Defending Against Fake News

An internet browser plugin to help in the fight against misleading information

Booz | Allen | Hamilton

PennState College of Engineering

Abstract

- Booz Allen Hamilton is an information technology consultancy that wants to help mitigate the rising fake news problem by offering a Google Chrome plugin that can inform readers about the validity of news.
- They requested Penn State to take their plugin and add a fresh approach to their existing methods.
- Accuracy can be added by implementing a machine learning algorithm based on the sentiment of an article which infers bias by judging overt positive or negative tones related to the article.
- The final product consists of a model which can take in an article from a webpage or news source and output a rating of fake or real. Adding this to the plugin can increase accuracy of the plugin's rating.

Motivation - Is this true?

- Social media and smartphones allow for fake news to be distributed quicker than ever before. It takes little effort to share an article or graphic that has no backing, but appears real.
- It takes facts over 6x longer to reach an audience of 1,500 people than sensationalized articles (Fox, 2008).
- By the time a fake article is circulated the damage is done. Retraction of articles gets little to no attention compared to the original news.
- Research and focus groups have helped identify good practices to validate a news source. This is illustrated below (Kiely, 2016).

Consider the source
- Reputation for serious and informative news
- Satirical or unreliable sources

Title of article
- Overt positive or negative can indicate bias
- Clickbait
- Unrelated to article content

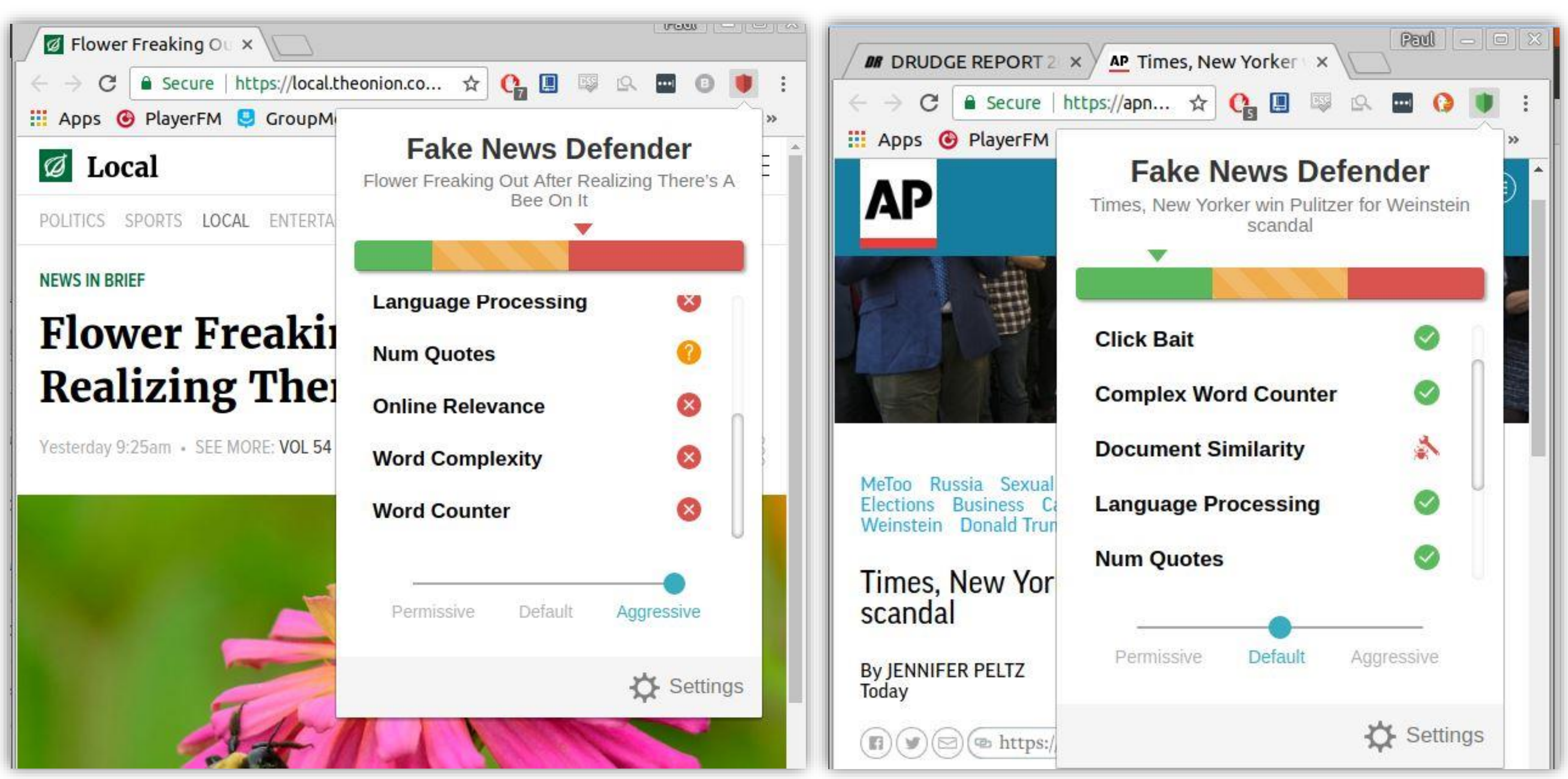
Author's track record
- Track record of good reporting

