Topic Modelling on Powerful Women: Twitter Analysis

Fatima Loumaini- Metis Summer 2019

Introduction

- Women are by far the biggest influencers on Twitter

- Dataset including 11 women from different backgrounds: liberals, conservatives, politicians, actresses and tv personalities

- Data scraped from January 2018

Project Goal

- Understand what are the topics these women are interested in

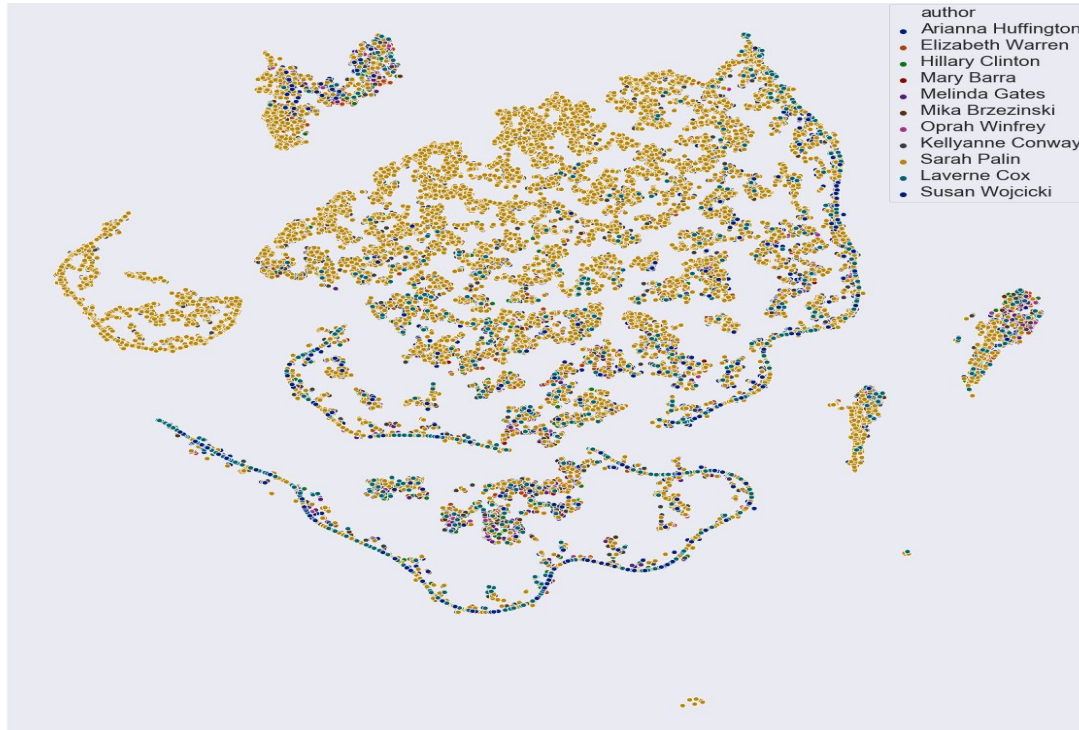- Find similarities and differences amongst a specific set of women

Methodology

- Twint package for data scraping from Twitter

- Each tweet is considered as a document

- TFIDF for word frequency in the documents

- NMF for dimensionality reduction

- t-sne for visualization in 2D

Most discussed topics on the dataset

	Topic 1: Facebook	Topic 2: Women	Topic 3: Communication	Topic 4: Politics	Topic 5: Mental Health
0	posted new	great	think	trump	need
1	photo facebook	love	wow	president	know
2	posted new photo	thank	wow think	just	people
3	new photo facebook	wa	did	president trump	need know
4	new photo	time	say	donald	help
5	posted	woman	dont think	donald trump	people need
6	facebook	day	did think	hillary	health
7	photo	just	just think	clinton	dont
8	new	today	just	obama	right
9	facebook album	good	dont	says	dont need

Tweets by authors in 2D space

author
- Arianna Huffington
- Elizabeth Warren
- Hillary Clinton
- Mary Barra
- Melinda Gates
- Mika Brzezinski
- Oprah Winfrey
- Kellyanne Conway
- Sarah Palin
- Laverne Cox
- Susan Wojcicki

How Sarah Palin and Laverne Cox compare?

Most used Words:

- ● Sarah Palin: Alaska(prob 10%)

- ● Laverne Cox: trans (prob 91%)

Most dominant topics:

- ● Sarah Palin: Politics

- ● Laverne Cox: LGBTQ+ rights

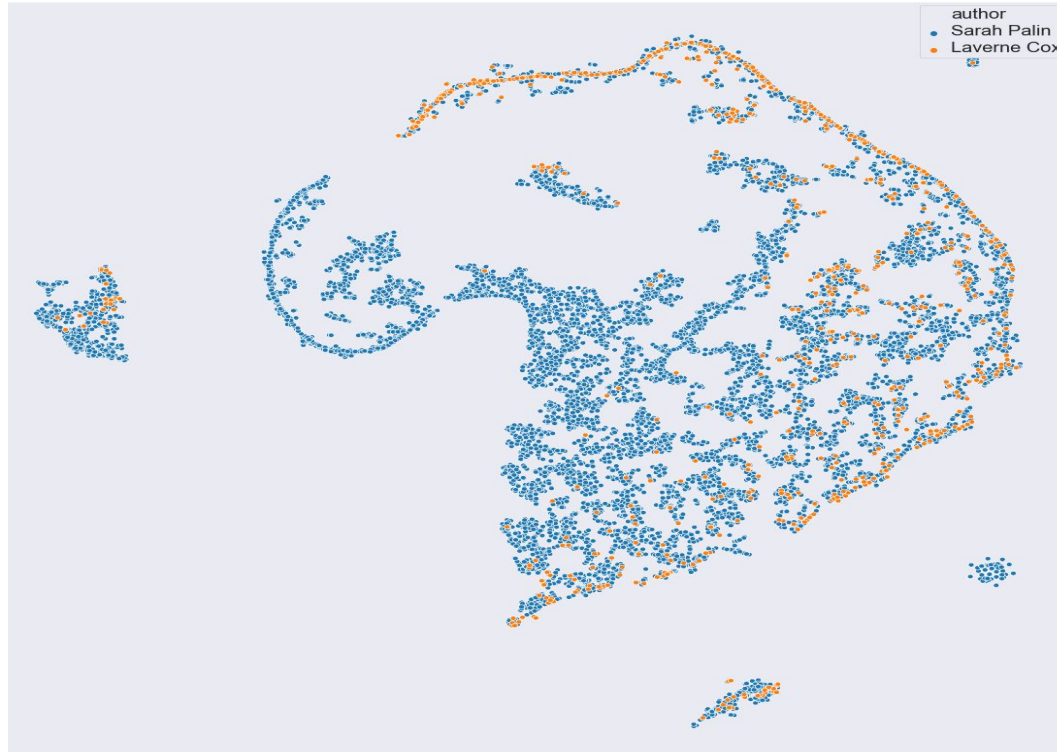# How Sarah Palin and Laverne Cox compare?

Common Words/Topics:

- Words: N/A

- Topics: Celebration

Opposite Words Topics:

- Words: trans (91% Cox, 0% Palin)

- Topics: LGBTQ+

How Sarah Palin and Laverne Cox tweets compare in 2D space?

Conclusion

- Women from different backgrounds have a lot in common

- Potential to analyze how words are used in different topic spaces

- Potential for a supervised NLP to recommend to a common woman(aka someone like me) on what words to use to sound like a powerful women

- Build a FlaskWeb App to make it available to the public

Thank you