Fake News During the 2016 Presidential Election

Understanding the factors that contribute to the pervasiveness of fake news

By Phillip Chau

With the increasing role of social media platforms such as Twitter in the political landscape, recent political and social events such as the 2016 Presidential Election have seen an ever increasing presence of “fake news”.

Typically, the goal of fake news is to spread fabricated and deceptive content with the intent of bolstering public opinion, and especially during the 2016 Presidential Election, Twitter appeared to be a popular medium to propel “fake news” by both candidates. However, while we can easily note that “fake news” left its mark during the election, questions such as *how does fake news propagate so quickly* and *why does fake news have such widespread impact* remain unclear. Thus, through this project, I hope to use data visualization, natural language processing and data modeling in order to answer some of these questions and **gain a better understanding on the factors that contribute to the pervasiveness of fake news**.

Data Sources

The dataset utilized for this project was created by Julio Amandor, Axel Oehmichen and Miguel Molina-Solana of Imperial College London and contains a collection of over 9000 tweets related to the 2016 US Election. One important thing to keep in mind is that the authors gathered two groups of communications scholars in order to classify each tweet as “fake” or “real” news (there are two columns that each contain booleans that denote whether the group concluded that the tweet is fake or real). Ideally, we would want to categorize a tweet as “fake news” only if both groups agreed it was fake. However, this method resulted in an extremely insignificant amount of fake news compared to real news and as a result, the approach I took to create reasonably sized fake and real news dataframes was simply consider a tweet as “fake news” if at least one of the groups agreed it was fake. Thus, if both did not agree the tweet is fake, the tweet is considered real news.

In the end, I ended up with 1975 fake news tweets and 6445 real news tweets. While there is still some class imbalance, this was the most optimized configuration possible and I at least had a sizable amount of fake news to work with.

Data Cleaning and Engineering

Before working with the data, I had to do some basic text cleaning in order to prevent trivial tokens from skewing my results. This process involved removing punctuation, numbers and stop words (common words in natural language) and lowercasing text. Furthermore, the authors stated that they created this dataset by querying the Twitter API for tweets with the hashtags **#MyVote2016**, **#ElectionDay**, **#electionnight** and user handles **@realDonaldTrump** and **@HillaryClinton**. Obviously, these words will be most common amongst all tweets in both real and fake dataframes. Thus, it was necessary to remove these specific strings from the dataset in order to prevent them from overpowering the counts of other words. Using the common NLP library TextBlob, I created columns that rank the polarity of the tweets on a scale of -1 to 1 (1 being the most positive tweets and -1 being most negative) and rank the subjectivity of the tweets on a scale of 0 to 1 (0 being most objective).

Finally, to define pervasiveness, I decided to use the median retweet count of the cleaned fake news dataframe. Retweets quantify the spread of a post on Twitter and perfectly allow me to capture the notion of the popularity of a tweet. I chose the median over the mean since there's a chance that a single tweet could heavily skew the results, and in the end, my threshold for success became any tweet with over 2000 retweets. Thus all the data below only includes "pervasive/successful news" based on my metric.

Exploratory Analysis

Who are the users with the most fake tweets?

	user_screen_name	post_count	retweet_count	user_followers_count
0	realDonaldTrump	62	5977.419355	1.749375e+07
1	JamesOKeefeIII	30	5798.733333	2.193511e+05
2	FoxNews	19	2987.000000	1.196921e+07
3	DineshDSouza	17	2611.352941	4.127712e+05
4	SheriffClarke	16	5430.000000	4.220609e+05
5	PrisonPlanet	16	3357.187500	3.367709e+05
6	DanScavino	16	3419.437500	2.299974e+05
7	Fahrenthold	15	5321.600000	2.289740e+05
8	elizabethforma	12	8171.166667	1.023658e+06
9	KellyannePolls	12	5035.000000	5.655594e+05

