Coronavirus tweets Analysis_MVP

Xufei Li

Part 1:

Applied Topic Modeling on March 2020 English tweets to get top 10 topics:

Topic1: Panic buying & Stocking food
Topic2: grocery & retail
Topic3: online shopping & delivery
Topic4: gas price decline
Topic5: Toilet paper hoarding
Topic6: supermarket shelves are empty - Food shortage
Topic7: Mask & Hand sanitizer shortage
Topic8: Thanks ALL front-line workers
Topic9: Covid impact on business & economy
Topic10: StayAtHomeSaveLives?

Then compare the top 10 topic for each country - I chose 'United States', 'Italia', 'United Kingdom' (3 countries with most tweets) for this comparison

Part 2:

Using results from Part 1 to group similar tweets to 6 clusters(chose K based on inertias & silhouettes coefficient). Label the clusters with "group topic"

What to do next?

Apply topic modeling & clustering to May 2021 coronavirus tweets(obtained from twitter API). Then compare results with March 2020. Visualize result.