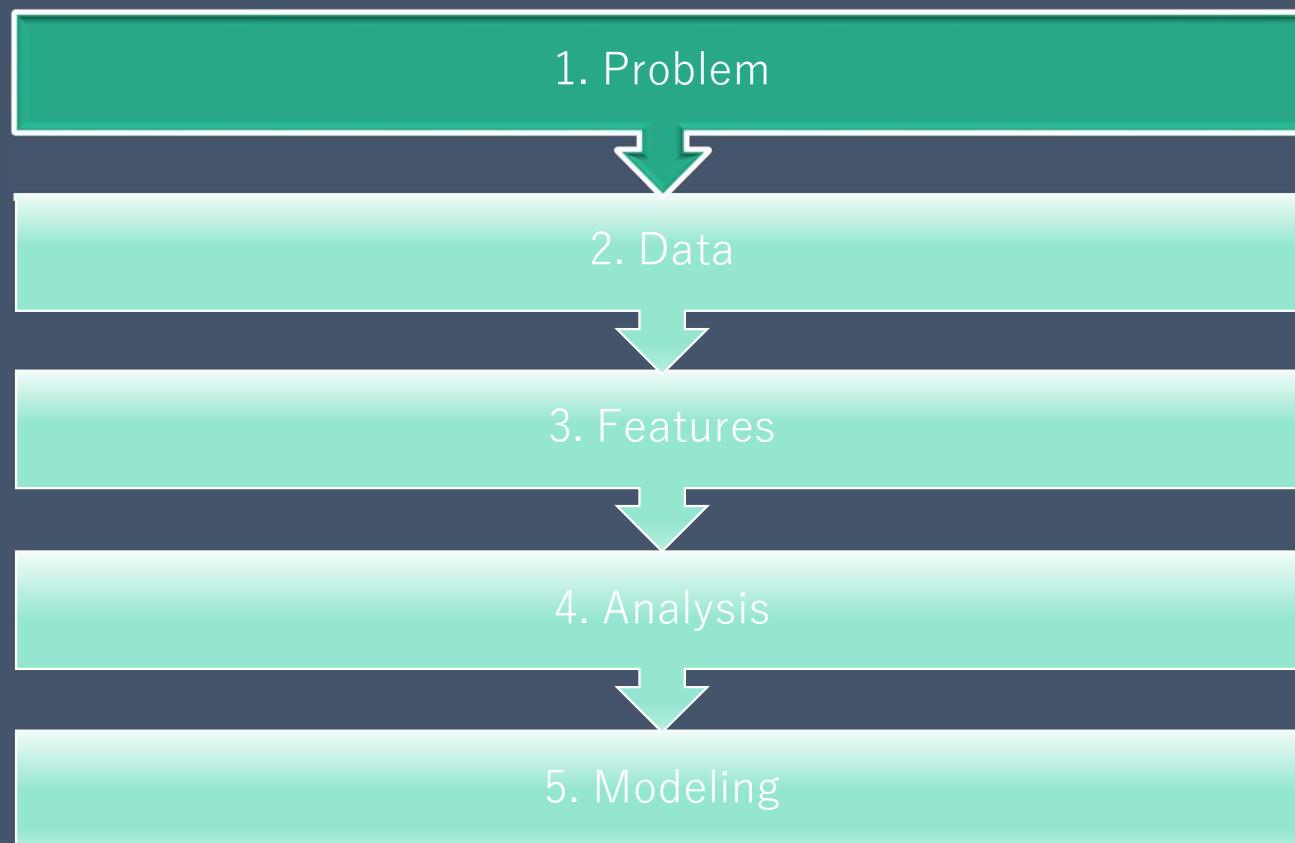


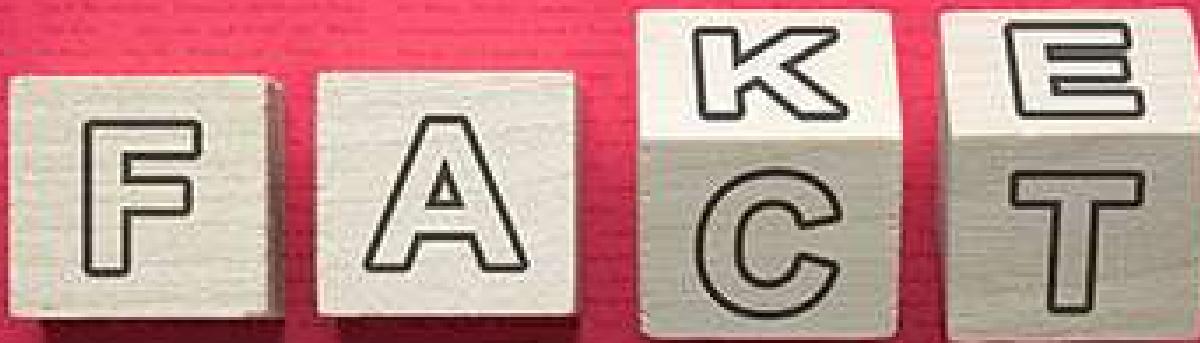
Fake News Detection

By Anna Douville



Problem Statement

In order to stop the spread of fake news, can social media websites use an algorithm that outperforms human fake news detection by using NLP to determine whether an article is real or fake?

