Detecting Fake News

CS522 – Data Mining Final Project

Brandon Lasher

Outline

- Project Overview
- Project Implementation
- Conclusions

Project Overview

Problem

- With the widespread acceptance of social media as a source of news, the occurrence of fake news has become far more prevalent. It is easy to simply share an interesting story without any verification process.
- This has affected everything from major news networks, policies and political campaigns; even to discussions around the dinner table.
- Is it possible to detect Fake news based on the language used instead of time-consuming fact checking?

Goals

- Look features of fake news vs true news.
- Create a model which can predict if a text is fake news.
- Apply the model to a separate dataset, to see how it hold up.

Project Implementation

Data Collection and Overview

Data Cleaning

Data Exploration

Model Building

Model Testing

Data Collection and Overview

- Data Source: https://www.kaggle.com/code/josutk/only-one-word-99-2/data
  - Fake.csv and True.csv
- About the data
  - Text and subject fields includes urls, tags, images, html and other abbreviations.
  - No null fields
- Data shape
  - Fake.csv - (23481, 4) , True.csv - (21417, 4)

Feature Name	Data Type	Description
title	object	Title of the article
text	object	Text of article, may include publisher and author
subject	object	Type of News
date	date	Date of publication

Data Cleaning

1

Clean the date field and reformat

2

Add a type field to keep track of the news type

3

Merge title and text fields into one called news

- Remove the original title and text fields

4

Clean the newly created news field

- Remove tags, urls and html
- Remove any additional punctuation
- Remove any return lines or newlines
- Remove any numbers
- Make everything lowercase
- Remove stopwords
  - Stopwords are commonly occurring words such as ‘the’, ‘a’, ‘in’, etc.

Data Cleaning

- Before Cleaning

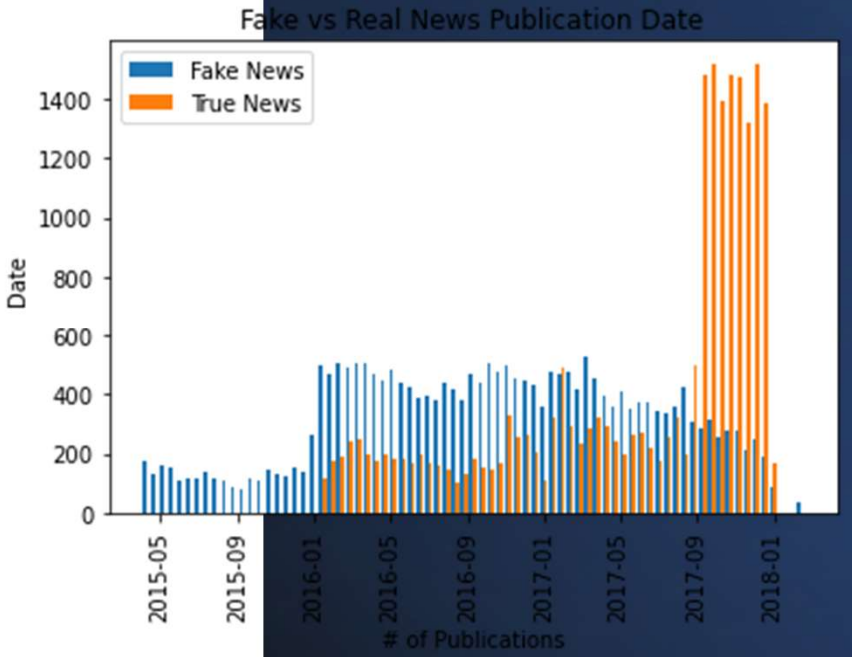
' Donald Trump Sends Out Embarrassing New Year’s Eve Message; This is Disturbing Donald Trump just couldn t wish all Americans a Happy New Year and leave it at that. Instead, he had to give a shout out to his enemies, haters and the very dishonest fake news media. The former reality show star had just one job to do and he couldn t do it. As our Country rapidly grows stronger and smarter, I want to wish all of my friends, supporters, enemies, haters, and even the very dishonest Fake News Media, a Happy and Healthy New Year, President Angry Pants tweeted. 2018 will be a great year for America! As our Country rapidly grows stronger and smarter, I want to wish all of my friends, supporters, enemies, haters, and even the very dishonest Fake News Media, a Happy and Healthy New Year. 2018 will be a great year for America! Donald J. Trump (@realDonaldTrump) December 31, 2017Trump s tweet went down about as welll as you d expect.What kind of president sends a New Year s greeting like this despicable, petty, infantile gibberish? Only Trump! His lack of decency won t even allow him to rise above the gutter long enough to wish the American citizens a happy new year! Bishop Talbert Swan (@TalbertSwan) December 31, 2017no one likes you Calvin (@calvinstowell) December 31, 2017Your impeachment would make 2018 a great year for America, but I ll also accept regaining control of Congress. Miranda Yaver (@mirandayaver) December 31, 2017Do you hear yourself talk? When you have to include that many people that hate you you have to wonder? Why do the they all hate me? Alan Sandoval (@AlanSandoval13) December 31, 2017Who uses the word Haters in a New Years wish?? Marlene (@marlene399) December 31, 2017You can t just say happy new year? Koren pollitt (@Korencarpenter) December 31, 2017Here s Trump s New Year s Eve tweet from 2016.Happy New Year to all, including to my many enemies and those who have fought me and lost so badly they just don t know what to do. Love! Donald J. Trump (@realDonaldTrump) December 31, 2016This is nothing new for Trump. He s been doing this for years.Trump has directed messages to his enemies and haters for New Year s, Easter, Thanksgiving, and the anniversary of 9/11. pic.twitter.com/4FPAe2KypA Daniel Dale (@ddale8) December 31, 2017Trump s holiday tweets are clearly not presidential.How long did he work at Hallmark before becoming President? Steven Goodine (@SGoodine) December 31, 2017He s always been like this . . . the only difference is that in the last few years, his filter has been breaking down. Roy Schulze (@thbthttt) December 31, 2017Who, apart from a teenager uses the term haters? Wendy (@WendyWhistles) December 31, 2017he s a fucking 5 year old Who Knows (@rainyday80) December 31, 2017So, to all the people who voted for this a hole thinking he would change once he got into power, you were wrong! 70-year-old men don t change and now he s a year older.Photo by Andrew Burton/Getty Images.'

