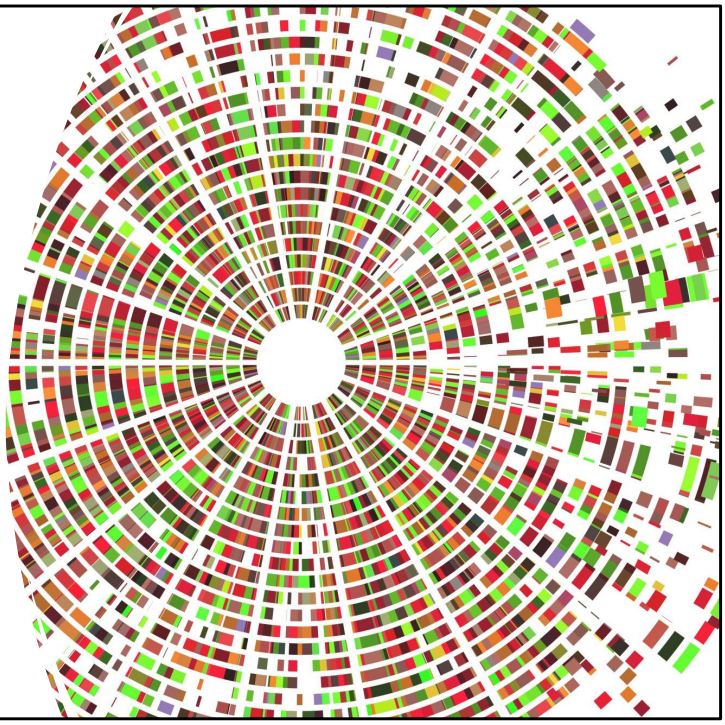
# The Rise of Fake News in the Political Scene

*By Michael Mahoney*



# Presentation Overview

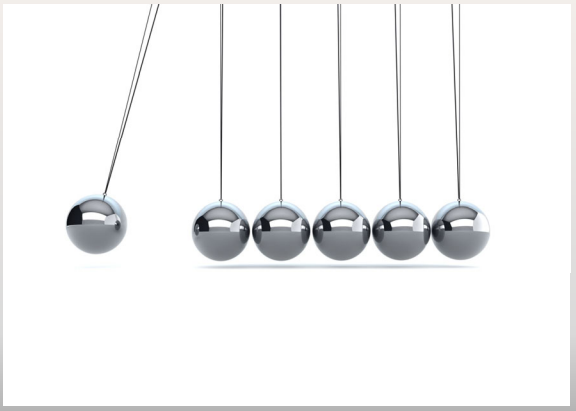
- Quick intro to the data
- Talk about misinformation and what we found
- Explore some recommendations
- Talk about the future


The short answer is everyone but some should be more tunned in than others.

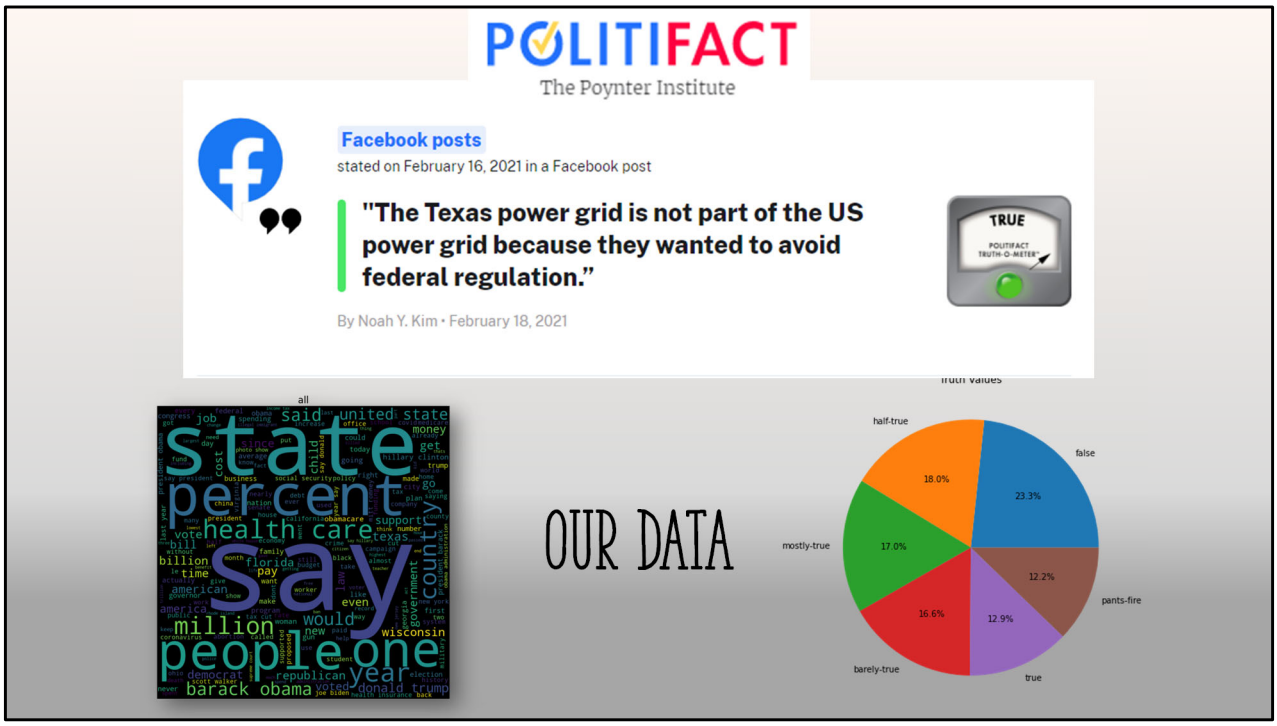
- Social Media Companies
- Journalists who work in the political realm
- Politicians and their staff