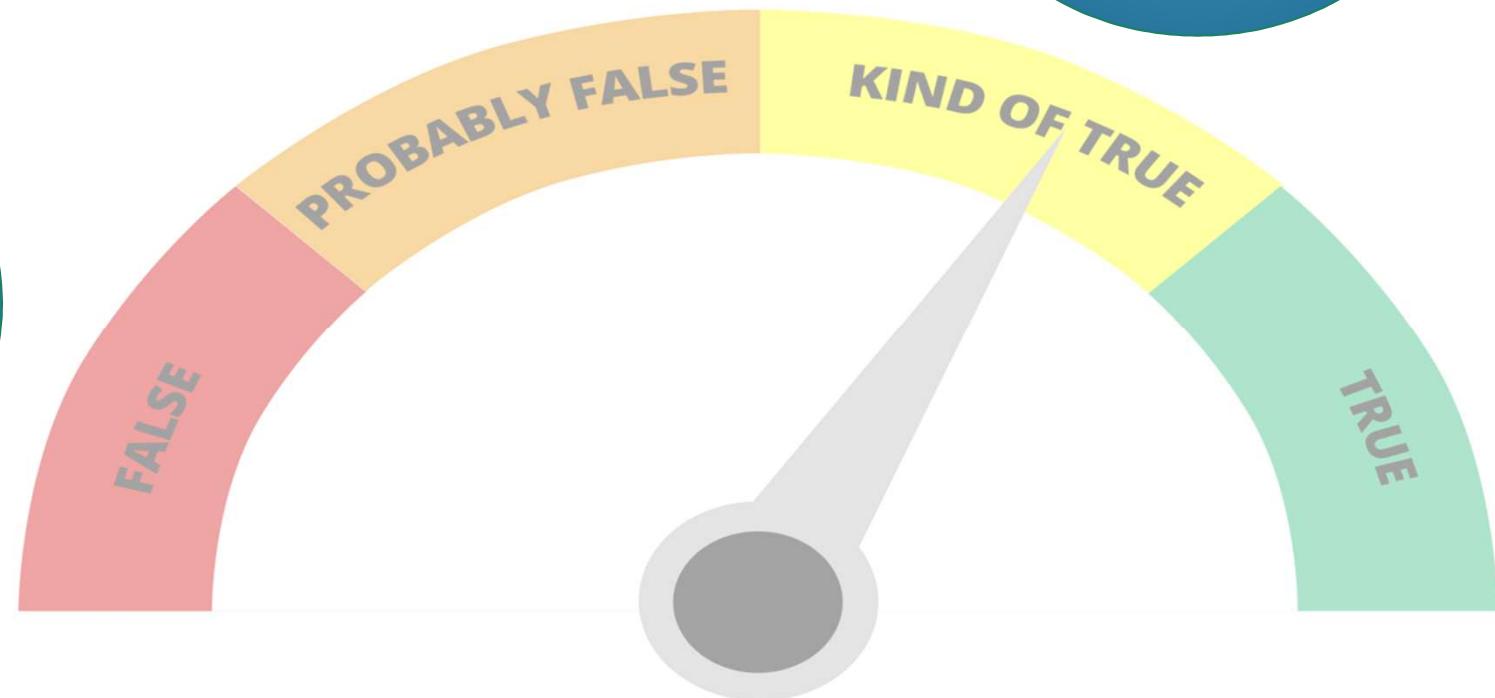


Background

Social media has made it easier for fake news to spread

54%

Average accuracy
rate of human fake
news detection



67%

Percent of Americans
that use social media
as their main news
source

Effects

Political influence

- 2016 presidential election

Fear Mongering

- Hurricane Irma

Public Safety

- COVID-19

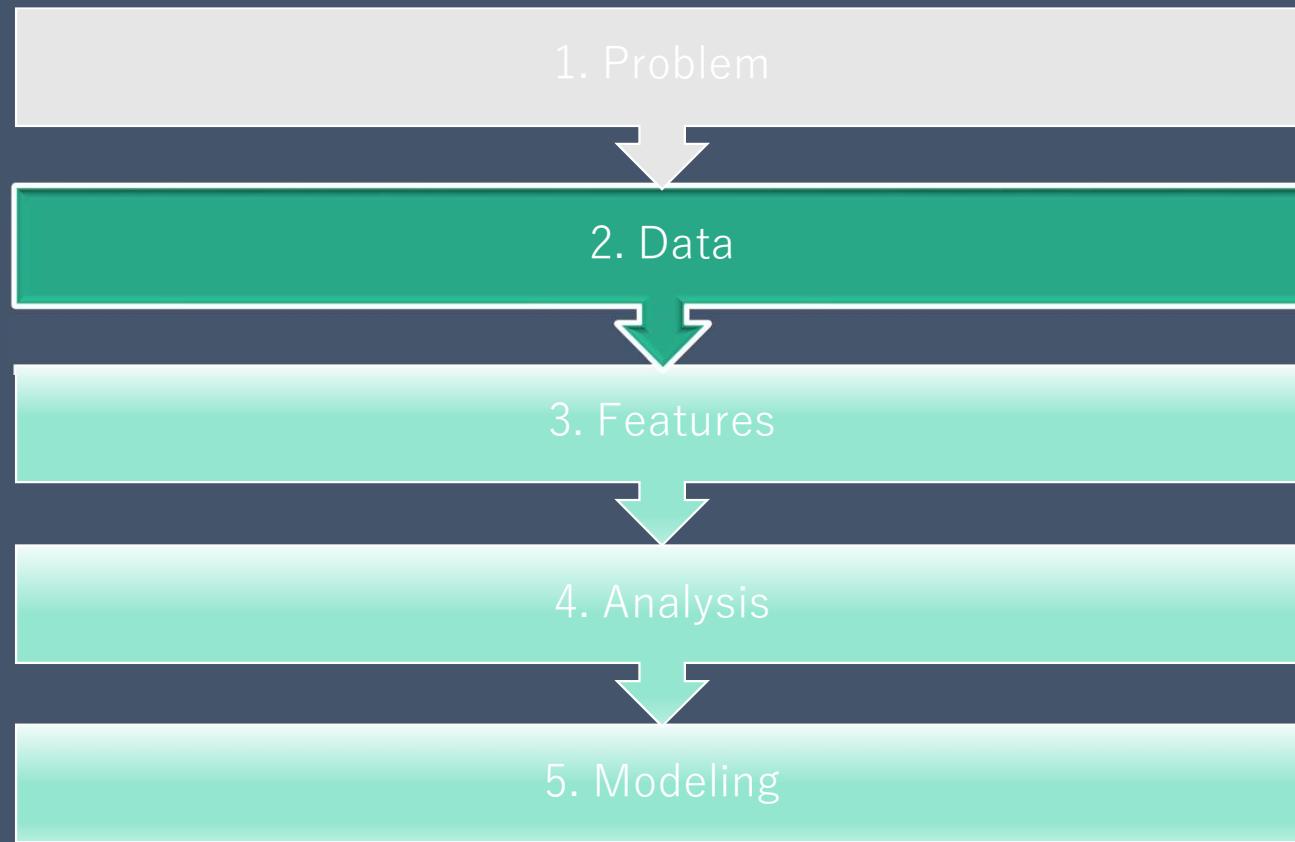


Stakeholders

Social Media Websites

- Facebook
- Instagram
- Twitter
- Snapchat





The Data

- University of Victoria ISOT Lab
- 21,417 real articles from Reuters
- 23,481 fake articles from various fake news websites
- 4 columns
 - Date published
 - Subject
 - Title
 - Text



