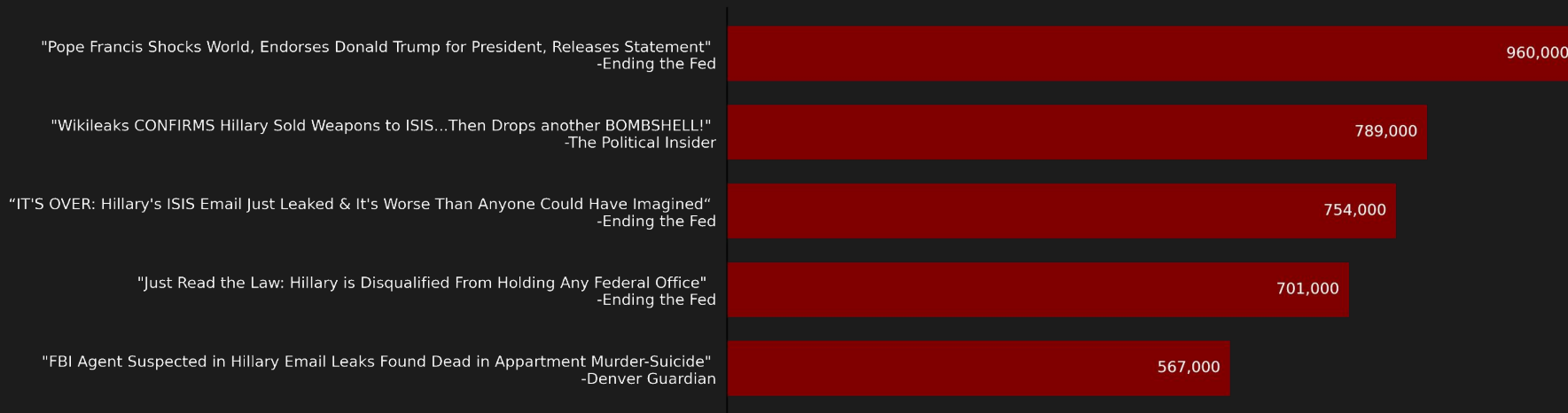


Detecting Misinformation in News Media

Julia Nguyen

● Misinformation is a Real Problem

Facebook Engagements for Top 5 Fake Election Stories (Aug-Nov 2016)



● Misinformation is a Real Problem

Total Facebook Engagement for Top 20 Election Stories (Aug - Nov 2016)