# Misinformation

- The Cost 79 Billion USD
  - 37 billion USD in losses on the stock market
  - 17 billion USD in losses annually from false/misleading financial information
  - 9.54 billion USD spent on reputation management for false information about companies.
  - 3 billion USD has been already spent to combat fake news.
  - 400 million USD spent by US public institutions defending against misinformation.
  - 250 USD million spent on brand safety measures.



In the study titled “The Economic Cost Of Bad Actors On The Internet: Fake News 2019, ” author Professor Roberto Cavazon from the University of Baltimore explores the cost of fake news from a high level perspective. The economic effects have hit every sector of the economy. As social media sites are the host for much of the content that circulates, they have already begun to invest in moderation to help fight mis-information. Facebook invested 3 billion USD in 2018 and has stated more investment is on the way.

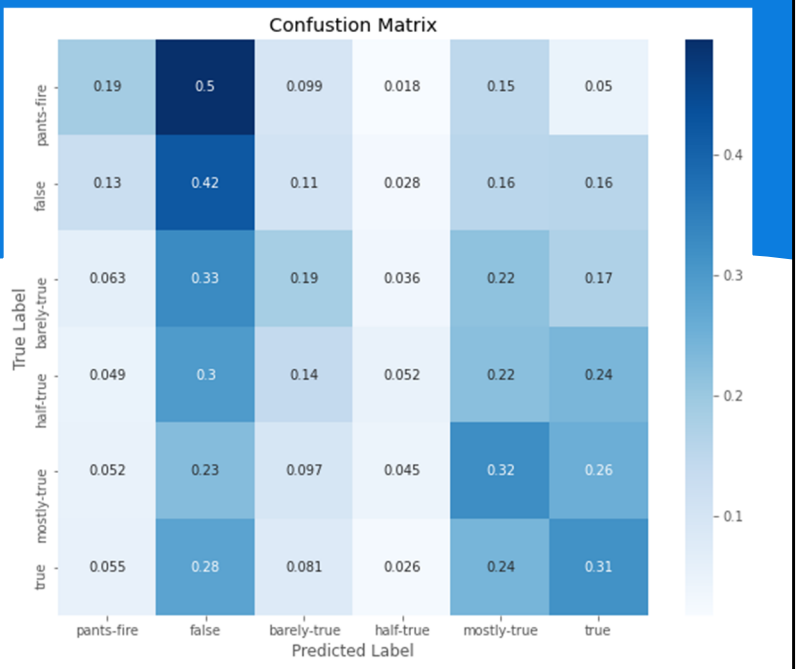



1. The data was pulled from a website called Politifact.com. PolitiFact researches political and social context that is impactful to the public. Almost all of the content comes in the forms of quotes or general statement. I will generally refer to individual pieces of text as a quote.
2. In total there were around 18500 quotes that fell into 6 distinct categories. The most numerous is the “false” category and the least numerous was the “pants-fire” category.
3. Pants-fire is truly a subcategory of false including quotes that were false with the added caveat of being particularly outlandish.
4. Shown is an example of a Politifact.com fact check. On the website you can click the quote and pull up the article that explains the rating given.