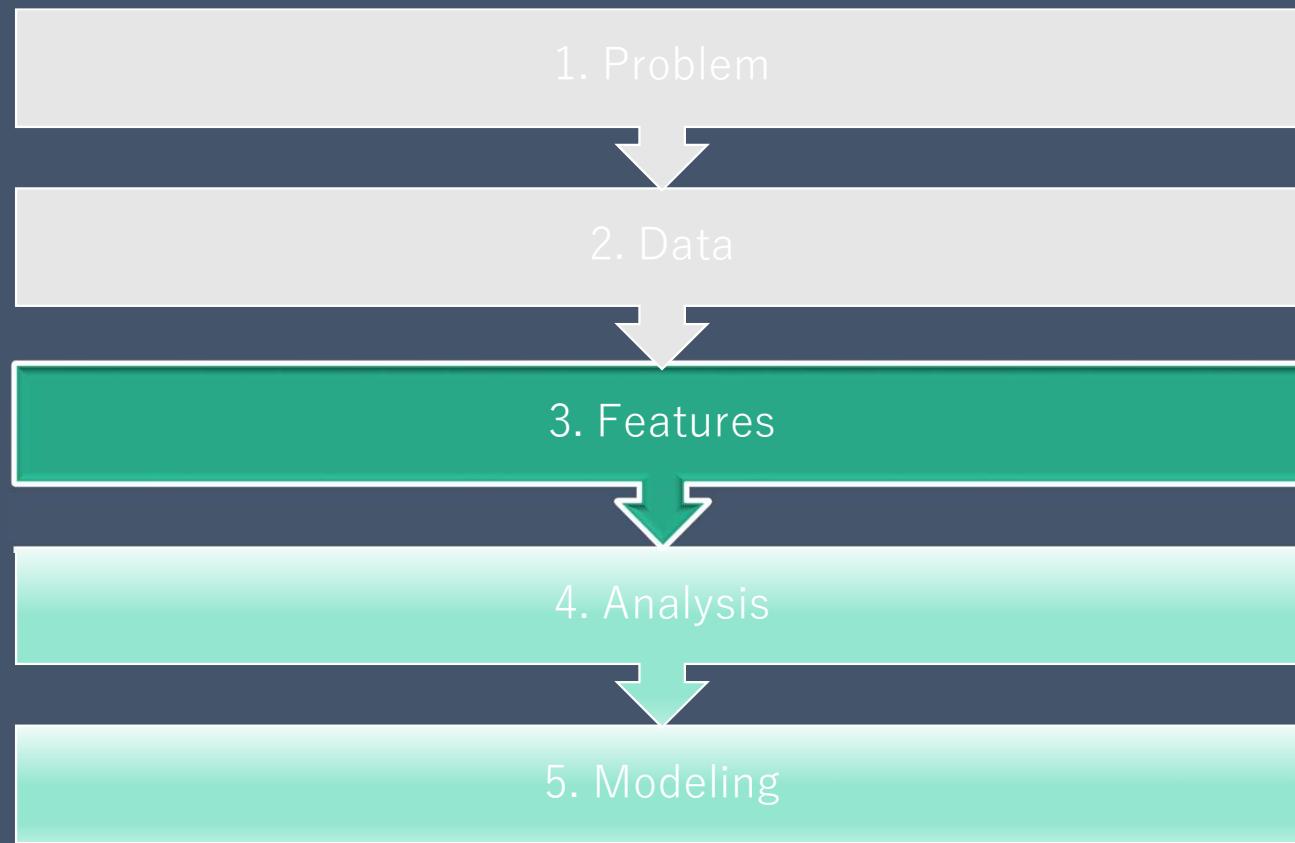
Data Cleaning

- Removed beginning phrase from Reuters articles

text
WASHINGTON (Reuters) - The head of a conservat...
WASHINGTON (Reuters) - Transgender people will...
WASHINGTON (Reuters) - The special counsel inv...
WASHINGTON (Reuters) - Trump campaign adviser ...
SEATTLE/WASHINGTON (Reuters) - President Donal...

- Dropped subject column
- Dropped 5,793 rows of duplicates
- 39,104 rows total





Basic Feature Extraction

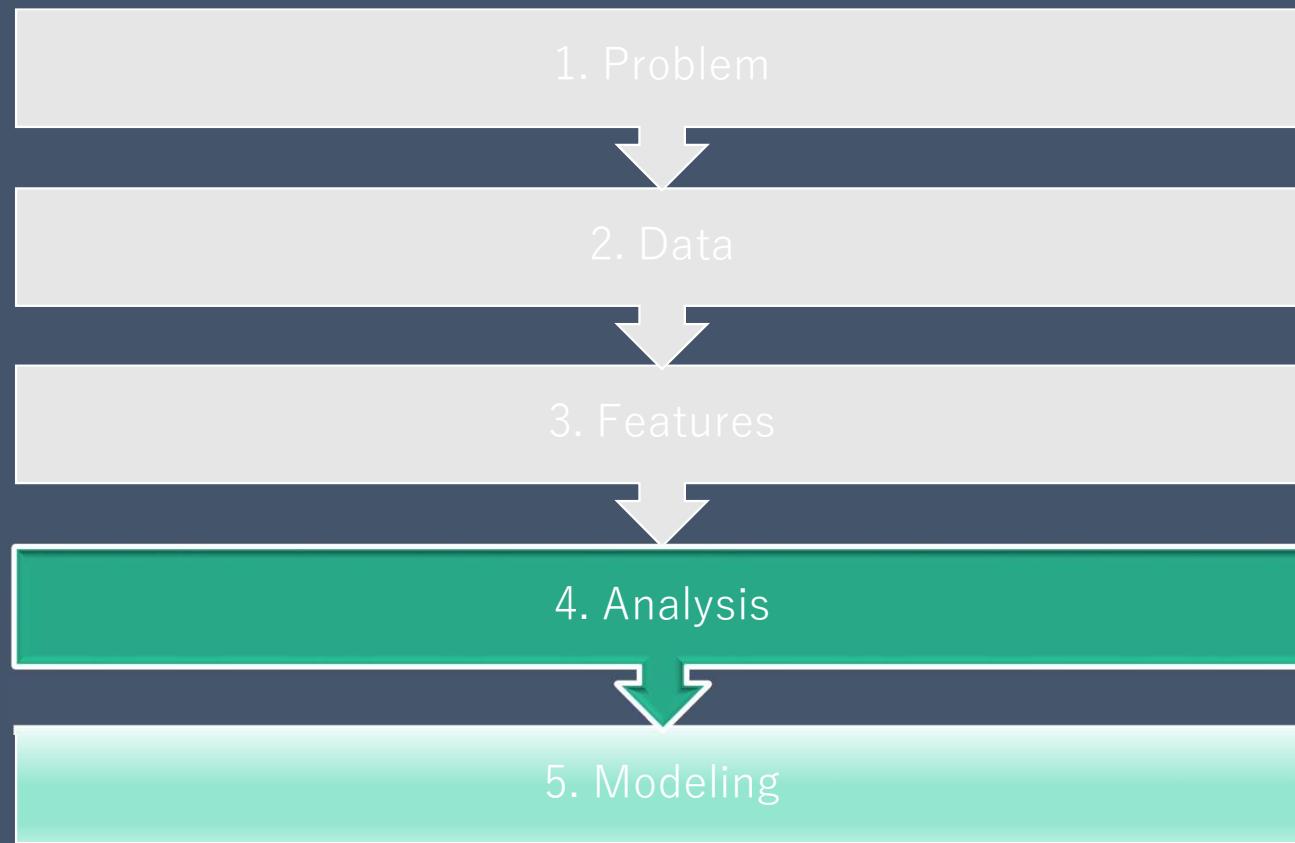
- Word counts
 - Total number of words
 - Stop words
 - Upper case words
- Average word length
- Special character counts
 - Exclamation marks
 - Question marks
 - @ signs
 - Numbers
- Date features
 - Weekend
 - Time of year
- 16 basic features total



TF-IDF Vectorization

- Text pre-processing
 - Made all lowercase
 - Removed punctuation, numbers, and stop words
 - Lemmatization
- Included unigrams and bigrams
- 800 text n-grams
- 200 title n-grams