- After Cleaning

‘donald trump sends embarrassing new year’s eve message disturbing donald trump wish americans happy new year leave instead give shout enemies haters dishonest fake news media former reality show star one job country rapidly grows stronger smarter want wish friends supporters enemies haters even dishonest fake news media happy healthy new year president angry pants tweeted great year america country rapidly grows stronger smarter want wish friends supporters enemies haters even dishonest fake news media happy healthy new year great year america donald j trump realdonaldtrump december tweet went welll expect kind president sends new year greeting like despicable petty infantile gibberish trump lack decency even allow rise gutter long enough wish american citizens happy new year bishop talbert swan talbertswan december one likes calvin calvinstowell december impeachment would make great year america also accept regaining control congress miranda yaver mirandayaver december hear talk include many people hate wonder hate alan sandoval december uses word haters new years wish marlene december say happy new year koren pollitt korencarpenter december trump new year eve tweet happy new year including many enemies fought lost badly know love donald j trump realdonaldtrump december nothing new trump years trump directed messages enemies haters new year easter thanksgiving anniversary pic twitter com daniel dale december holiday tweets clearly presidential long work hallmark becoming president steven goodine sgoodine december always like difference last years filter breaking roy schulze thbthttt december apart teenager uses term haters wendy wendywhistles december fucking year old knows december people voted hole thinking would change got power wrong year old men change year older photo andrew burton getty images'

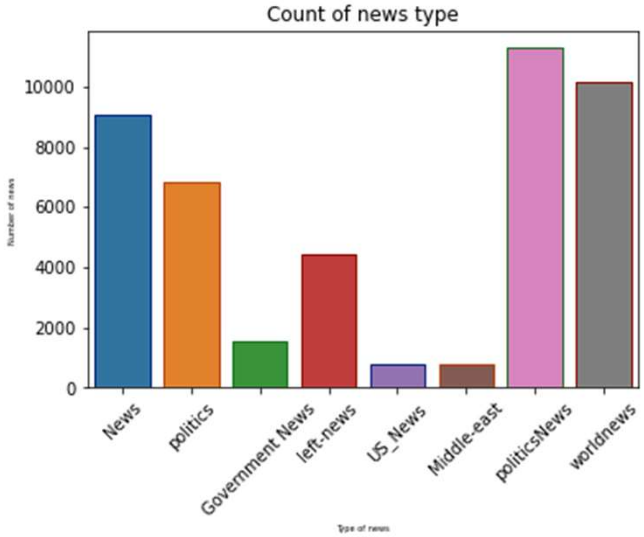
Data Exploration - Publication Date

- Due to the uneven distribution of the publication dates between the two sets it is likely their vocabularies have a bias due to the topics covered



