



#MeToo, a term coined by social activist Tarana Burke, to promote "empowerment through empathy" among sexual assault and harassment survivors, became a widespread movement on social media when actress Alyssa Milano encouraged spreading this hashtag to attempt to draw attention to workplace sexual violence. With more than 500,000 retweets and 12 million posts in Facebook within first 24 hours, #MeToo soon became a voice strong enough to bring down 201 Powerful Men according to The New York Times. Through this project we intend to analyze the virality of data from Data Science perspective and add value to the social cause behind this movement.

Goal

We have often observed that disrespectful behavior towards women is due to lack of awareness or due to certain disbeliefs in the society.

Our Goal: Suggest Targeted Approach to spread awareness and help in the society.

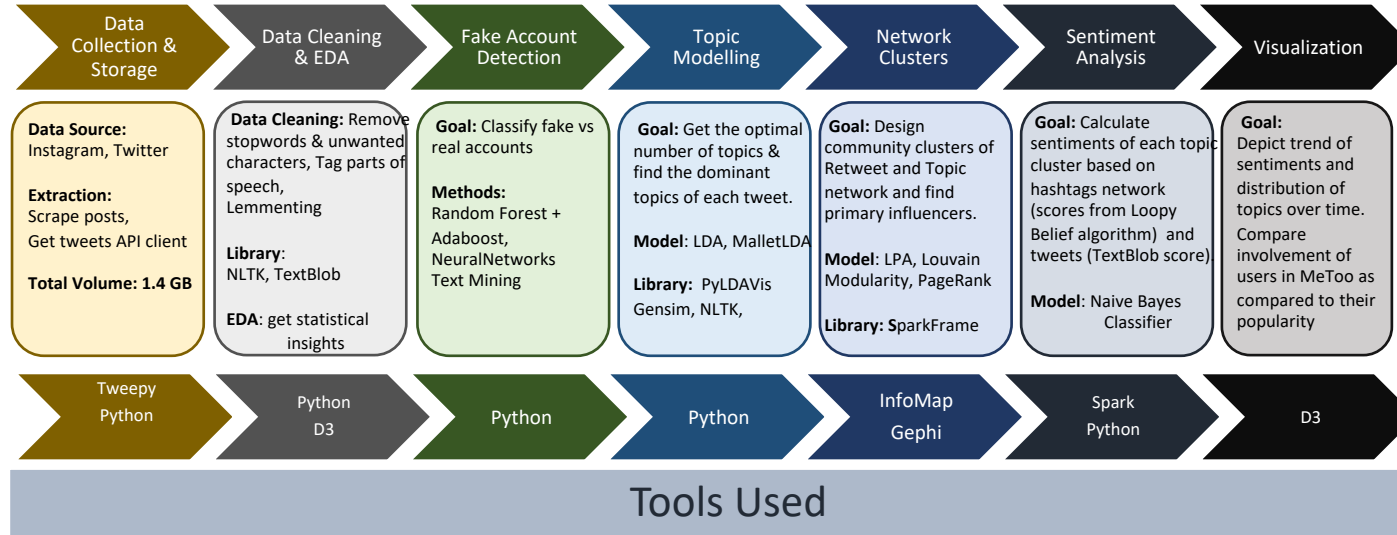
Questions we intend to answer:

- Can we identify who is real and who's fake
- Can we find any similarity in beliefs and emotions of different users and group them together
- Who is more involved in this movement.
- What are the different countries and languages that became part of #MeToo
- How can we put together all these information and spread awareness in a faster and targeted way
- Can we learn something new from technology perspective.

Future Goal

- Analyzing videos and real time data streams
- Tuning the models to have higher accuracy
- Experiment with other variants of the algorithms
- Have further insights like how many police reported cases were observed before and after the #MeToo movement
- Visualize geolocations and tag nearby NGOs near the major cities.
- Social media sites have introduced several restrictions, get enterprise access to carry out deeper analysis.

Data Science Pipeline



Tools Used

Resolution

Our analysis attempts to

- Collectively analyze multiple social media platforms and identify real accounts
- cluster people based on the common topics that they are discussing,
- find the main influencers of each of these clusters and
- analyze the sentiments of these clusters

Our Product would help the government or the NGOs easily identify group of people with similar beliefs and rapidly spread awareness or broadcast legal information/helpline numbers to the intended clusters through the influencers

Key Features: An interactive web App depicting

- trend in sentiments over time
- Topic distribution
- Statistical insights

Learnings

Lessons Learnt:

- Social network analysis metric like Modularity described how many sub-communities can be detected. The more sub-communities and smaller they are, the slower viral content spreads
- Using message propagation algorithm to determine sentiments from hashtags gave additional insights
- Apply filter to ignore words judiciously to get better topic clusters.
- Text mining performed better than ML algorithms for identifying Fake accounts

Results