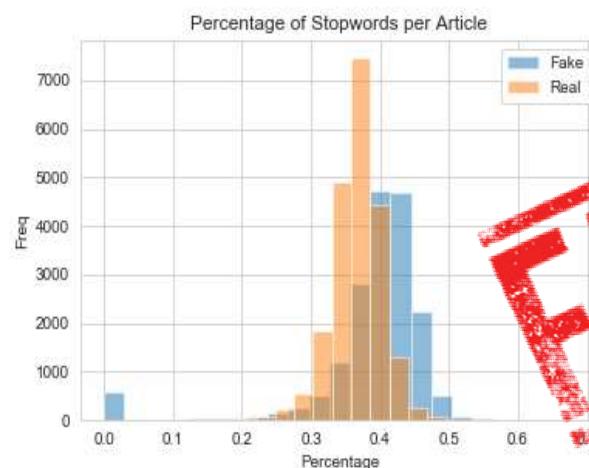
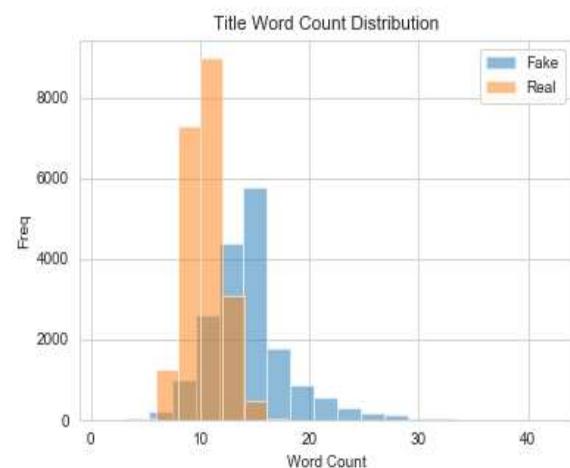
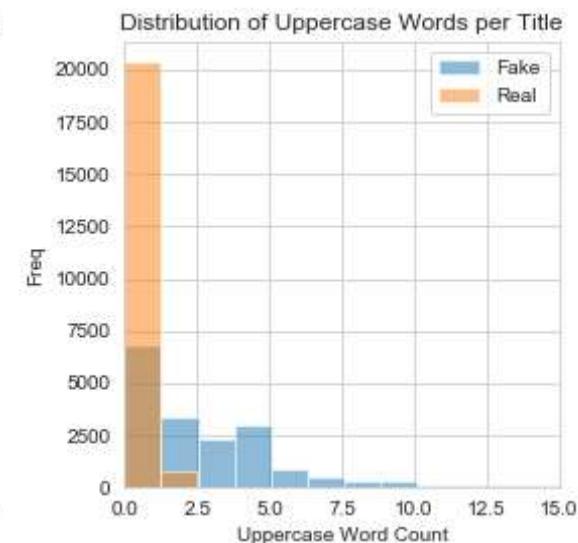
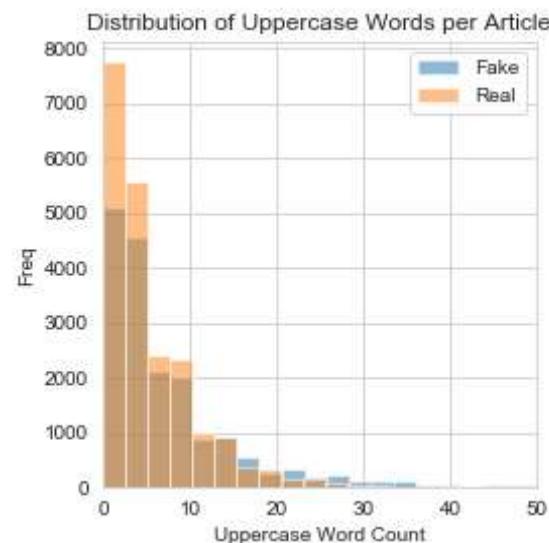




Word Counts

Fake news has...

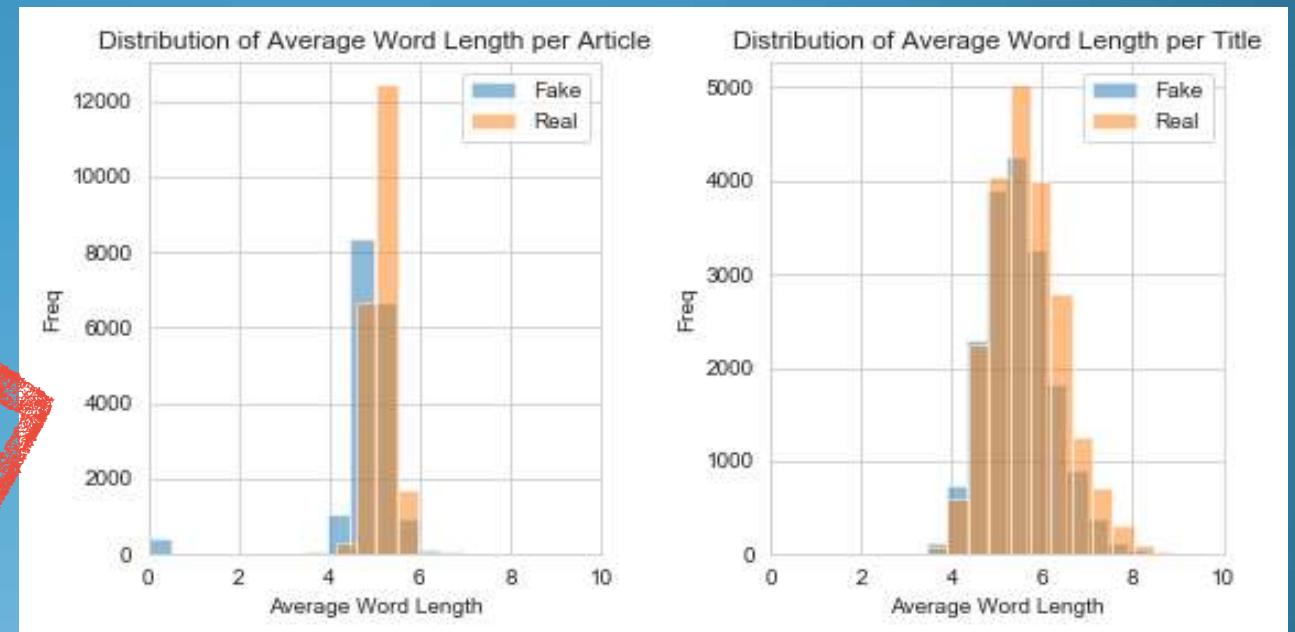
- Longer titles
- More stop words
- More upper case words