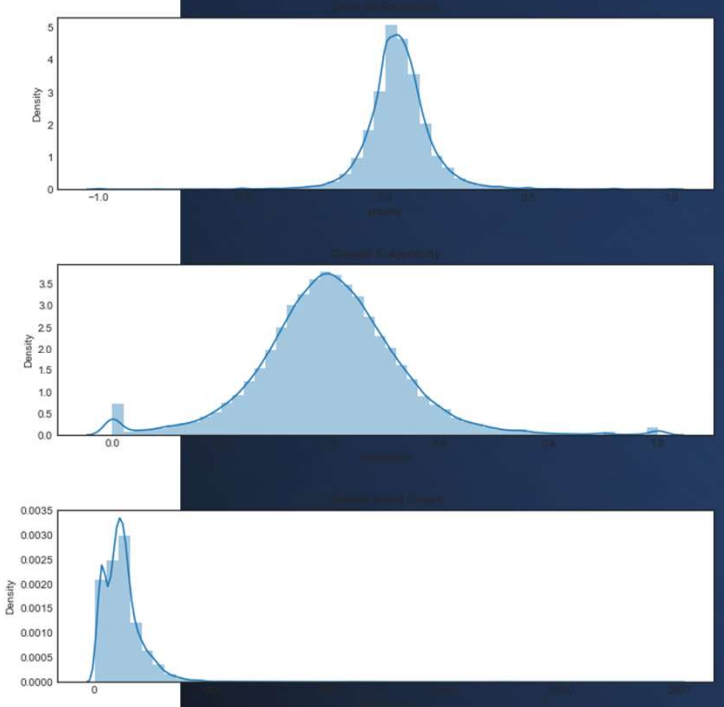
Data Exploration – Count of News Type

- Fake news and the True news files used different classifier names
  - True news (1) only had politicsNews and worldnews
  - Fake news (0) had news, politics, government news, left-news, us_news and middle-east.



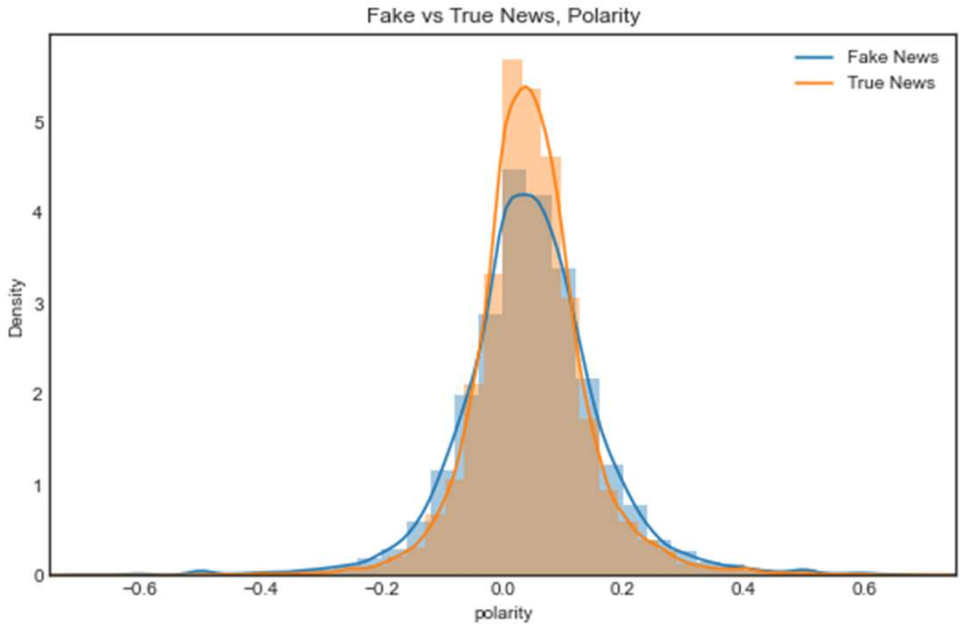
Data Exploration – Sentiment, Subjectivity and Word Count

- Using a Sentiment (polarity) measure we can determine if a text is happy or sad.
- Using a Subjectivity measure we can determine if a text is more objective or subjective
- Counting the number of words