Figure 1. Top 10 Users with Fake News Tweets

For the most part, the top 10 users appear to be prominent figures in the political landscape. Not surprisingly, **Donald Trump** ranks first with the most amount of fake news. Following him are **James O' Keefe**, who is a conservative political activist, **Fox News**, which is a right leaning news outlet that was heavily favored by Trump, **Dinesh D Souza**, a right wing political commentator and conspiracy theorist, and **David Clarke** (SheriffClarke), who is a former law enforcement official of Milwaukee County and is a heavy conservative, Trump supporter. Notice how even within the top 5, all the users are right leaning (with further research the remaining 5 are also right leaning for the most part). This makes us suspect that perhaps "fake news" was very popular amongst conservatives, who did appear more receptive of Trump's unprofound claims during the election and his presidency. One thing to note is that the users that post the most fake news also have a very large user follower base. The positive relationship between having a large follower count and posting more "fake news" makes sense as these users are leveraging their popularity on social media to reach out and influence more individuals, which contributes to the pervasiveness of fake news.

When do users post most often?

The next thing I was interested in about users that post fake news was when do they like to post? From the heatmap below, it appears that fake news is primarily posted in the beginning of the week such as Monday and Tuesday. While there was barely a difference in months, October and November seemed to be the most popular. These results also make sense, as in order to get the most coverage, users would want to post in the beginning of the week and let the tweet propagate throughout the rest of the week. November and October were the peak months of the election season, which would likely drive up fake news as candidates race to bolster their images.

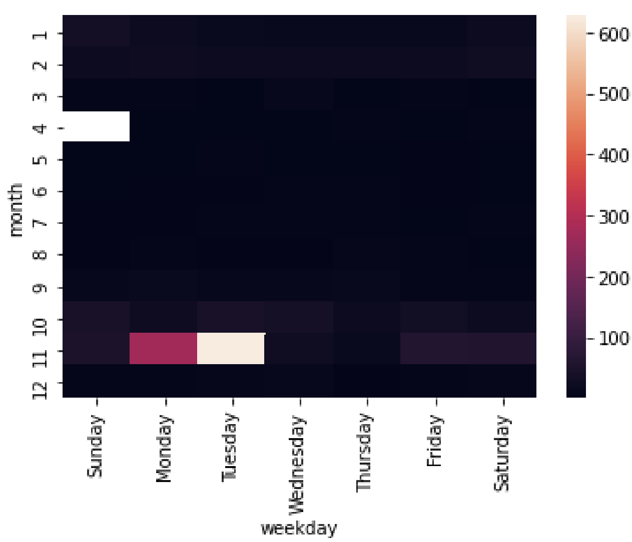

Figure 2. Heatmap of Fake News Counts against when its posted

What kind of content defines successful fake news?

Now that we know a bit about the users behind pervasive fake news, I wanted to understand what factors about fake news drive its success. To do this, I first created a word cloud of the cleaned fake news dataframe to get a better idea of what kind of words I should expect to work with.