Check the date(s)
- Does timeframe make sense

Read beyond the headline
- Does body match title
- Is the content absurd or fantastical

The Booz Allen Hamilton Plugin

- Booz Allen began work on the project in 2017.
- The Penn State team received their work and then built their own plugin server, installed the plugin, and then proceeded to experiment with the existing code repository.



Instances of the plugin operating reading both fake (left) and true (right) articles

Objectives

- Create and test a fake news rating model which can be implemented in BAH plugin.
- Use features that are not already encompassed in the existing plugin.
- Find a more comprehensive data set to use for model training.

Methodology

- Sentiment indicates feelings in a certain direction of bias on the topic (positive or negative).
- Vader Sentiment analyzer used to gauge the sentiment and compound ranking of sentences.
- Sentiments of each sentence saved to "scores" vector = [Positive, Negative, Compound].
- The scores vector allows the following features to be generated to tune the model (right).
- Data consisted of 3,539 real articles and 3,646 fake (McIntire). This is illustrated below (left).

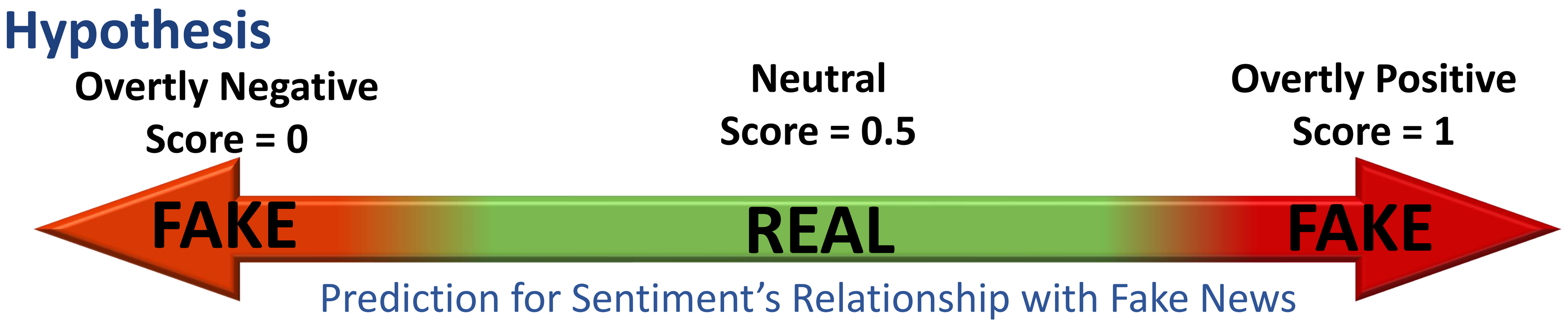
No.	Title	Article	Label
1	Slim Prospects for Ch	Politicians can't agree on	Real
...	...	..	...
3539	The NY Giants Pick S	The Giant's Head Coach A	Real
3540	Superbowl Cancled D	In a conference the comis	Fake
...	...	...	...
7818	Canada Declares Wa	The war between Mexico	Fake