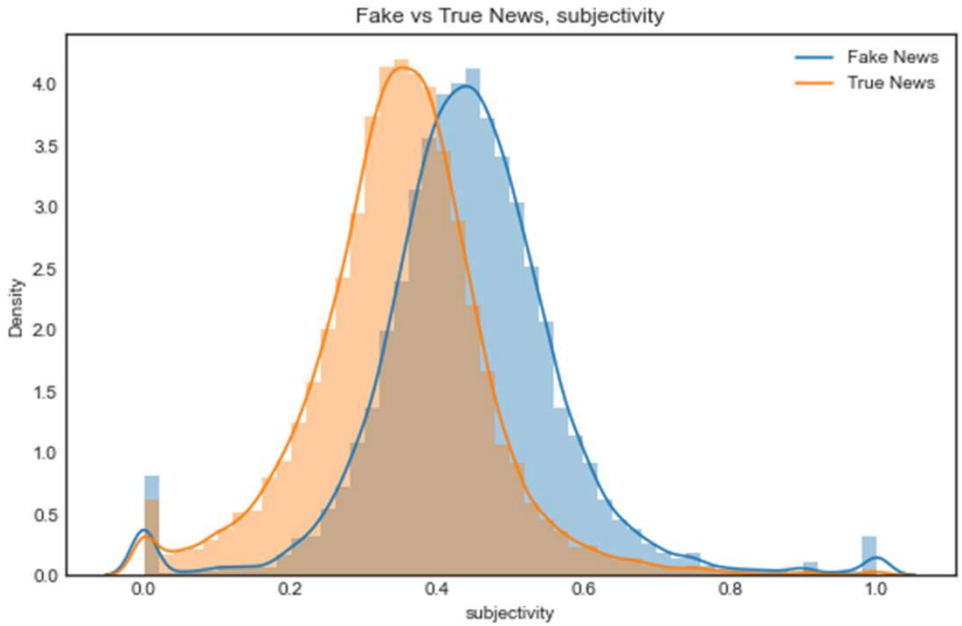
Data Exploration – Sentiment

Both types of news have a slightly positive overall sentiment



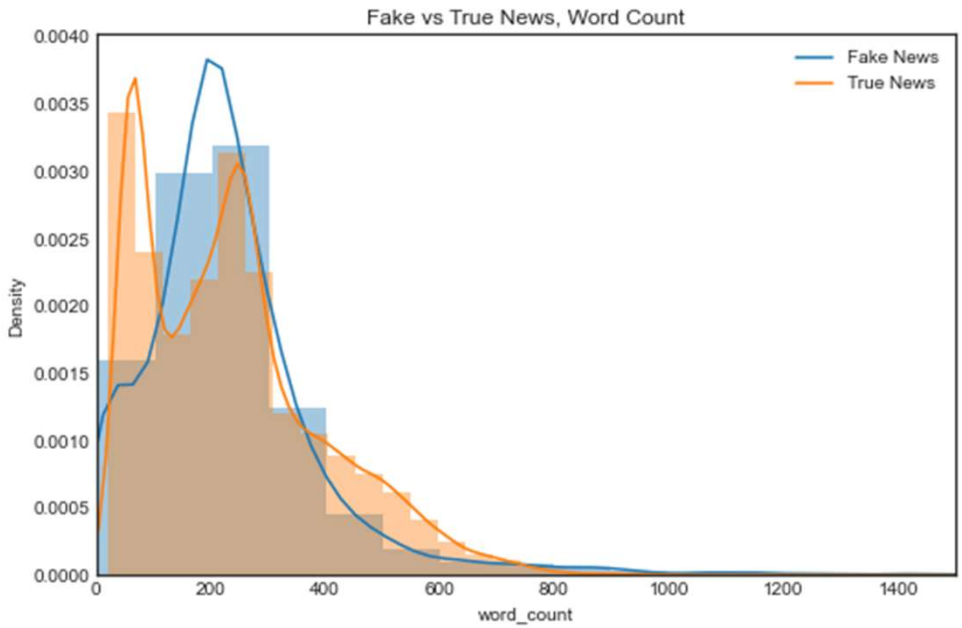
Data Exploration – Subjectivity