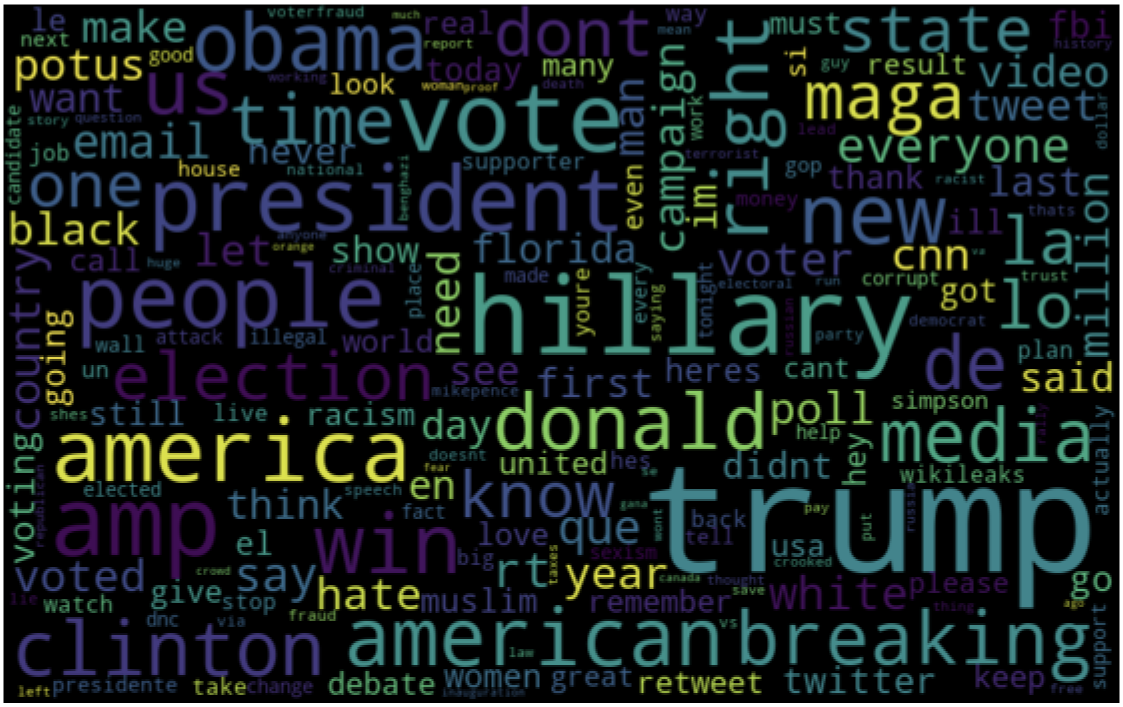

Figure 3. Word Cloud of Fake News Tweets

As seen above, some of the most prominent words include “trump”, “hillary”, “obama”, who were all relevant figures during the Presidential Election. From here, it would seem that mentioning a prominent political figure is involved in “successful fake news”. However, to get a better understanding of what successful fake news tweets are saying, I decided to do topic modeling using Latent Dirichlet Allocation (LDA) in order to create topics with the vocabulary derived from all possible fake news tweets.

As seen by the chart below, I created 10 different topics and took the top 10 most probable words within each topic. It appears that of the 10 different topics, some obvious categories include the following: 1) Attacks on Hillary Clinton (Topic 3, 6, 7). 2) News on the Hillary Clinton Email Controversy. 3) Trump’s attack on news platforms (Topic 2). Other pervasive fake news topics involve mentioning some unprofound fact about race such as “muslim” in topic 1 and “black” in topic 2 or about voting in states like Florida (Topic 10). In general, successful fake news appears to often involve polarizing opinions and blatant attacks on given groups.

Topics
Topic 1: muslim elected media stop american said twitter americans potus people
Topic 2: new got great black let cnn trump dont like amp
Topic 3: mikepence know shes question lives criminal fbi clinton hillary media
Topic 4: hace god truth climate start tweet los change want simpsons
Topic 5: time usa term dnc state woman trumps news country breaking
Topic 6: voted clinton hate obama donald hillary vote president america trump
Topic 7: fear different didnt video today sexism emails day racism campaign
Topic 8: inauguration actually thats real trying guy new house white amp
Topic 9: tell times maga debate paypal thank rts ill hillary wins
Topic 10: voting states history going florida united que man time right

Figure 4. Chart of Top 10 Topics

In order to get a better understanding of what exactly these fake news tweets look like, I've included some of the raw tweets below.

Damira Maricic
@DamiraMaricic

Dear @Starbucks u would rather hire muslim refugees then our #veterans? 62 million of us voted 4 @realDonaldTrump bcuz WE SUPPORT #USA 🇺🇸

7:43 AM · Feb 12, 2017 · Twitter for iPad

Figure 5. Fake News Tweet 1

David A. Clarke, Jr.
@SheriffClarke

Dems blasted @realDonaldTrump for calling ghettos crime infested hell holes. He cared more about blacks than they do hotair.com/archives/2017/...

12:41 AM · Feb 1, 2017 · Twitter for iPad

Figure 6. Fake News Tweet 2.