# MODELS

LSTM with Non-Lemmatized Tokens

Accuracy: 26%



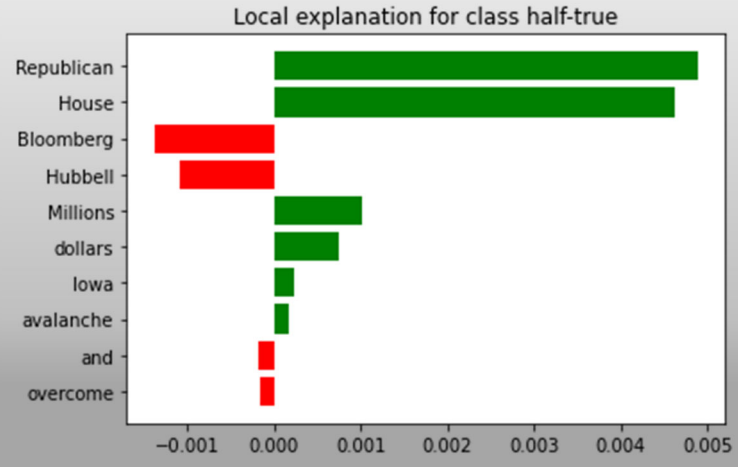

The best performing model to date was the LSTM neural network with non-lemmatized tokens. Don't worry if you don't know what this means! I tried many different types of models and I wanted to make sure to call out the winner.

Looking at the confusion matrix, the categories are logically ordered from left to right with pants-fire on the left and getting progressively more truthful with true on right. The total accuracy of the model was 26% with false, mostly-true and true being the best performing categories. Because pants-fire is more of a subset of the false category, I'm not disappointed at the misclassification between the two. Ultimately, the distinction between false and outlandishly false is more academic than practical.

# Example Classification

## Half-True Example

Millions of Hubbell and Bloomberg dollars wasn't enough to overcome the Iowa House Republican avalanche of success.'




Here we see an example of a correct classification by our model (half-true). The graph on the right show which words pushed the model towards classifying half-true – green – and which pushed the model away from half-true – red. Because we are dealing with a multi-class problem the direction in which the red words push the model isn't exactly clear, the best we can say at a high level is "away from picking the half-true label."

What about the words that aren't highlighted? Does this mean the model ignored them?

Not exactly. The model uses every word during the classification task. In the case of the unhighlighted words, the model determined that their affect on picking half-true is so small that they aren't worth showing in the graph. In some cases, we're talking about considering down the 32nd decimal place.

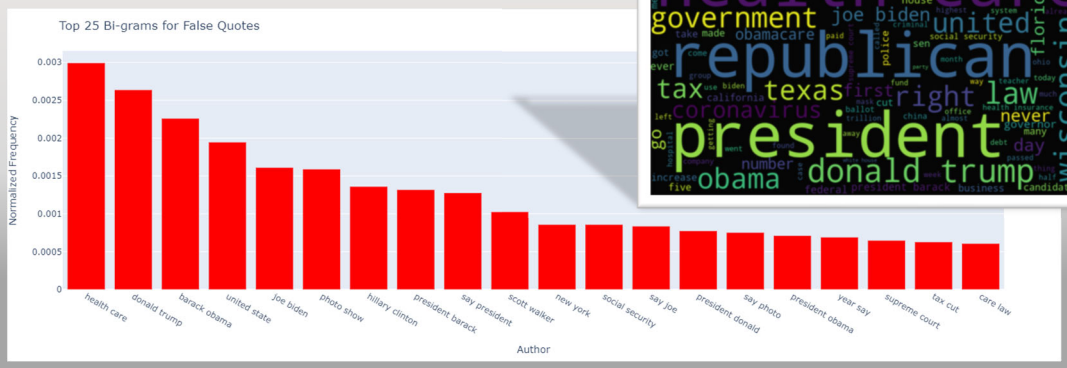
# The Truth

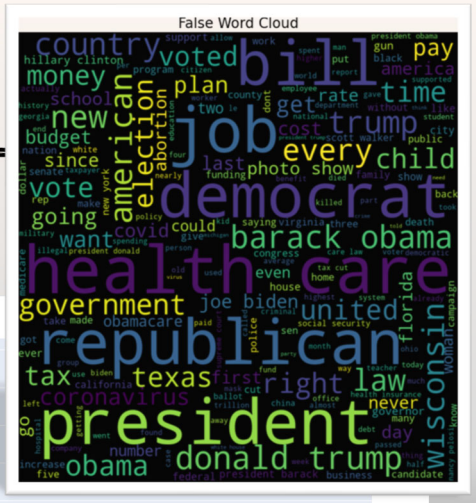
See something similar?