FAKE NEWS

Average Word Length

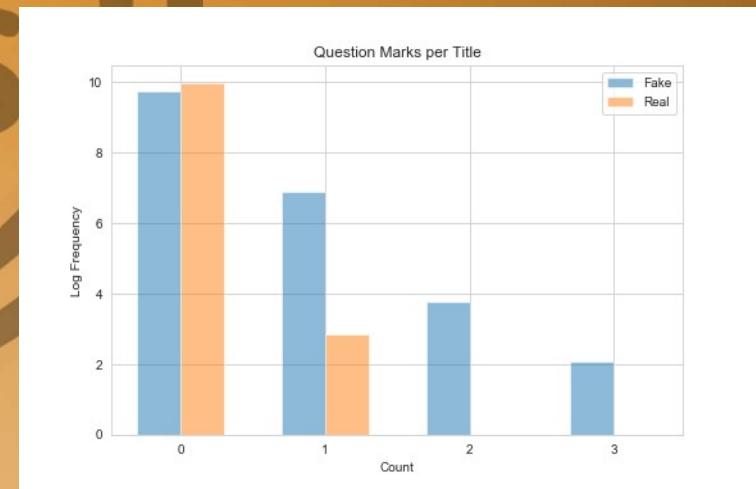
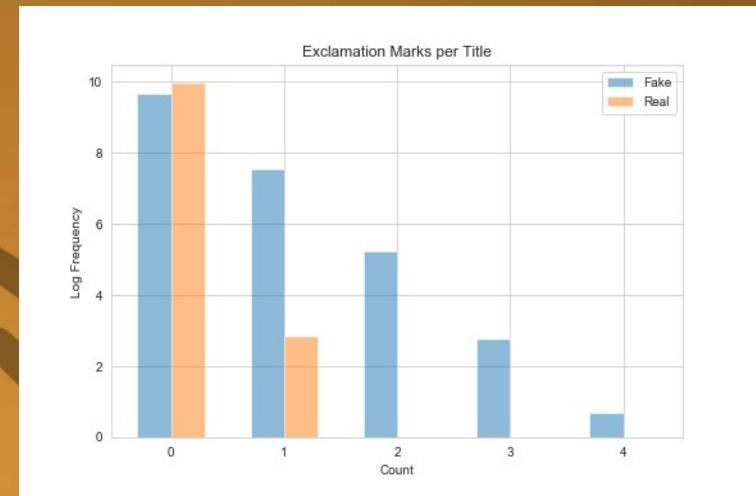
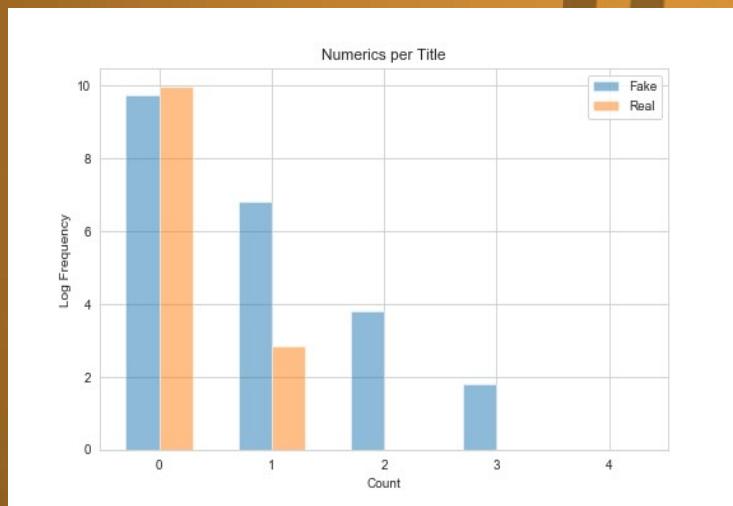
Real News has slightly longer words on average



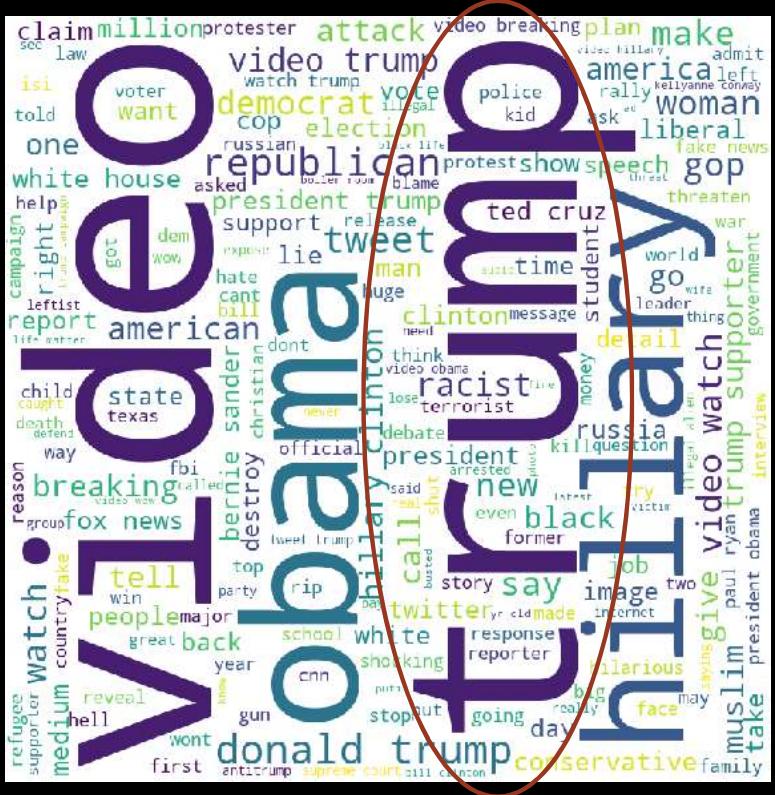
Special Characters

Fake News has more...

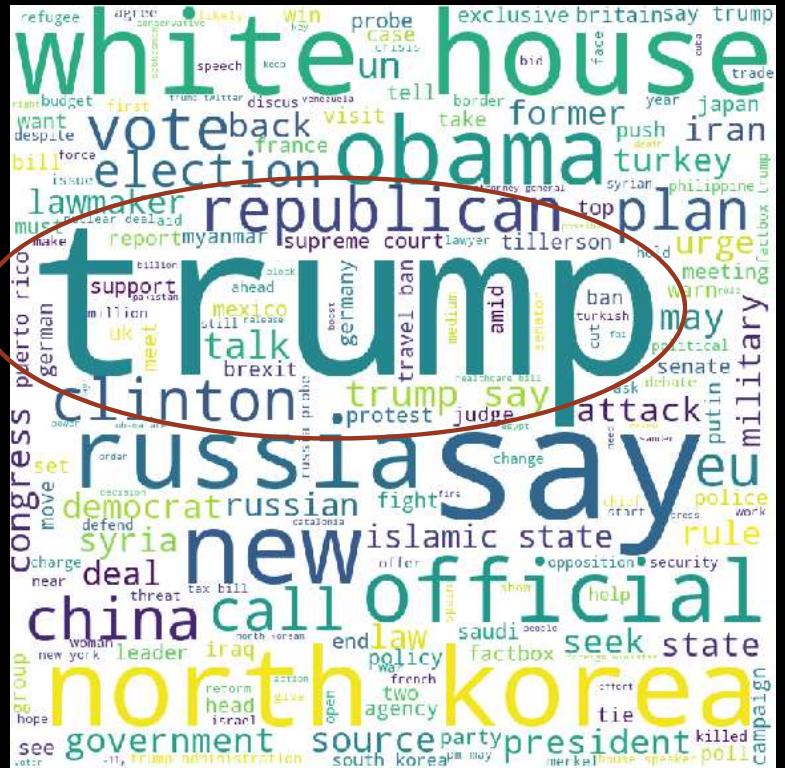
- Exclamation marks
- Question marks
- Numbers