What Distinguishes Successful Fake News from Successful Real News?
The final question of interest is determining what attributes of pervasive fake news distinguish it from all other news. Fake news obviously seems to have unique features to it that make it very popular amongst certain populations, but there's a possibility that the features that drive successful fake news are different than those that drive real news. In order to rigorously answer this question, I conducted a 2 sample t-test, which is a type of hypothesis test that allows me to determine whether the difference between the means of two groups is statistically significant or not. Testing all possible features at an alpha level of 0.05 (allowing for 5% risk of incorrectly rejecting the null hypothesis), I noticed that it appears on average that fake news is significantly shorter than real news. At the same time, while both successful real and fake news tend to be neutral sentiment, real news is significantly more positive than fake news. Finally, on average, posters of fake news tend to have a much higher user friend count than those with real news and tend to post tweets later in the year than those with real news. Having a higher friend count often means that a user is able to extend their message to a larger network, which reaffirms the notion that fake news is able to successfully propagate by a user having an extensive social network.

	month	polarity	text_length	user_friend_count
fake	8.864810	0.033487	115.566582	10970.822785
real	7.331575	0.088493	118.777347	6736.276338

Figure 7. Significant Results of T-test

Given that we defined "pervasiveness" as tweets that have a high retweet count, I was now curious to see if the polarity of a fake news and real news tweet impacted its success.

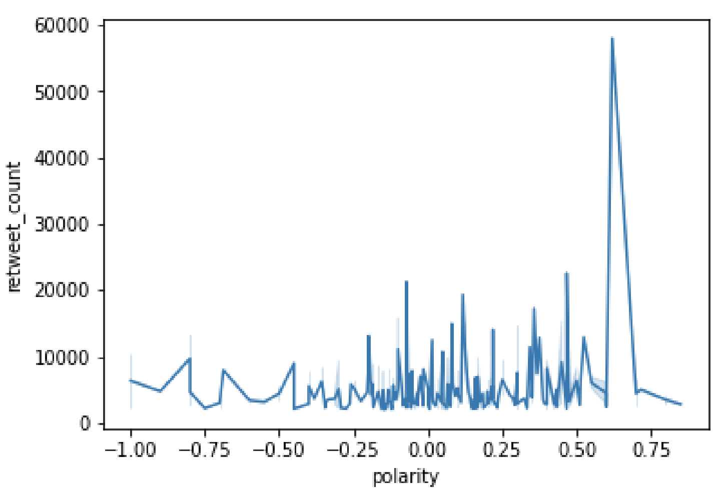

Figure 8. Polarity vs. Retweet Count of Fake News

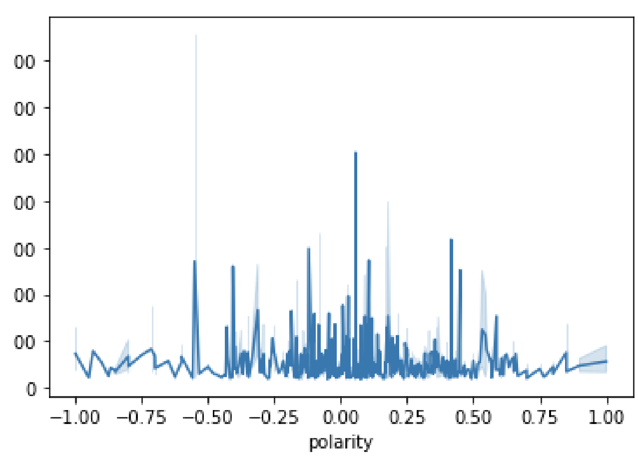

Figure 9. Polarity vs Retweet Count of Real News

Plotting the polarity scores against the retweet counts of successful fake and real news tweets above, we see that as the polarity score of fake news increases, the retweet count increases as well. This result honestly came as a bit of a surprise to me, as I initially assumed that successful fake news consists of hateful messages and attacks, but it appears that "fake news" promoting and bolstering a candidate contributes more to the pervasiveness of fake news. Not too surprisingly, however, retweet count of real news appears to peak at 0 polarity (neutral sentiment) and dips down at the edges, which is a sign that these tweets compose of factual claims.

Data Pre-Processing

Now that we have a much stronger understanding of what fake news is and what makes it so pervasive, I want to make a classifier to determine fake vs. real news and to determine the most important features that contribute to this distinction.

