lawsprivacy-free-speech-on-trial-government-overreach-in-the-infodemic/</a>
- How to Identify Fake News in 10 Steps. (n.d.). Retrieved October 3, 2020, from <a href="http://library.pfeiffer.edu/Fake-News-Worksheet.pdf">http://library.pfeiffer.edu/Fake-News-Worksheet.pdf</a>
- ISOT Fake News Dataset. (n.d.). Retrieved October 3, 2020, from <a href="https://www.uvic.ca/engineering/ece/isot/assets/docs/ISOT">https://www.uvic.ca/engineering/ece/isot/assets/docs/ISOT</a> Fake News Dataset ReadMe.pdf
- Kumar, V.P. (2020). Basic Text cleaning, WordCloud and N-gram analysis. Retrieved October 10, 2020 from <a href="https://www.kaggle.com/madz2000/nlp-using-glove-embeddings-99-87-accuracy">https://www.kaggle.com/madz2000/nlp-using-glove-embeddings-99-87-accuracy</a>
- Madz2000. (2020). NLP using GloVe Embeddings(99.87% Accuracy). Retrieved October 10, 2020 from <a href="https://www.kaggle.com/madz2000/nlp-using-glove-embeddings-99-87-accuracy">https://www.kaggle.com/madz2000/nlp-using-glove-embeddings-99-87-accuracy</a>
- Shane, T. & Noel, P. (n.d.). Data deficits: why we need to monitor the demand and supply of information in real time. Retrieved October 3, 2020, from <a href="https://firstdraftnews.org/long-form-article/datadeficits/">https://firstdraftnews.org/long-form-article/datadeficits/</a>