Word Clouds

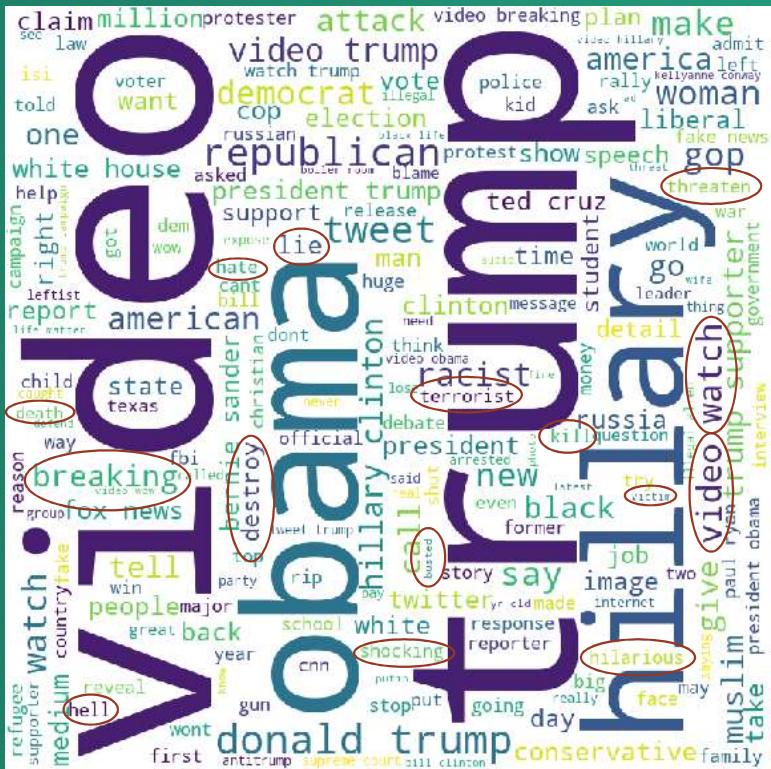


Fake Titles



True Titles

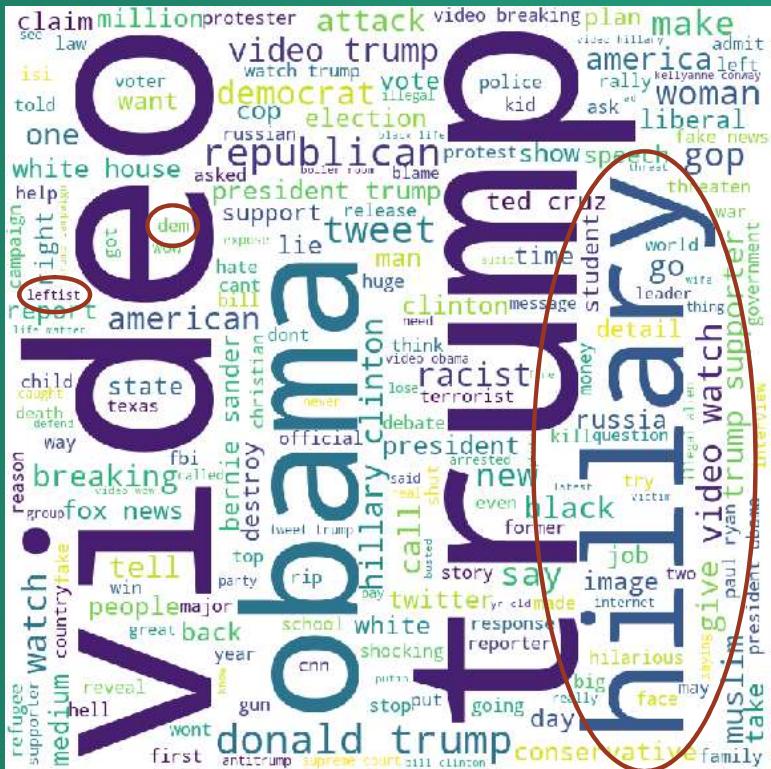
Word Clouds



Fake Titles

- Clickbait words
 - ‘watch’, ‘video’, ‘hell’, ‘terrorist’, ‘lie’, ‘breaking’, ‘threaten’, ‘busted’, ‘shocking’, and ‘destroy’

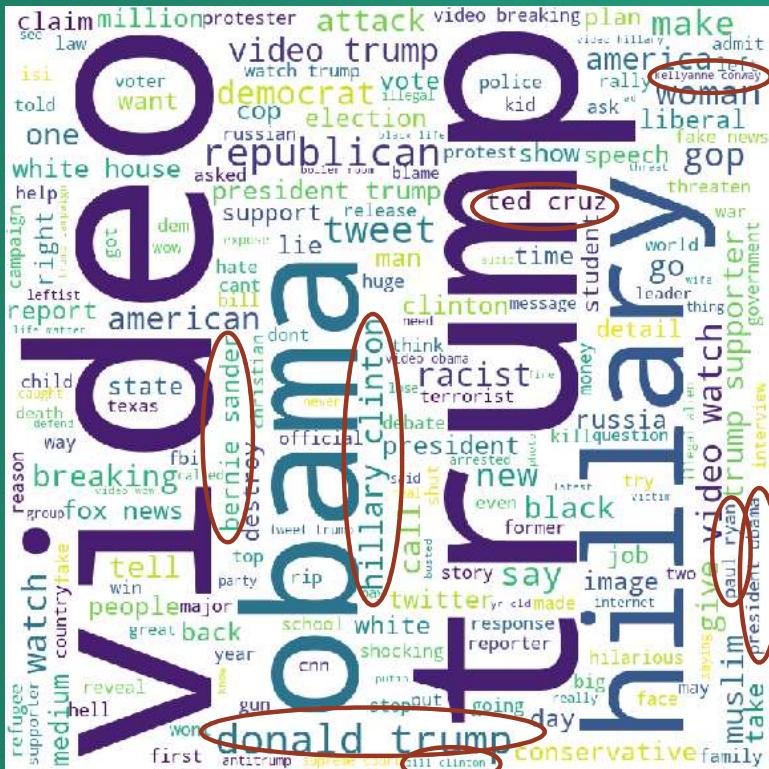
Word Clouds



Fake Titles

- Clickbait words
 - ‘watch’, ‘video’, ‘hell’, ‘terrorist’, ‘lie’, ‘breaking’, ‘threaten’, ‘busted’, ‘shocking’, and ‘destroy’
 - Informal words
 - ‘hillary’, ‘dem’, ‘leftist’

Word Clouds



Fake Titles

- Clickbait words

- ‘watch’, ‘video’, ‘hell’, ‘terrorist’, ‘lie’, ‘breaking’, ‘threaten’, ‘busted’, ‘shocking’, and ‘destroy’

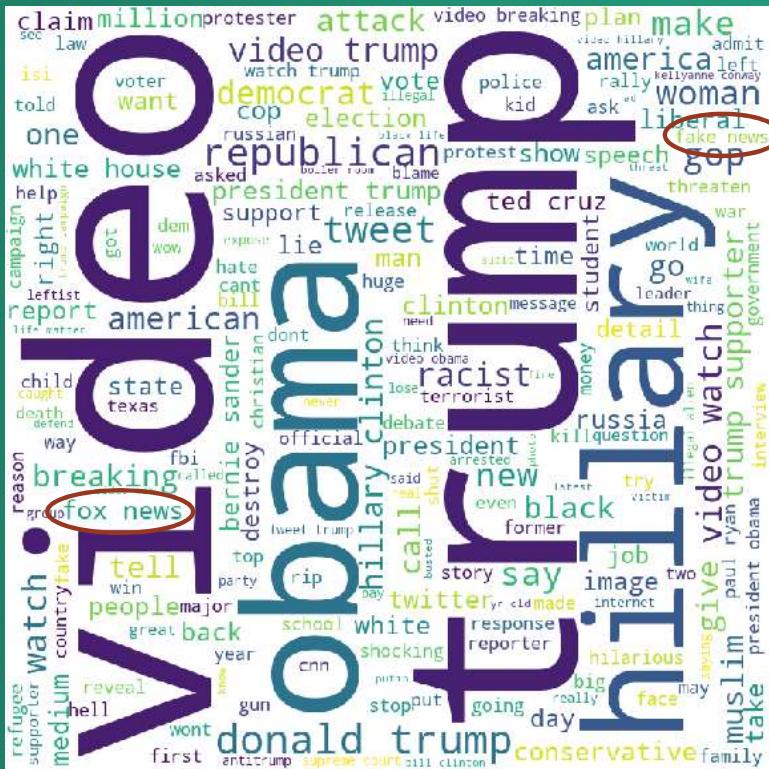
- Informal words

- ‘hillary’, ‘dem’, ‘leftist’