TF-IDF

TF-IDF is a measure of how relevant a word is to a document in a collection of documents. In order to convert text into numerical features, I used **Tfidfvectorizer** to get a TF-IDF score for each word in the vocabulary derived by my tweets. Essentially, my features for my training and test data became all possible words from the tweet-derived vocabulary, with their values being their TF-IDF scores.

After creating my features, to create my labels, I simply denoted real news as 0 and fake news as 1 and then appended my two dataframes to create a mixed dataset of labeled real and fake news. Finally, I split the data into training and test sets and removed any words from the feature list that did not appear in at least 1% of the documents (this step was needed to reduce the amount of irrelevant features in my dataset).

Class Imbalance

As seen by the figure below, the amount of real news (0) is almost 4 times higher than the amount of fake news, which would mean our classifiers could get a very good result simply predicting "real news" each time. To mitigate this, I oversampled the fake news data in the training set through sklearn's RandomOverSampler library in order to make the 2 classes have equal amounts of data.

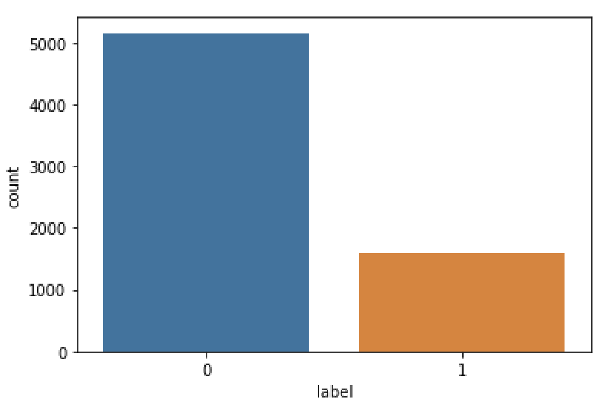

Figure 10. Counts of Real and Fake News

The final result was a training set with **10,304 tweets** and **101 features** and a testing set with **1684 tweets** and **101 features.**

Modeling

I fit 2 supervised classification models in order to classify fake vs real news. I specifically chose Naive Bayes and Random Forest since both of these are able to extract the most important features used to determine the classification, which would give me a better idea of what components of the tweet specifically lead to it being fake news.

Naive Bayes

This is a probabilistic classifier based on Bayes theorem that uses the naive assumption that all features are independent. This classifier allows me to compute the words that most likely belong to fake news compared to real news by calculating the frequency of the word across all tweets in the class and dividing it by the total number of tweets within the class. As seen below, I then took the top 10 words that had the highest fake/true ratio of their probabilities, which was simply an indication of the words that most likely would belong to a fake news tweet. Just like I mentioned before in the topic modeling, people such as “Trump”, “Obama” and “Hillary” and words like “media” appear to be highly favorable in fake news tweets. The popularity of these words could come from baseless attacks on “Obama” and “Hillary” or the “media” or unprofound claims about “Trump”, which we saw earlier.

token	true	fake	fake/true ratio
everyone	0.003761	0.017219	4.578324
breaking	0.007021	0.022751	3.240237
hillary	0.013917	0.041762	3.000843
media	0.008908	0.021519	2.415568
right	0.010925	0.025690	2.351404
que	0.005987	0.013306	2.222499
obama	0.010451	0.023068	2.207288
wins	0.007150	0.015086	2.109968
la	0.005894	0.012416	2.106334
trump	0.046417	0.086723	1.868350

Figure 11. Naive Bayes Classifier Results

Random Forest
The next model I trained was a random forest model. This model creates an ensemble of decision trees and outputs the mean result of the outcomes of these several decision trees in order to reduce overfitting. As seen by the feature importance plot below, just like with Naive Bayes, the most important features include “trump”, “hillary” and “maga”. Thus, from both models, it is quite apparent that mentioning important political figures at the time was a good indicator of fake news, which could make sense since fake news often involved attacking the opposing candidate or making false claims to boost the public image of a certain candidate.