● Goals

- Identify misinformation
- Determine misinformation topic clusters

● Workflow



Data: Scraped from
Various News
Outlets

Labels: Ad Fontes
Media and ASU

Logistic Regression
Model to classify
news as real or
misinformation

Perform Latent
Dirichlet Allocation
to find
misinformation
topic clusters

App:
- Classify news
articles
- Reports recent
misinformation
topic clusters

● Workflow



Data: Scraped from
Various News
Outlets

Labels: Ad Fontes
Media and ASU

Logistic Regression
Model to classify
news as real or
misinformation

Perform Latent
Dirichlet Allocation
to find
misinformation
topic clusters

App:
- Classify news
articles
- Reports recent
misinformation
topic clusters

● Workflow



Data: Scraped from
Various News
Outlets

Labels: Ad Fontes
Media and ASU

Logistic Regression
Model to classify
news as real or
misinformation

Perform Latent
Dirichlet Allocation
to find
misinformation
topic clusters

App:
- Classify news
articles
- Reports recent
misinformation
topic clusters

● Workflow



Data: Scraped from
Various News
Outlets

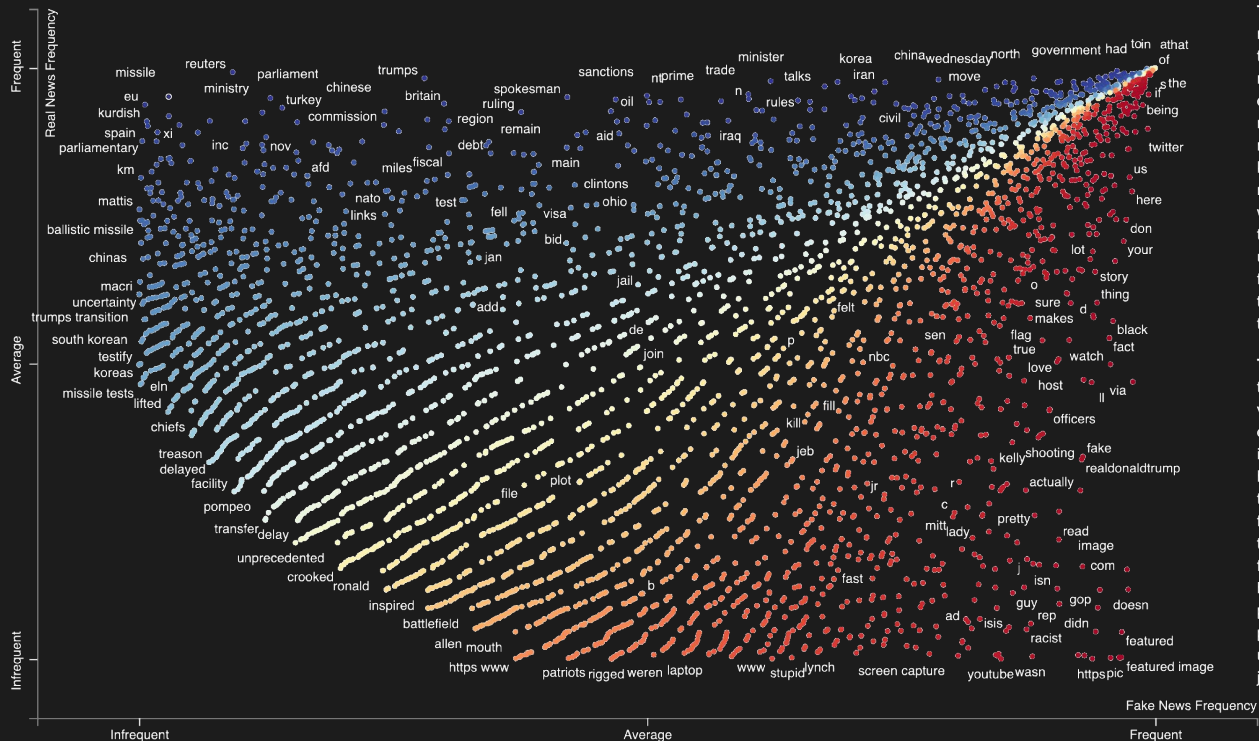
Labels: Ad Fontes
Media and ASU

Logistic Regression
Model to classify
news as real or
misinformation

Perform Latent
Dirichlet Allocation
to find
misinformation
topic clusters

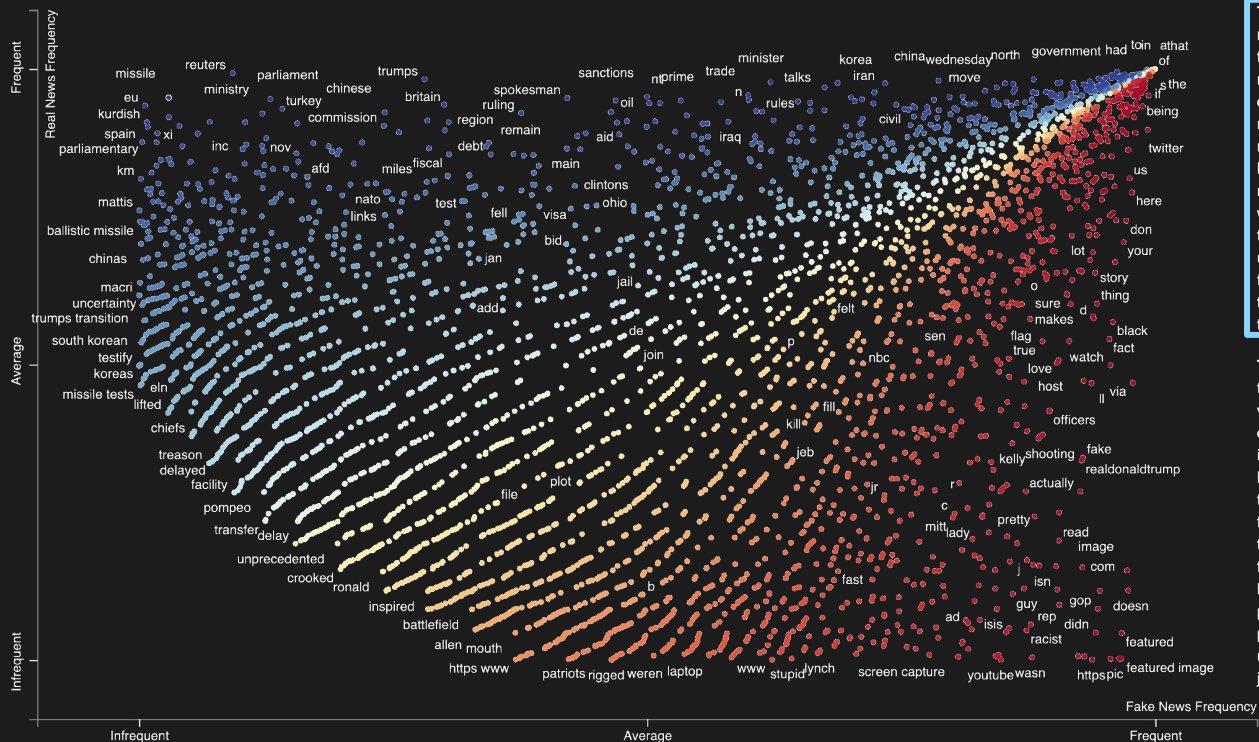
App:
- Classify news
articles
- Reports recent
misinformation
topic clusters