Dataset Visualization

Feature	Description	Feature	Description
article_pos	whole article positivity	end_pos	positivity of last 5 sentences
article_neg	whole article negativity	end_neg	negativity of last 5 sentences
article_comp	whole article compound rank	end_comp	comound rank of last 5 sentences
start_pos	positivity of first 5 sentences	pos_var	variance of positivity in article
start_neg	negativity of first 5 sentences	neg_var	variance of negativity in article
start_comp	compound rank of first 5 sentences	comp_var	variance of compound rank in article
tot_letters	total letters in article	sent_count	Average sentences per paragraph
tot_words	total words in article	word_count	Average words per sentence
tot_sent	total sentences in article	letter_count	Average letters per word
tot_para	total paragraphs in article		

Features Available for Tuning Model

Hypothesis



Prediction for Sentiment's Relationship with Fake News

Implementing Sentiment Analysis

- NLP can be implemented to infer positive, negative, and various other "feelings" in text.
- Vader Sentiment Analyzer returns three rankings for each sentence which can then be manipulated. The most successful combination is shown below.

	pos_diff	neg_diff	comp_diff	start_pos	end_pos	end_neg	word_count	letter_count	label
0	-0.030951	0.061333	-0.122836	0.08925	0.0946	0.1544	14.896552	4.776235	FAKE
1	0.010590	0.037960	-0.114158	0.05225	0.0742	0.0846	17.840000	4.923767	FAKE

Example Feature Rankings of Most Accurate Model

Testing Algorithm

- Fitting the model one time is not sufficient to gain a confident estimate on accuracy.
- To gain the best accuracy learning was implemented 100 different times.
- This repetition confirms that the model did not simply get 'lucky' the first time it was ran.

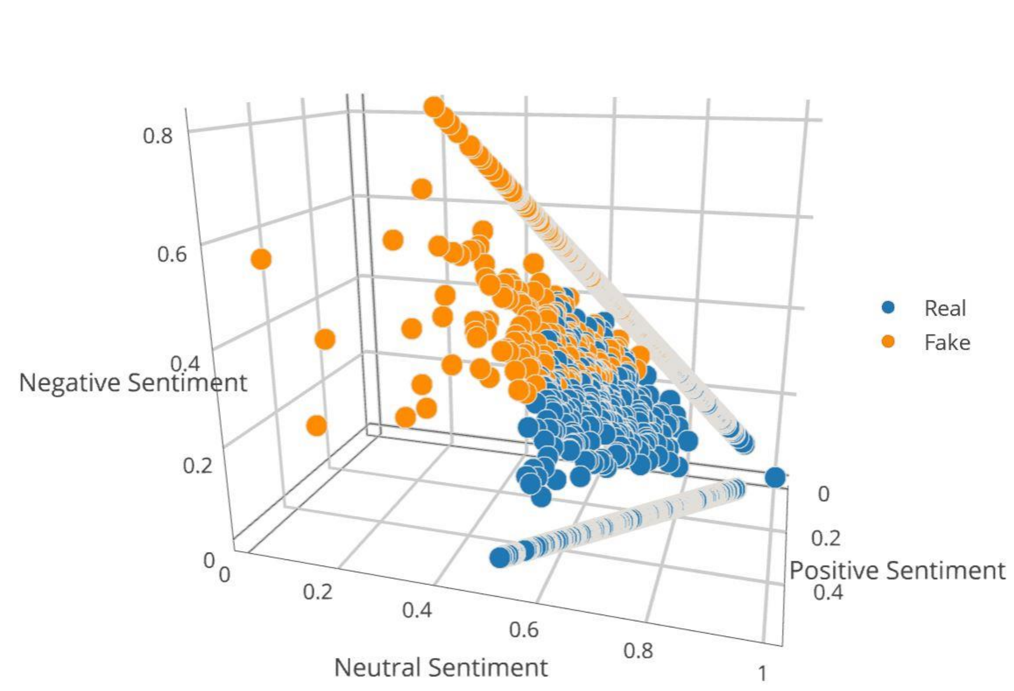
```
accuracys = []
gnb = GaussianNB()
for state in range(100):
    x_train, x_test, y_train, y_test = train_test_split(scores[:,:], label, test_size=0.2, random_state=state)
    gnb.fit(x_train, y_train)
    accuracys.append(accuracy_score(y_test, gnb.predict(x_test)))
print(np.var(accuracys))
print(np.mean(accuracys))

0.00470844444444
0.722666666667
```