Fake news tends to use more subjective language than true news

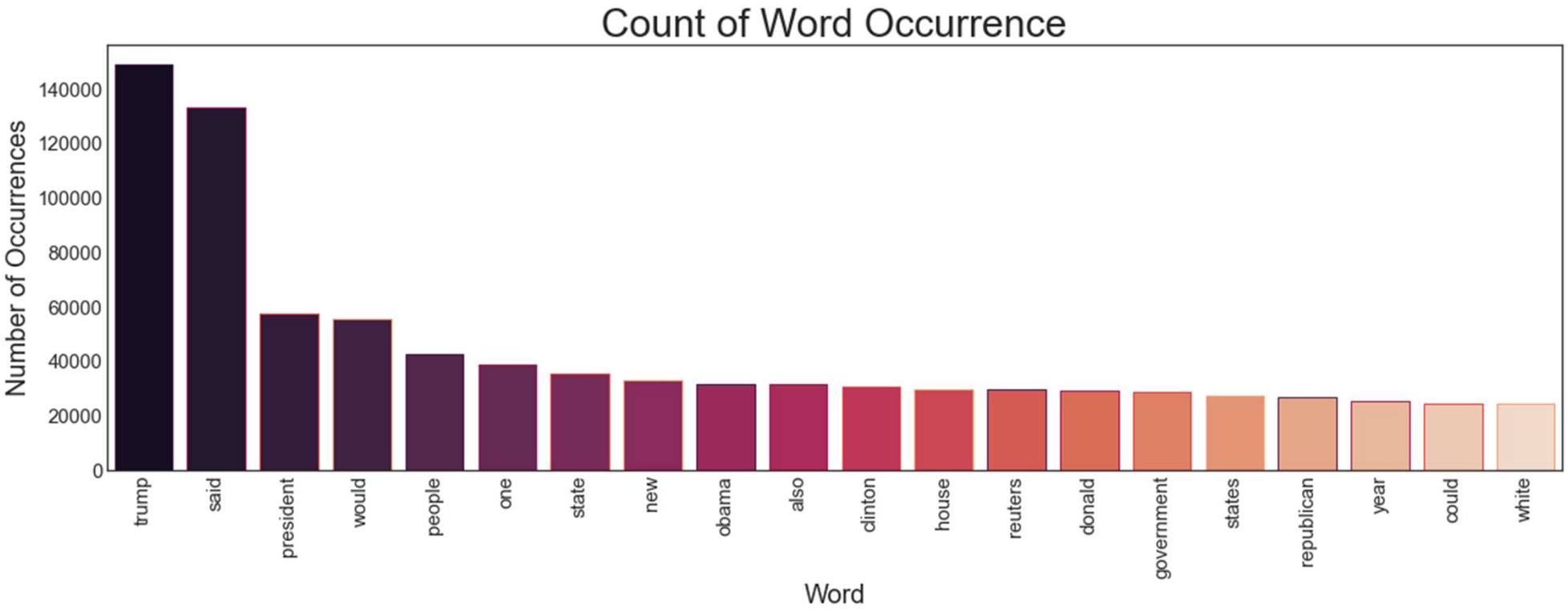


Data Exploration – Word Count

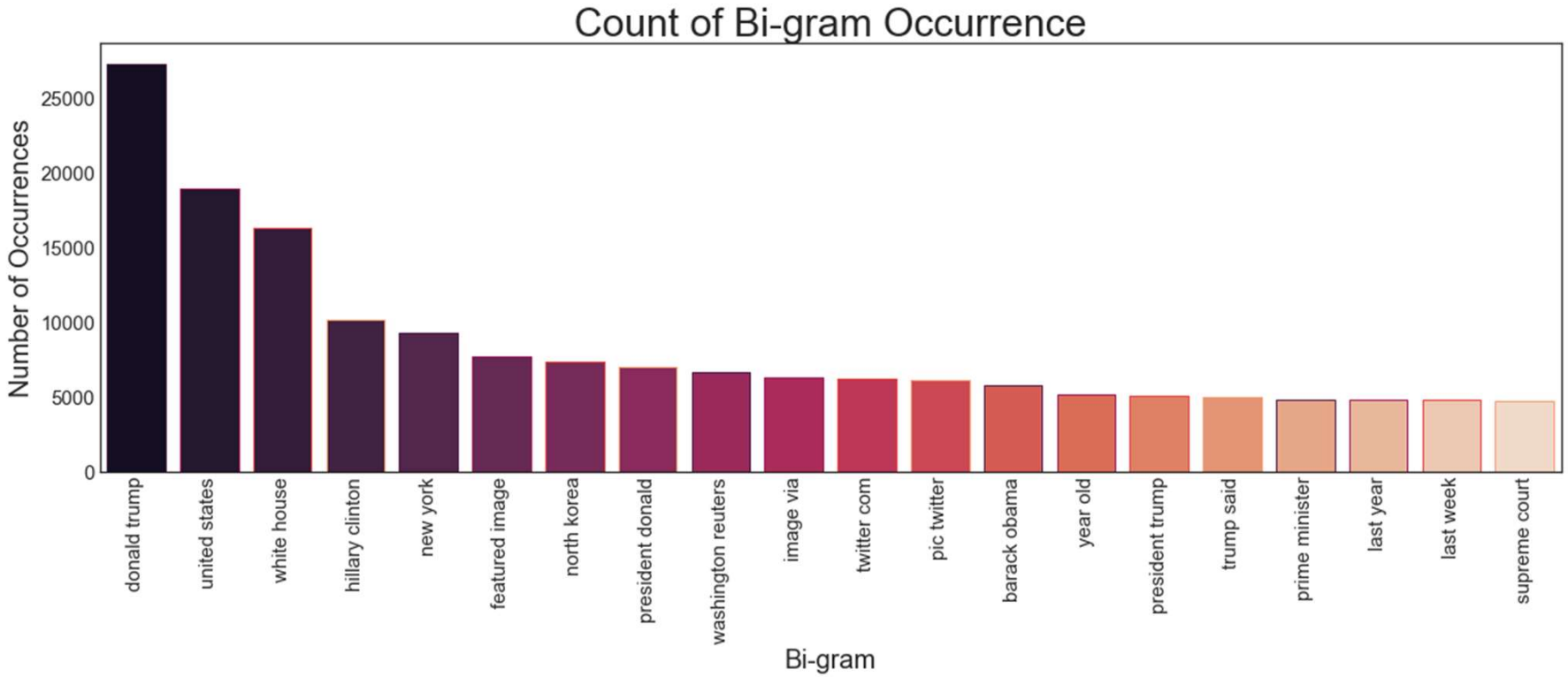
No clear distinction between the fake and true news