- People

- ‘trump’, ‘hillary’, ‘obama’, ‘ryan paul’, ‘kellyanne conway’, ‘ted cruz’, ‘bernie sanders’

Word Clouds



Fake Titles

- Clickbait words

- ‘watch’, ‘video’, ‘hell’, ‘terrorist’, ‘lie’, ‘breaking’, ‘threaten’, ‘busted’, ‘shocking’, and ‘destroy’

- Informal words

- ‘hillary’, ‘dem’, ‘leftist’

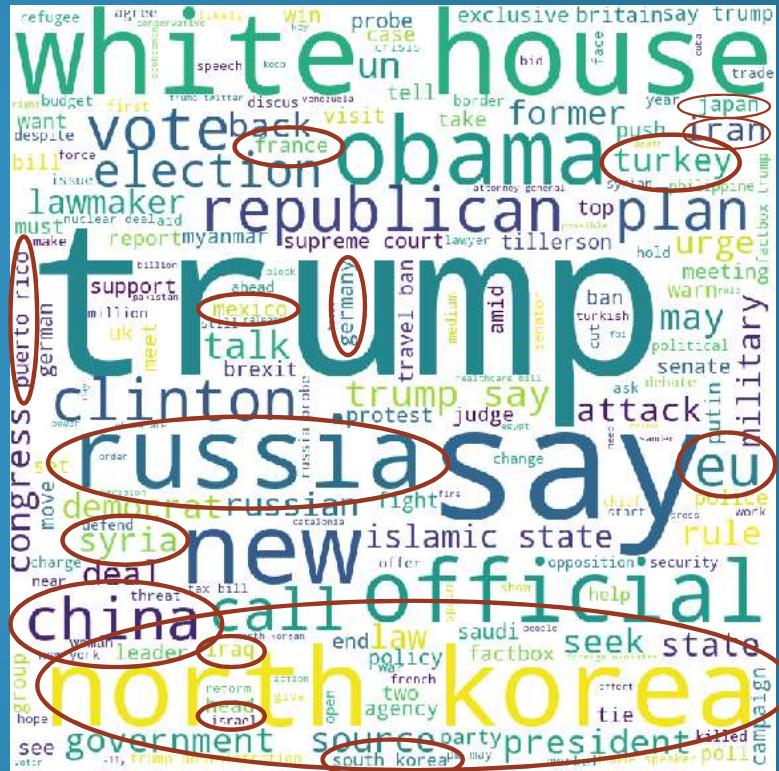
- People

- ‘trump’, ‘hillary’, ‘obama’, ‘ryan paul’, ‘kellyanne conway’, ‘ted cruz’, ‘bernie sanders’

- ‘fake news’
- ‘fox news’

Word Clouds

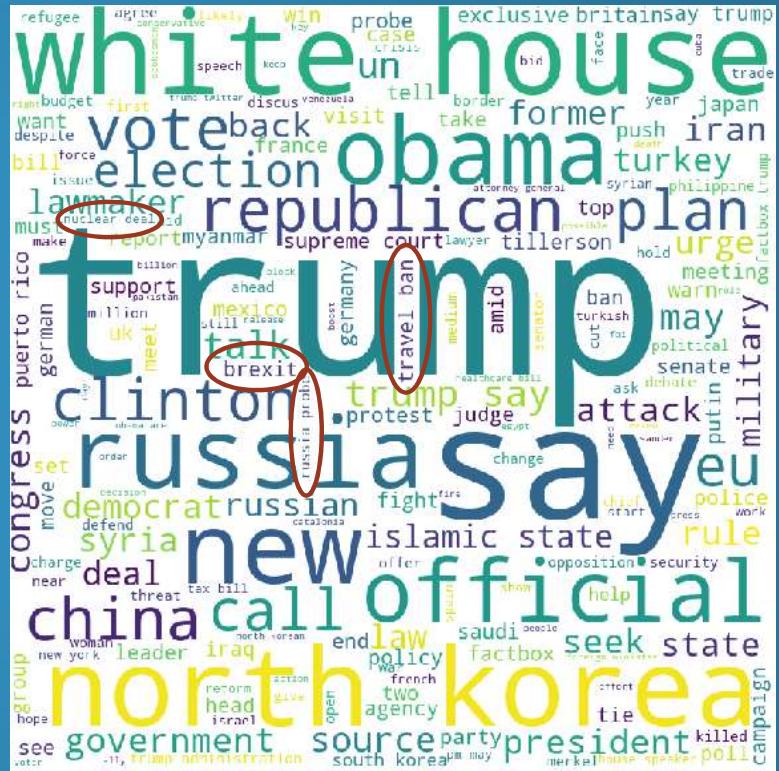
- Places
 - , ‘north korea’, ‘germany’, ‘iran’, ‘turkey’, ‘japan’, etc.



True Titles

Word Clouds

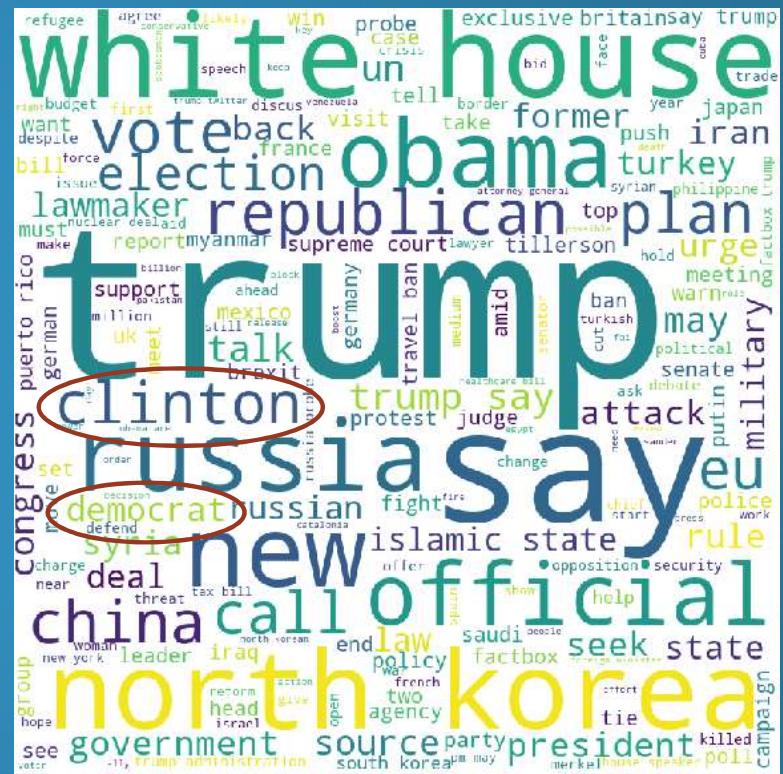
- Places
 - , ‘north korea’, ‘germany’, ‘iran’, ‘turkey’, ‘japan’, etc.
 - World news events
 - ‘brexit’, ‘travel ban’, ‘russian probe’, ‘nuclear deal’



