Fake News Classifier

"Truthiness (noun) - the belief in what you feel to be true rather than what the facts will support"
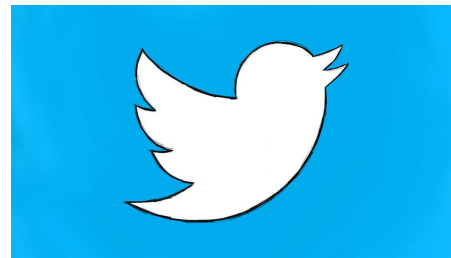
Stephen Colbert

Agenda

1. Overview
2. The Approach
3. Findings
4. Next Steps
5. Conclusion

Overview

- Twitter has a fake news problem.

- EMG Consulting has been retained to build a model to predict when an article may include misleading information to inform users when that is the case.

The Approach

- Model built based on roughly 38,500 articles from 2016-17 on US and world news, with over 9 million words in total — approx. 55/45 real/fake split

- Make predictions by assigning numerical values to words in each article, then training model to draw a distinction between real and fake articles based on these values

- Emphasis on creating a model that is accurate overall - costs associated with incorrect predictions in either direction!
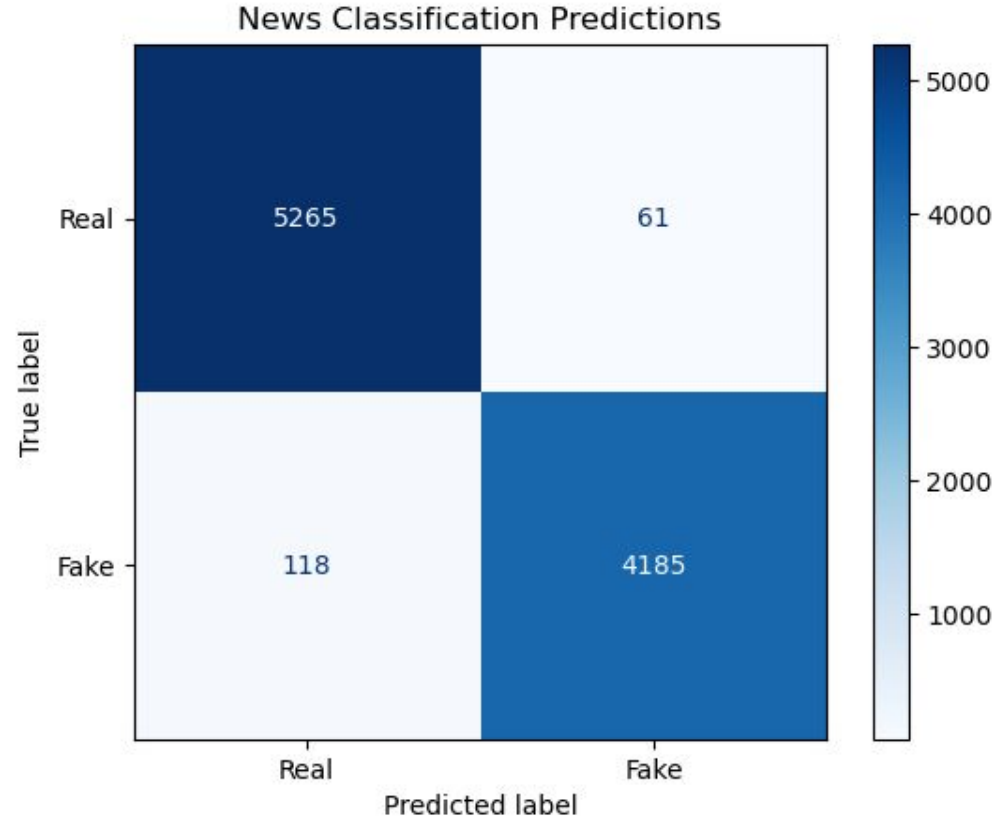
Finding #1: The model works - but we'll need to scale it.

The model is very accurate when making predictions on topics it is trained on.

- Final model was a Linear SVC.

- Roughly 98.1% accuracy on unseen data.

- However, narrow scope of training data means model would have to be scaled before deployment to public.

News Classification Predictions

Finding #2: When predicting "real" vs. "fake" news, pay attention to adjectives and adverbs - they can tip you off.

Some stories are easy to predict...