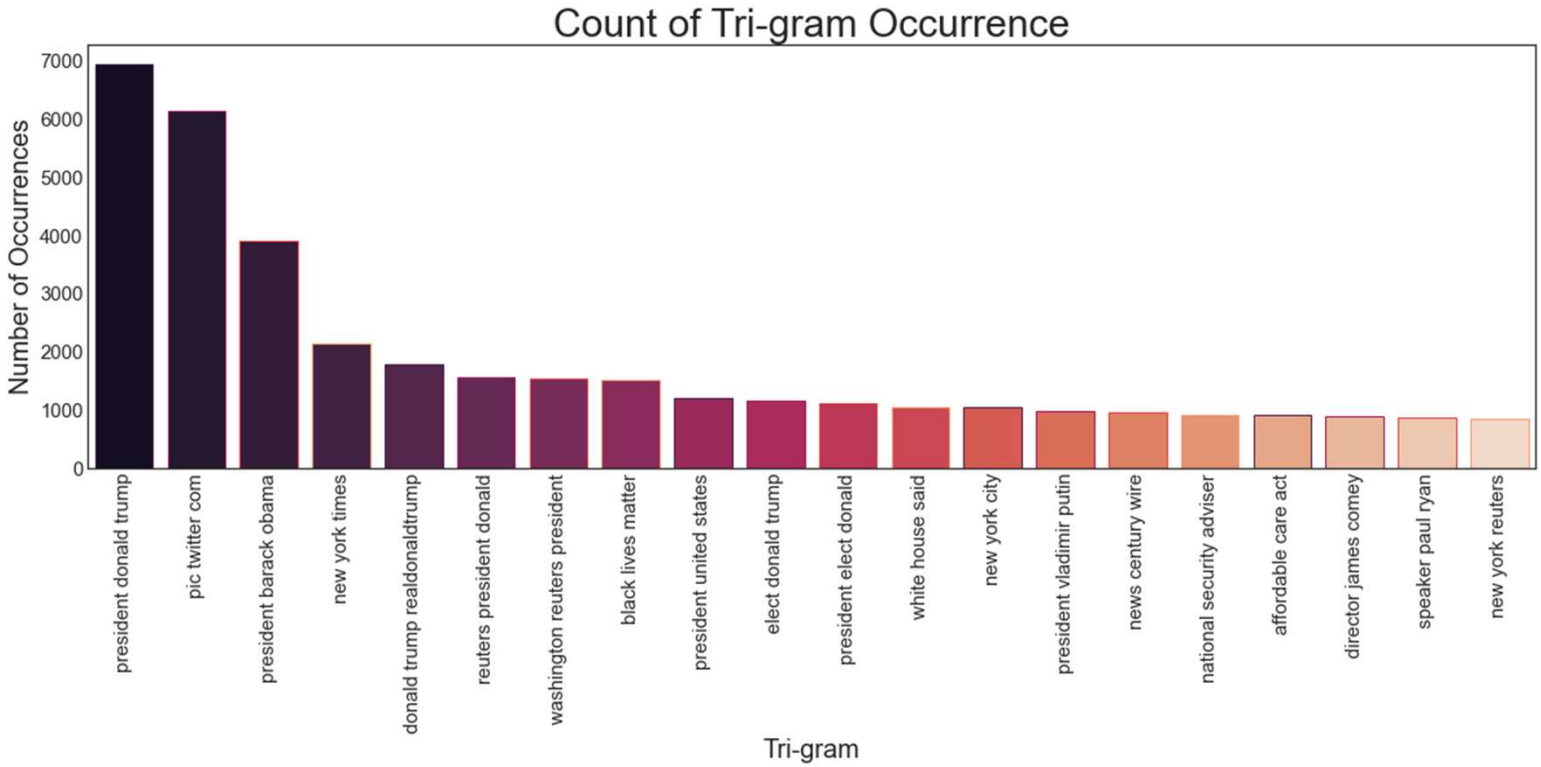
Data Exploration – Common Words



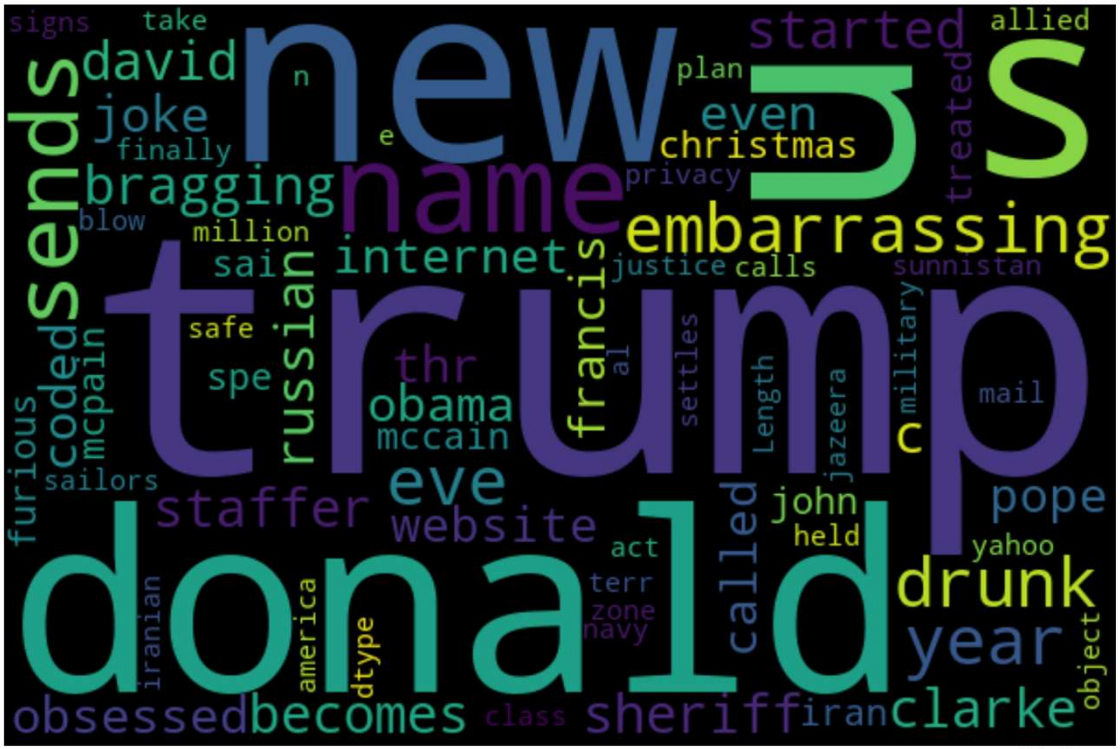
Data Exploration – Bi-grams



Data Exploration – Tri-grams

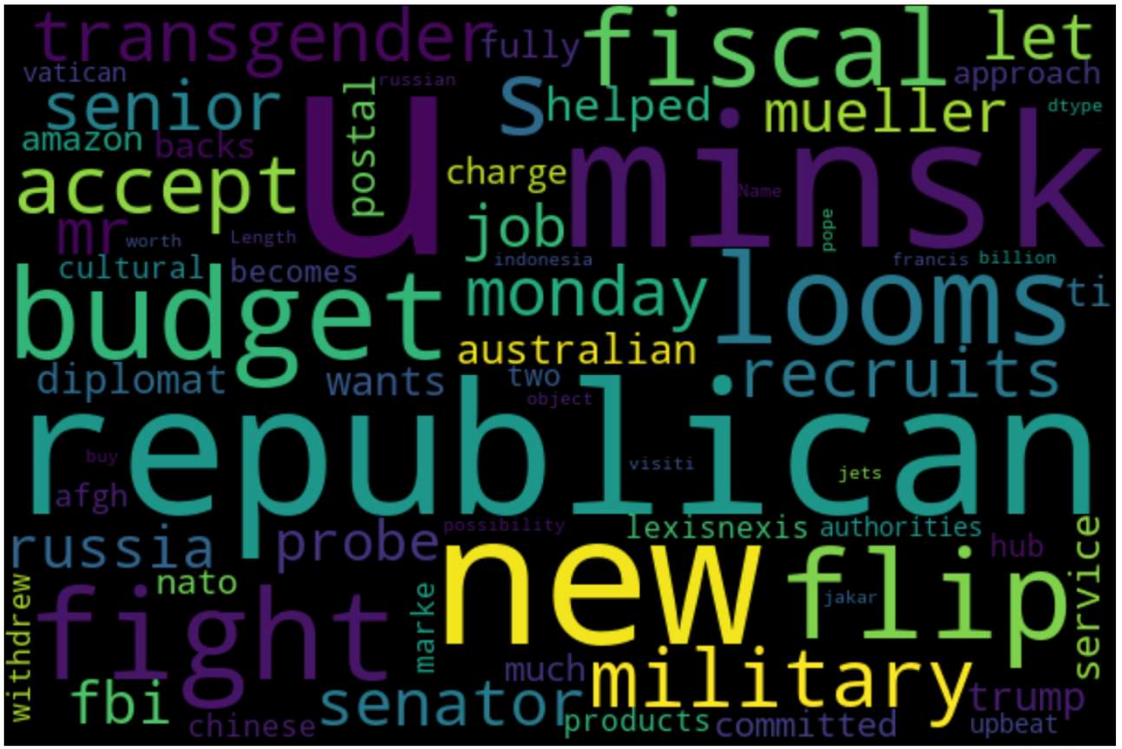


WordClouds – Fake News



Fake News

WordClouds – True News



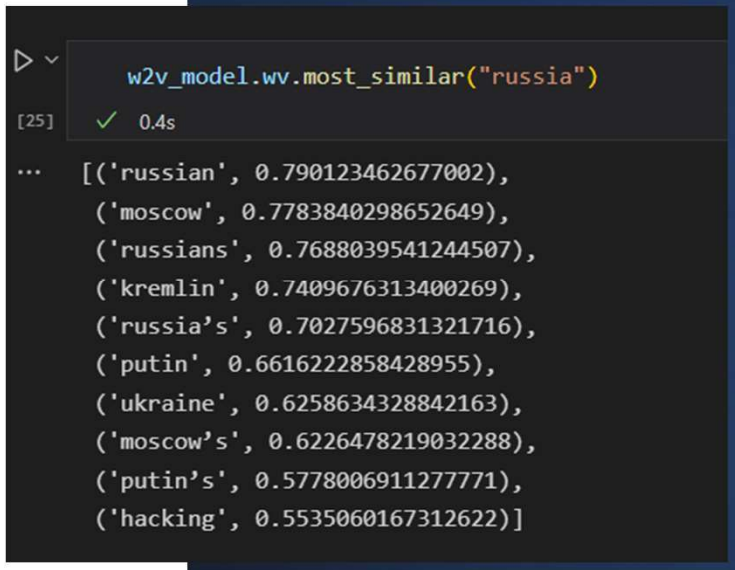
True News

Model Building - Long Short-Term Memory(LSTM)

- What is an LSTM?
  - A type of Recurrent Neural Network (RNN)
    - RNNs have internal state which can be used to represent contextual information.
    - RNNs keep information about past inputs of an amount of time which is defined by the weights of the input data
  - The LSTM architecture aims to provide a short-term memory for a RNN that can last thousands of timesteps
  - Developed to address the vanishing and exploding gradients seen with standard RNNs

Model Building – Word2Vec

- Word2Vec is a NLP technique which NN model is used to learn word associations from large sets of text.
- Using a cosine similarity function it can indicate the level of semantic similarity between words.
- The similarity is represented by a vector
- This is used as a word embedding in the LSTM to give context to the words



Model Building – Compilation

- Train / Test Ratio - .75 /.25
- LSTM units = 128

```
## Divide the dataset into Train and Test
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.25,
random_state=0)

#Configure the model
model = Sequential()
model.add(Embedding(vocab_size,output_dim=EMBEDDING_DIM,weights =
[embedding_vectors],input_length = maxlen , trainable=False))