● Exploratory Data Analysis



Top Real News	Characteristic
reuters	twitter
trumps	trump
minister	obama
china	comey
nuclear	trumps
north	realdonaldtrump
korea	barack
north korea	obamacare
wednesday	tweeted
tuesday	obamas
u	facebook
foreign	tweet
tax	hillary
thursday	manafort
	breitbart
Top Fake News	
tillerson	
t	brexit
via	macron
com	youtube
image	duterte
here	tweets
twitter com	megyn
featured	rubio
twitter	mcmullin
featured image	gorsuch
hillary	clintons
pic	rohingya
pic twitter	potus
us	cuomo
just	reuters

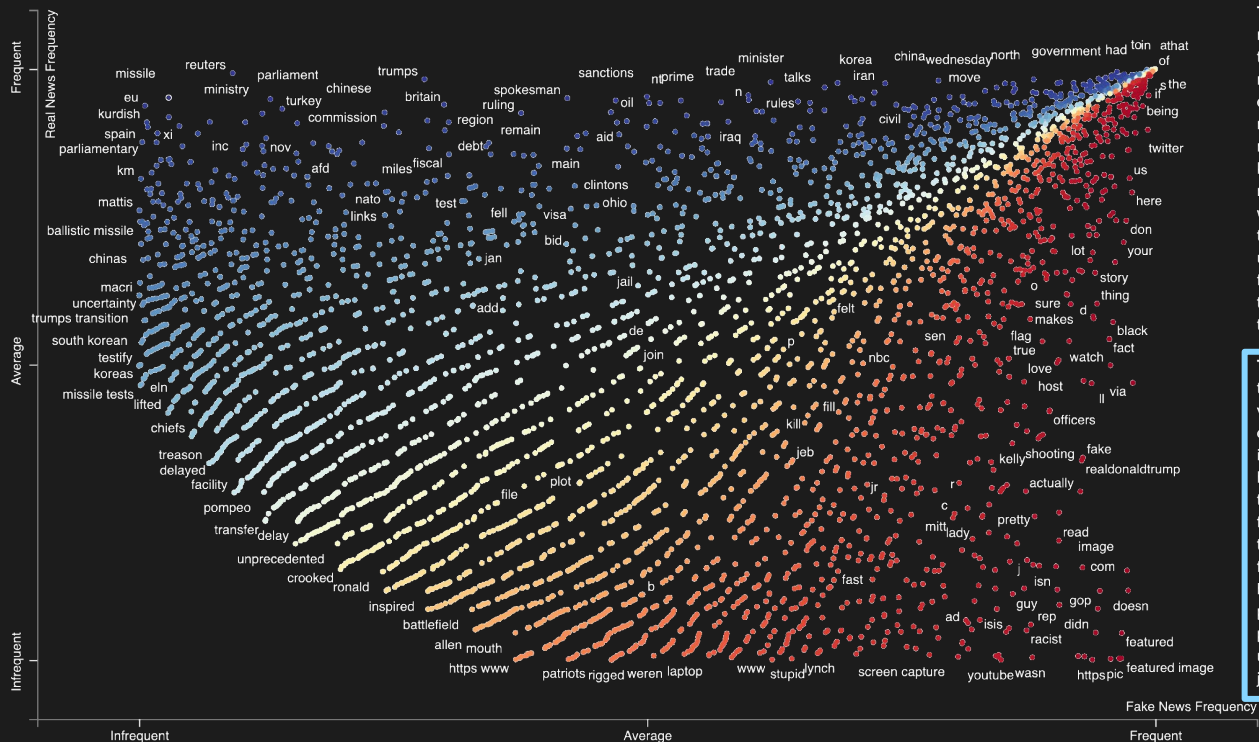
● Exploratory Data Analysis



Top Real News	Characteristic
reuters	twitter
trumps	trump
minister	obama
china	comey
nuclear	trumps
north	realdonaldtrump
korea	barack
north korea	obamacare
wednesday	tweeted
tuesday	obamas
u	facebook
foreign	tweet
tax	hillary
thursday	manafort