Unfortunately, when it comes to the non-false categories, there isn't much in the way of uniquely determining features. Looking at the word cloud, the vast majority of the commonly used words are repeats and with similar concentrations. This also holds true for the bi-grams in each case. What does stand out, not of any individual category but within the group, is the reference to numbers and statistics. Examples like "1 percent" or "national debt" only show up in the non-false categories and

# False Quotes

**Has no unique word pairs in the top 25**



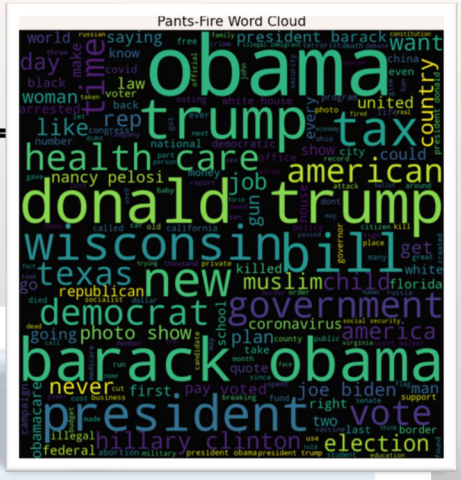

In addition to the unique word pairs, here are some statistics

- 10 of 25 include individuals or mentions of individuals (40%)
- 6 of 25 are political/social issues (24%)
- 1 is a state (4%)
- 0 are number based (0%)
- 2 talk about photos (8%)
- 3 include the word "says" (12%)

# PANTS-FIRE

**Word pairs that are unique in the top 25:**

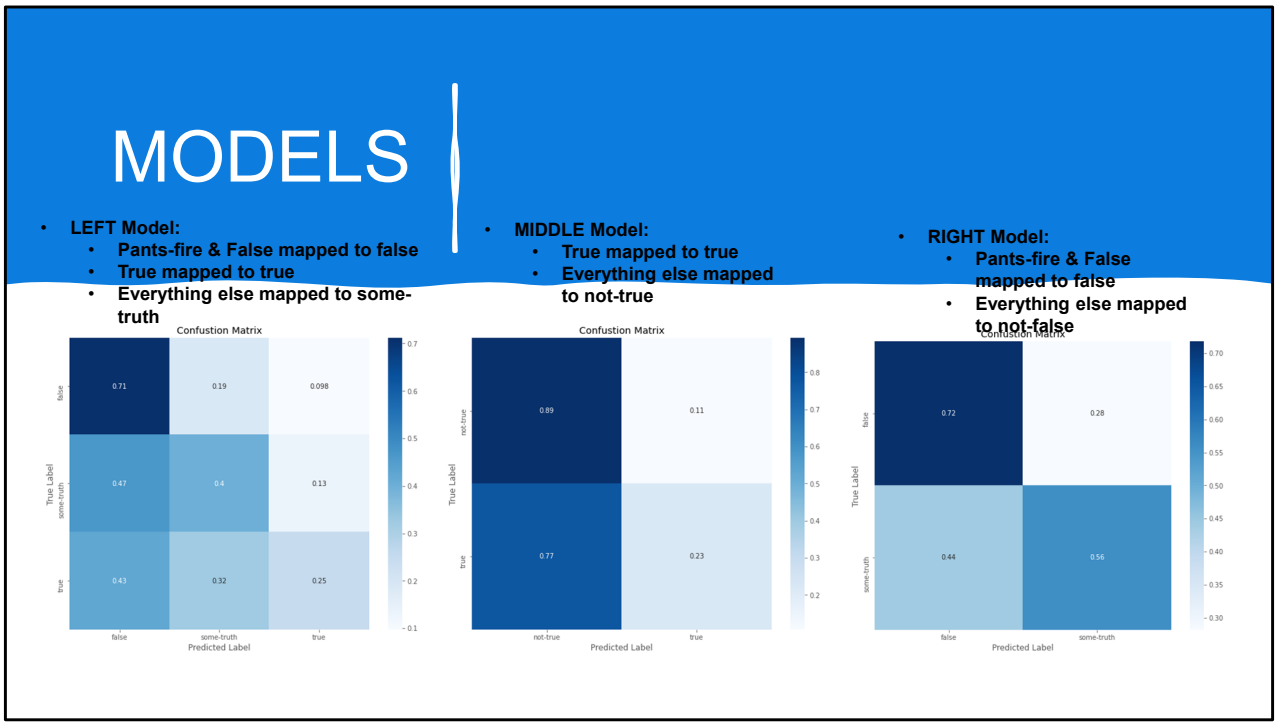
- **George Soros**
- **Says Barack**
- **Nancy Pelosi**