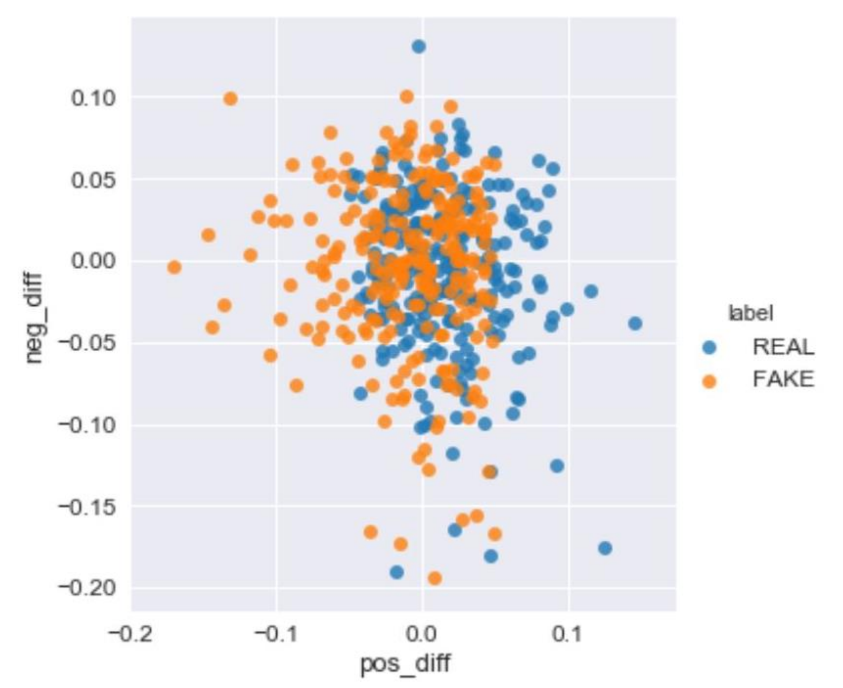
Script Used to Gauge Model's Accuracy with Output

Results

- The following graphs illustrate various features and how they relate to real vs fake news.



Real and Fake vs. Various Features



Real and Fake vs. Positive and Negative Sentiment

- The table below shows the top 3 models and the associated ratings that were used to train them

Accuracy	Features for Machine Learning Model						
72.20%	article_pos	article_neg	article_comp	letter_count	word_count	sent_count	para_count
71.38%	pos_var	neg_var	comp_var	letter_count	word_count	sent_count	para_count
70.43%	start_pos	start_neg	start_comp	end_pos	end_neg	end_comp	letter_count

Table of top 3 performing models with features included

Contributors

Team Members: Nick Baker, Paul Boehringer, Nina Brajovic, Peter Mathews, & Josh Norton

Mentors: Matt Kremer, Jooyong Lee, Ryan Fischbach, & Wahab Jilani

Faculty Advisor: Dr. Soundar Kumara

Acknowledgments

A special thanks to the project sponsor Booz Allen Hamilton for supporting this endeavor. Along with the Learning Factory for administrative support with the capstone process.

Citations

Fox, Maggie. *Want something to go viral? Make it fake news*. (2018). *NBC News*. Retrieved 18 April 2018, from https://www.nbcnews.com/health/health-news/fake-news-lies-spread-faster-social-media-truth-does-n854896

McIntire, G. "Georgemcintire/Fake_Real_News_Dataset". 2018. *Github*. Accessed April 18 2018. https://github.com/GeorgeMcIntire/fake_real_news

Kiely, Eugene, and Lori Robertson. 2016. "How To Spot Fake News - Factcheck.Org". Factcheck.Org. Accessed April 18 2018. https://www.factcheck.org/2016/11/how-to-spot-fake-news/.