Top Fake News	
tillerson	
t	brexit
via	macron
com	youtube
image	duterte
here	tweets
twitter com	megyn
featured	rubio
twitter	mcmullin
featured image	gorsuch
hillary	clintons
pic	rohingya
pic twitter	potus
us	cuomo
just	reuters

● Exploratory Data Analysis



Top Real News

reuters
trumps
minister
china
nuclear
north
korea
north korea
wednesday
tuesday
u
foreign
tax
thursday

Characteristic

twitter
trump
obama
comey
trumps
realdonaldtrump
barack
obamacare
tweeted
obamas
facebook
tweet
hillary
manafort
breitbart

Top Fake News

t
via
com
image
here
twitter.com
featured
twitter
featured image
hillary
pic
pic twitter
us
just

tillerson
brexit
macron
youtube
duterte
tweets
megyn
rubio
mcmullin
gorsuch
clintons
rohingya
potus
cuomo
reuters

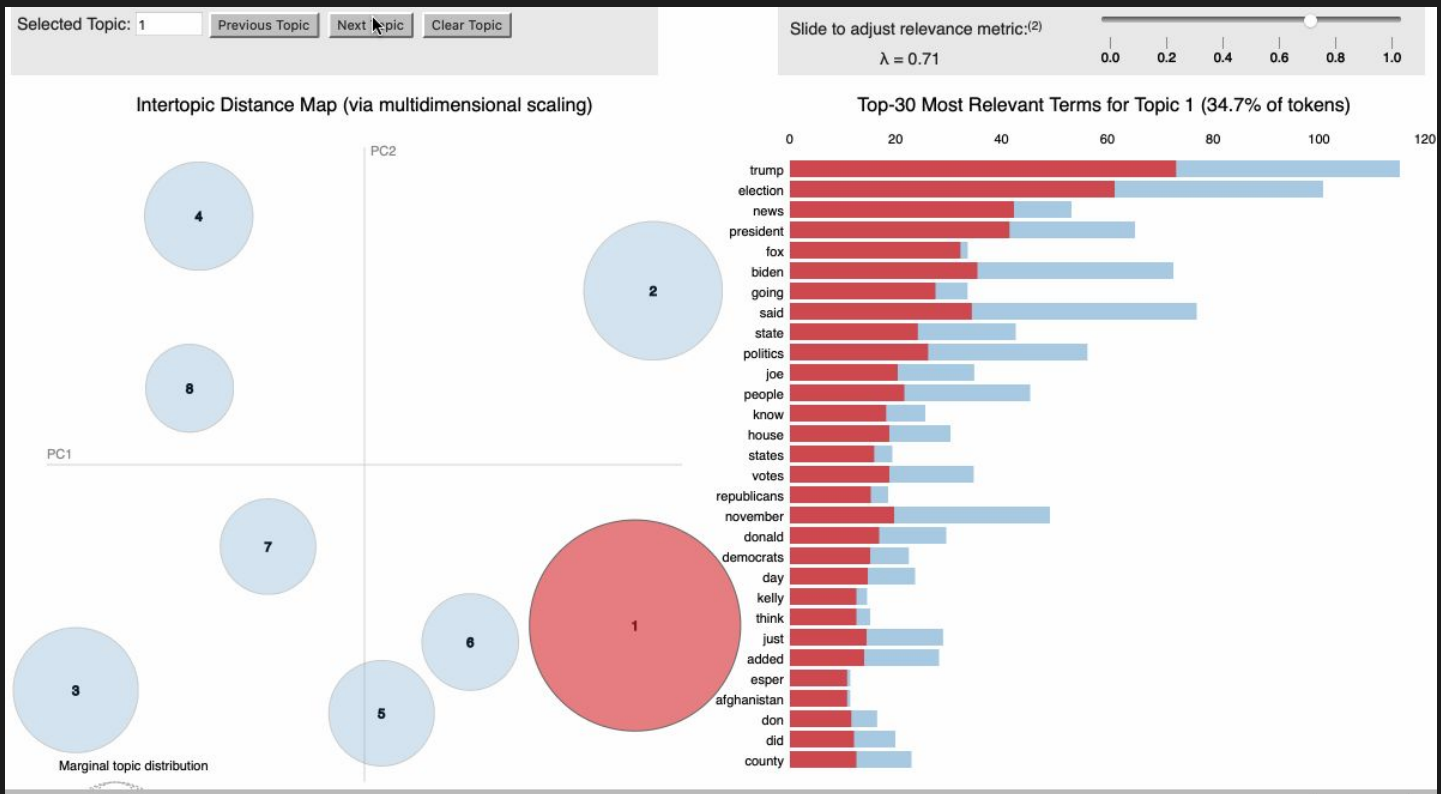
● Logistic Regression Model

	Precision	Recall
Fake News	0.71	0.78
Real News	0.89	0.85

Today's Misinformation:

1. "AL Senator-Elect Tuberville: Joe Biden is a 'Socialist' - 'Can't Wait to Get His Hands on Your Money'" -Breitbart
2. "Schiff and CNN Spread Disinformation about Hunter Biden Scandal: Greenwald" -FoxNews.com

● Topic Clustering



● Top Categories: Then and Now

2016-2017

1. Donald Trump
2. Race relations/Police
Violence
3. Hillary Clinton
4. Foreign Policy
5. Law and Order

Present

1. Trump/2020 Election
2. Democrats
3. Foreign Policy
4. Race relations
5. Religion



Navigation

Go to

- ☒ Misinformation Classification
- ☐ Misinformation Clusters

Misinformation Classification

Please enter the URL of the article you are interest in

Press Enter to apply

Start

Made with Streamlit



Navigation

Go to

- ☒ Misinformation Classification
- ☐ Misinformation Clusters

Misinformation Classification

Please enter the URL of the article you are interest in

Start

Made with Streamlit

● **Future Work**

Improve linear regression model

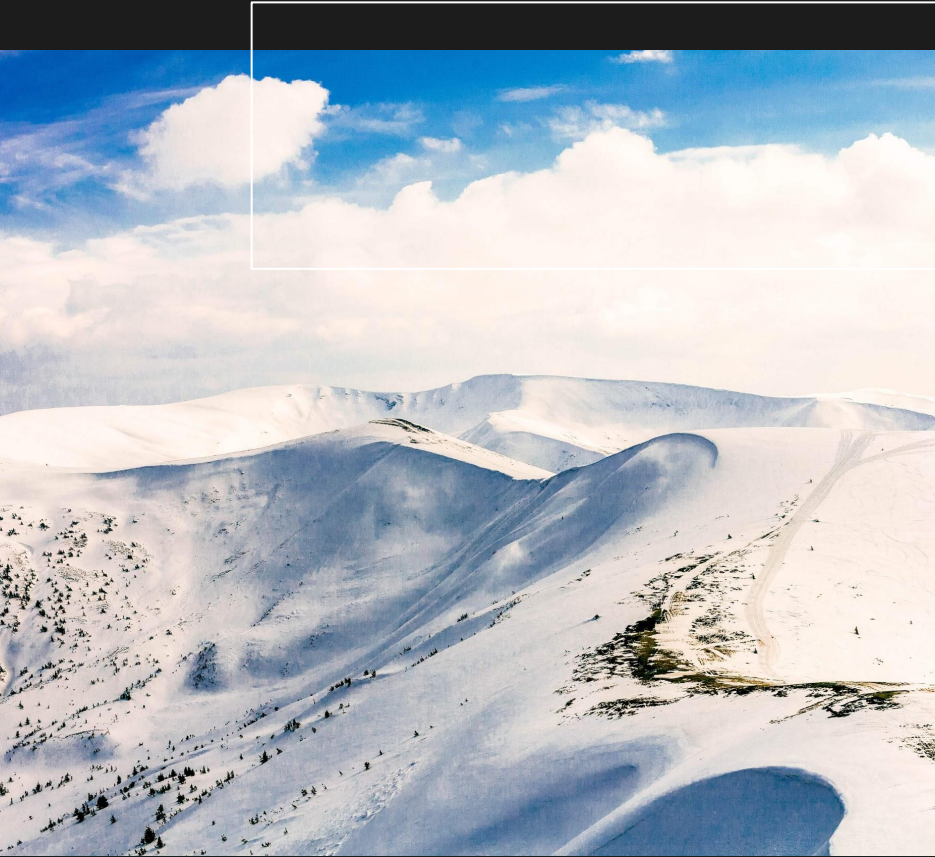
- More and better data
- Incorporate sentiment analysis

Look at misinformation topic clusters at different time periods

- **Questions?**

● Appendix

actual_negative	173	49
actual_positive	72	394
	predicted_negative	predicted_positive



Thanks •



CREDITS: This presentation template was created by [Slidesgo](#), including icons by [Flaticon](#), and infographics & images by [Freepik](#).

Please keep this slide for attribution.