In addition to the unique word pairs, here are some statistics

- 15 of 25 include individuals or mentions of individuals (60%)
- 2 of 25 are political/social issues (8%)
- 0 are states (0%)
- 0 are number based (0%)
- 2 talk about photos (8%)
- 5 include the word "says" (20%)




Because of the lackluster performance of the 6 category model I thought it prudent to collapse labels into a coarser structure. The attempt was to try and increase total model accuracy without compromising the practicality of the model. Here's how I combined the labels.

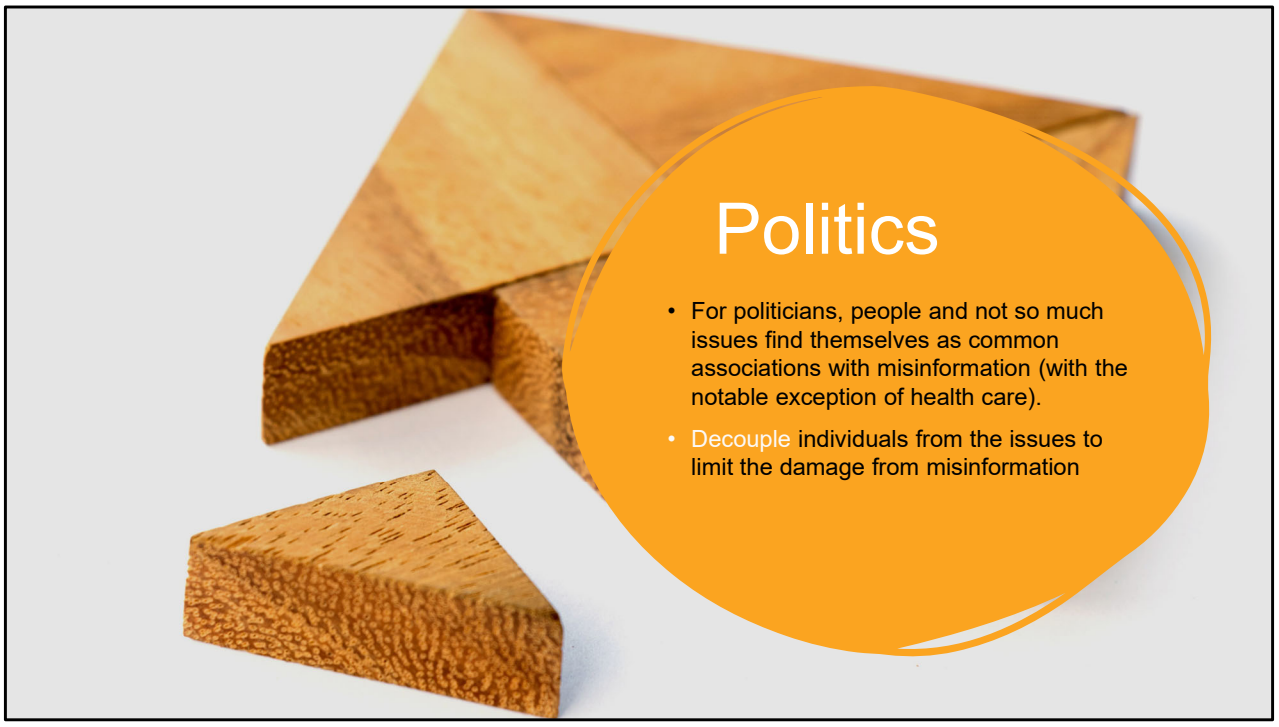
- LEFT Model:
  - Pants-fire & False mapped to false
  - True mapped to true
  - Everything else mapped to some-truth
- MIDDLE Model:
  - True mapped to true
  - Everything else mapped to not-true
- RIGHT Model:
  - Pants-fire & False mapped to false
  - Everything else mapped to not-false

Once again, it's clear that the false categories are captured at a higher rate than any other category. Of all the label mappings, the model to the far right is the most

useful, being able to spot 72% of all false quotes. This process is not very precise yet, hence, why I claim it is not ready for a production environment.