model.add(LSTM(units=128))
model.add(Dense(1,activation='sigmoid'))
model.compile(optimizer='adam',loss='binary_crossentropy',metrics=['acc'])
```

- Model Summary

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 embedding (Embedding)       (None, 1000, 100)         12564900

 lstm (LSTM)                 (None, 128)               117248

 dense (Dense)               (None, 1)                 129

=================================================================
Total params: 12,682,277
Trainable params: 117,377
Non-trainable params: 12,564,900
_________________________________________________________________
```

Model Building – Fitting & Accuracy Check

```
model.fit(X_train,y_train,validation_split=0.3,epochs=1)
```

- Training data set
  - Loss: 0.0948, Accuracy: 0.9693
- Validation data set
  - Loss: .0284, Accuracy: .9926

```
y_pred = (model.predict(X_test)>=0.5).astype(int)
accuracy_test = accuracy_score(y_test,y_pred)
accuracy_test
```

- Prediction on the Test data set
  - Accuracy: .9924

Additional Model Testing

- I found a second dataset used for a Kaggle competition on Fake news detection, so I decided to use that to test the model I created
  - https://www.kaggle.com/competitions/fake-news/data

- **Dataset Information**
  id: unique id for a news article
  title: the title of a news article
  author: author of the news article
  text: the text of the article; could be incomplete
  label: a label that marks the article as potentially unreliable
    1: unreliable
    0: reliable

- **Other notes**