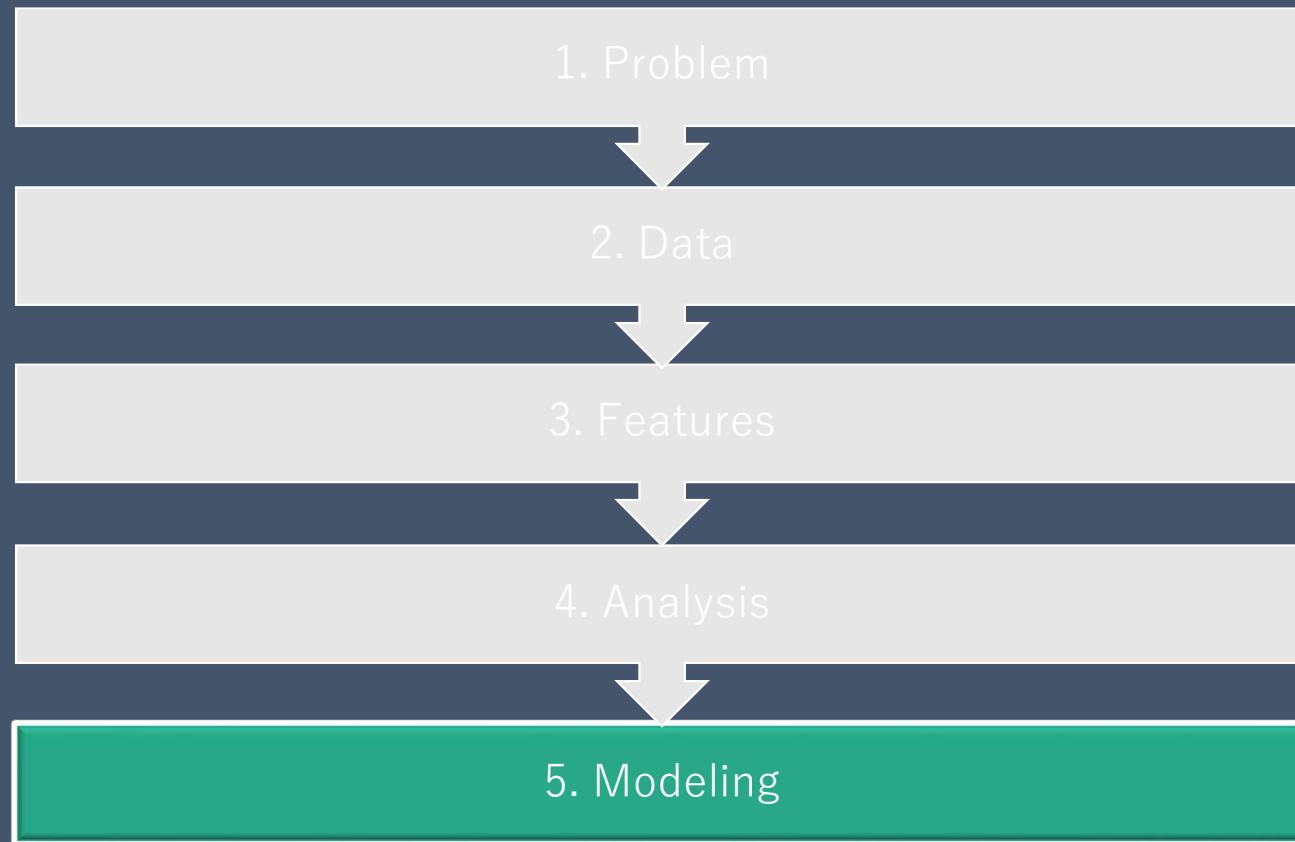
True Titles

Word Clouds

- Places
 - , ‘north korea’, ‘germany’, ‘iran’, ‘turkey’, ‘japan’, etc.
- World news events
 - ‘brexit’, ‘travel ban’, ‘russian probe’, ‘nuclear deal’
- Formal words
 - ‘clinton’, ‘democrat’



True Titles

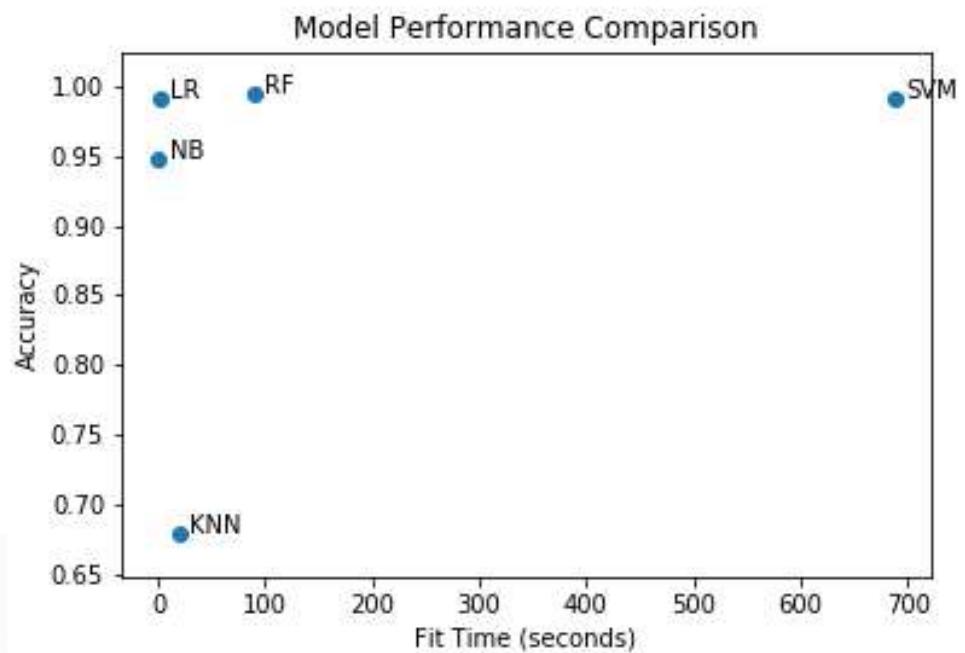


Model Metrics

Logistic Regression performed the best overall



FAKE
NEWS



	accuracy	fit_seconds
NB	0.948	0.11
RF	0.994	89.49
LR	0.991	1.40
SVM	0.992	687.00
KNN	0.679	19.00

Summary

- Key Differences
 - Word counts
 - Special character counts
 - Verbiage
- Model
 - Basic features, TF-IDF
 - Logistic Regression
 - >99% accuracy
- Applications
 - Compile a list of fake news websites
 - Warn users of fake news
 - Filter out fake news