'Fully committed' NATO backs new U.S. approach on Afghanistan

Donald Trump Sends Out Embarrassing New Year's Eve Message; This is Disturbing
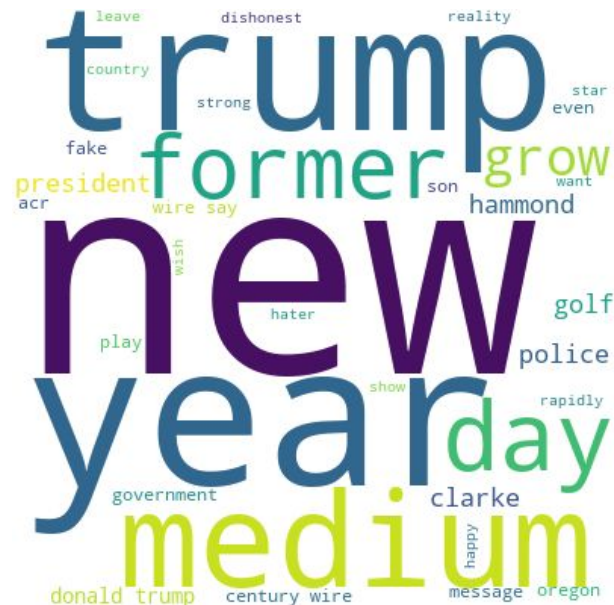
...some are not.

Potential Shift: Trump Warns Israel, 'New Settlements May Not Help Peace in Middle East'

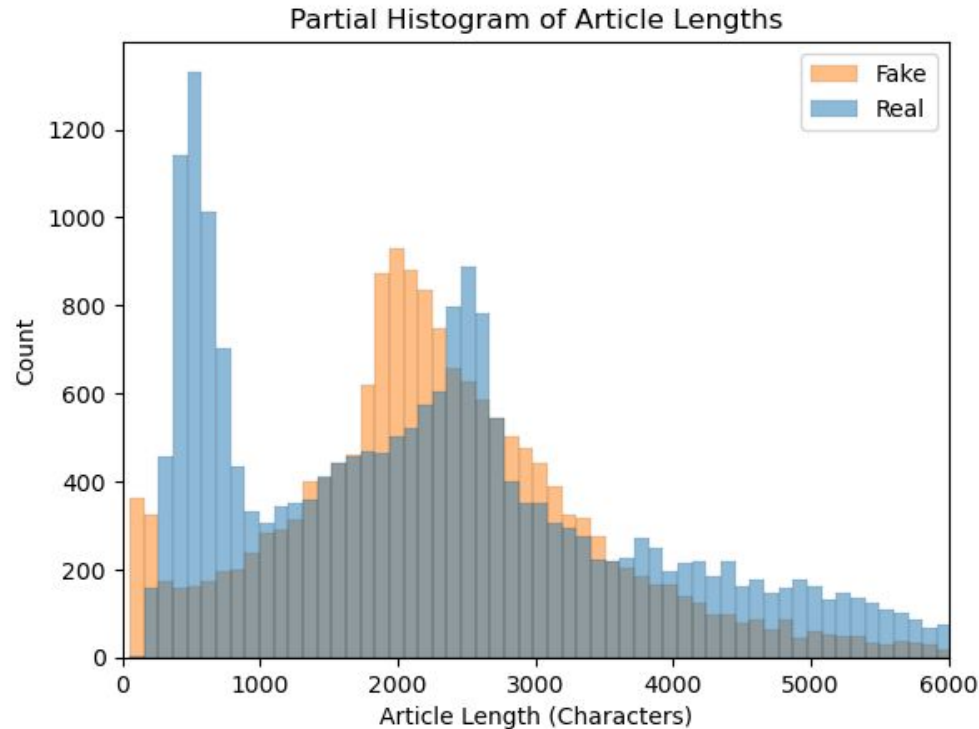**Trump hits back at Clinton, with a golf ball, on Twitter**

True positives
(fake news, predicted fake)

True negatives
(real news, predicted real)

Finding #3: Some fake news spreads in video form - we'll need a different approach to combat this.

Some "articles" in the data had no characters.

Partial Histogram of Article Lengths

Next Steps

- Massively increase the size of the training data, with articles from a wide array of organizations. Partner with impartial fact-checking organizations to ensure validity of labels in training data.

- Work with transcription services to quickly analyze and label misleading video content.

- Continue to refine preprocessing and modeling to further shrink number of incorrect predictions.

Thank you.

Contact:
emgerber94@gmail.com

GitHub Repository Link