  - The new data is formatted differently
  - The labels are reversed from what I trained with
  - Many Null text/title fields
  - I did not check for overlapping entries with my trained data

Additional Testing Results

Accuracy: .739

```
X_test2 = [d.split() for d in clean_test['news'].tolist()]
y_target2 = clean_test['label'].values

X_test2 = tokenizer.texts_to_sequences(X_test2)
X_test2 = pad_sequences(X_test2,maxlen=maxlen)

y_pred2 = (model.predict(X_test2)>=0.5).astype(int)
accuracy_test = accuracy_score(y_target2,y_pred2)
accuracy_test
```

Conclusions

- LSTM did very well when testing against a subset of our original dataset.
- When testing against another dataset, which it wasn't trained. It did better than expected.
  - Visual inspection of some of the first few entries seem to have overlap in topic from the training dataset.
- Some insights on the data itself
  - Trump dominated the topics of news during the 2016 cycle
    - This is shown strongly in the N-gram Analysis
  - Fake news shows more subjectivity
    - This is shown when the subjectivity feature was created

Future Work and Additional Information

- This project could be improved upon by utilizing the derived features
  - Specifically, the use of Sentiment and Subjectivity would likely improve the outcome
- The use of a bidirectional LSTM
  - This has been commented in several papers are the preferred method for NLP

- The notebook for this project can be found at:

https://github.com/blasher565/dm_cs522/tree/main/final%20project/Code

Thank you!

Any Questions?