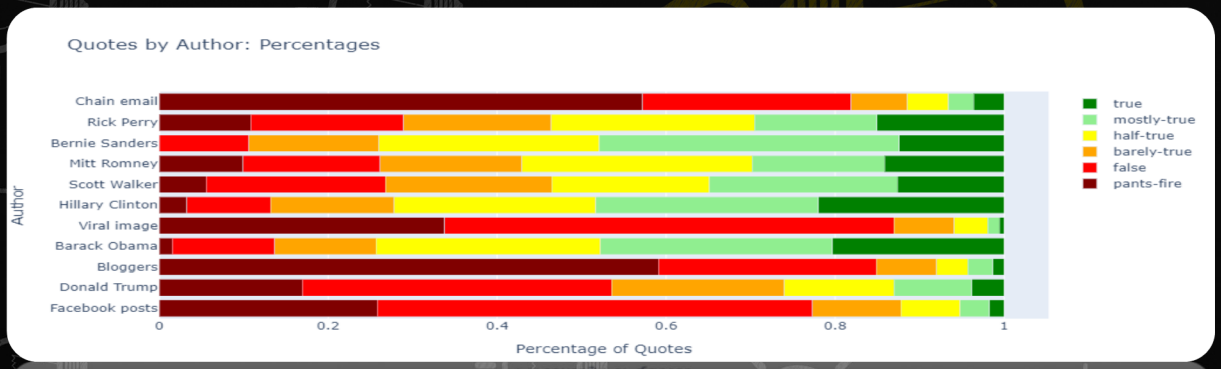
# Fighting Misinformation

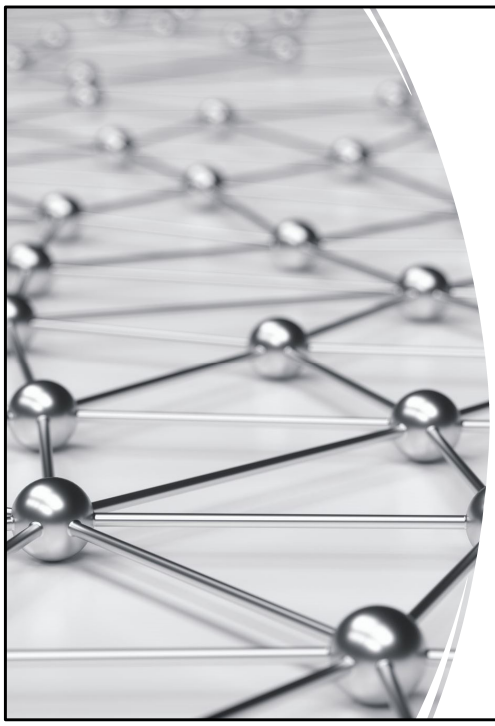
- Mentions of “**photo”** content were only mentioned in the top 25 word pairs for the false and pants-fire categories.
- Flag content with mentions of photos (not necessarily photos themselves) and follow up with moderator verification.


Politics
• For politicians, people and not so much issues find themselves as common associations with misinformation (with the notable exception of health care).
• Decouple individuals from the issues to limit the damage from misinformation

# Perpetuating Information

- **Don't repost/tweet/share info from social media without verifying sources. Viral content with anonymous authors have incredibly *high* rates of false-hood.**





# Future Work

- I would like to bring in some more quotes.
- The second modeling goal is to implement a v ersion of a BERT transformer.
- As for the data exploration; I would like to mak e more use of the various NLTK apis for text e ngineering.
- I would like to make the data more interactive on the app such that users can select specific quotes and interact with them through the vario us text models.

# References


- Politifact.com They are awesome and you should consider donating!
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukas z Kaiser, Illia Polosukhin: Attention is All you Need. NIPS 2017: 5998-6008
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL-HLT (1) 2019: 4171-4186
- Keras Developers: Keras API Docs: https://keras.io/api/
- Ploly Developers: Plotly API Docs: https://dash.plotly.com/
- Scikit Learn Developers: Scikit Learn API Docs: https://scikit-learn.org/stable/modules/classes.html
- Pandas Developers: Pandas API Docs: https://github.com/pandas-dev/pandas
- Jacob Kessler: ScatterText: https://github.com/JasonKessler/scattertext
- Marco Tulio Correia Ribeiro: Lime: https://github.com/marcotcr/lime
- Radim Rehrek, Petr Sojka: Software Framework for Topic Modelling with Large Corpora: Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks: p. 45 - 50 : 2010: ELRA: Valletta, Malta
- Andreas Mueller: word-cloud: https://github.com/amueller/word_cloud