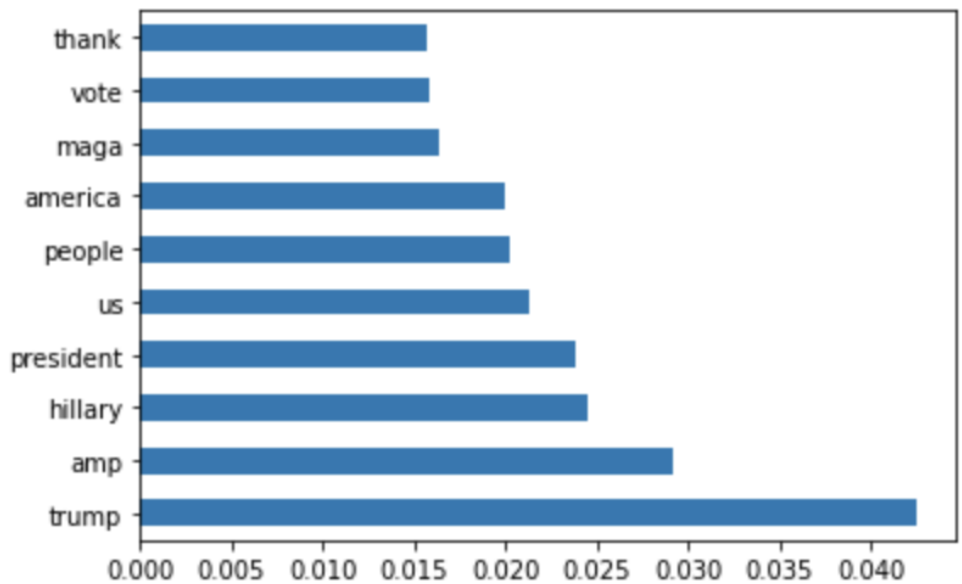

Figure 12. Most Important Features of Random Forest

Results

	Random Forest	Naive Bayes
Accuracy	59.3%	63.9%
False Positives	486	357
False Negatives	191	206
Mean Squared Error	0.63778	0.6008

As seen above, across all metrics, the Naive Bayes classifier outperformed Random Forest. Random forest tended to misclassify real news as fake more while Naive Bayes tended to misclassify fake news as real more. Although the random forest accuracy is currently low, keep in mind that no hyperparameter tuning was done prior. As this process takes an extensively long time, in the future I would like to increase the accuracy of my random forest model through GridSearch to find the optimal depth and number of estimators.

Findings and Conclusion

From our analysis, we first learned about the users behind fake news tweets and then took a deep dive into the factors that may contribute to the pervasiveness of fake news. Twitter users that posted the most fake news during the 2016 Presidential Election often were right-leaning, Trump supporters with an extensive follower count. Their large base popularity on social media perhaps was the engine that pushed their fake news through a wide social network, contributing to the pervasive nature of fake news and enabling them to continuously push fake news content. The concept of pervasiveness can be quantified in Twitter based on a high retweet count, and as a result, looking at the fake news tweets with a high amount of retweets, we noticed that these tweets often discussed attacks on Hillary Clinton or Obama or pushed unproven claims to bolster the public image of candidates like Trump. It would make sense that these types of tweets would find the most success, as based on our initial analysis, it seems that fake news posters are right-leaning figures who presumably have a large, right-leaning following. Thus, making right-wing appealing statements such as attacks on the left could be very popular amongst these users' follower base, driving the success of fake news. More specifically, successful fake news tweets often contained more polarizing sentiment in comparison to real news and the users of these fake news tweets often also had a significantly larger friend base than the average real news poster, which further contributes to the idea that "users are leveraging their popularity on social media to reach out and influence more individuals". Finally, extracting the most important features from our two classifiers, we reaffirmed the importance of making statements involving important political figures to generate well received fake news.

In conclusion, the success of fake news is determined not only by the content itself, but depends on the proper blend of sentiment, access to an extensive social network and catering to a target audience. With all of these in mind, we now understand more about why and who exactly propagated fake news during the 2016 Presidential Election.