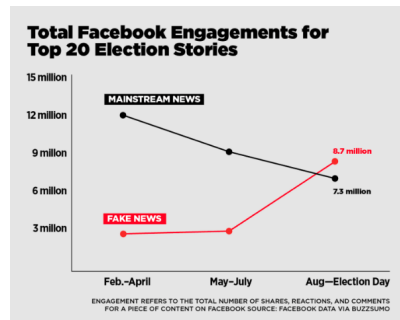




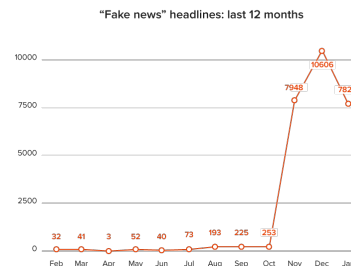
A Bayesian Approach to Fake News

1 Introduction and Background

"Fake news" became an increasingly coined term and was almost non-existent in the general context and media providers prior to October 2016 of the Presidential Election. In late 2016, the top 20 fake news stories on Facebook were reported to outperform the top 20 real news stories, which was determined by the number of comments, reactions, and shares. With websites such as Facebook who provide ad revenue to sites that host news stories, it should be the responsibility of these sites to filter out false information in order to correctly inform and educate our society, especially when it comes to presidential elections.



(a) Signal Media



(b) BuzzFeed News

Fake news posts have exploited the feeds of Facebook users'. Due to this problem, the data science community has looked to respond to this problem through a Kaggle competition called the Fake News Challenge ¹.

1.1 The Data

1.2 Methods

We performed supervised machine learning techniques such as naive bayes, count vectors, TF-IDF and SVM. With respect to language processing, tokenization, filtering, lemmatization, and stemming were all used for preprocessing the reviews. For our evaluation process, we used a confusion matrix to understand the accuracy of our models.

¹<http://www.fakenewschallenge.org/>

2. Exploratory Analysis

3 Supervised Analysis

3.1 Naive Bayes

3.2 Multinomial Naive Bayes

3.2.1 Count Vectorization

3.2.2 TF-IDF

3.2 SVM

3.3 Random Forest

3.4 Neural Network

4 Discussion

References

- [1] Russel and Norvig. *Artificial Intelligence: A Modern Approach*, 3rd Ed..
- [2] *This Analysis Shows How Viral Fake Election News Stories Outperformed Real News on Facebook*. [BuzzFeed News, November 2016]. https://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook?utm_term=.tnpJN34BM#.yfv6V10ZX.
- [3] *12 Months of Fake News Headlines, Dissected with Media Monitoring*. <https://signalmedia.co/media-monitoring-blog/fake-news-dissected-media-monitoring/>.