

CSC110 Project Proposal: Climate Change Sentiment on Twitter

Harsh Jaluka, Ronit Kumar, Thomas Liu, Wilson Sy

Friday, November 6, 2020

Problem Description and Research Question

Evidence of climate change and its potential harms date as far back as the early 19th century. Since then, after decades of research, the vast majority of scientists have reached the conclusion that the increase in global temperature is, in large part, due to human activities; particularly, the increase in greenhouse gas emissions after the industrial revolution.

Despite the evidence for anthropogenic global warming, or man-made climate change, there is a lack of agreement about the very existence of climate change among the general public. There is a vocal group of people around the world that denounces the legitimacy or the severity of climate change. These people may be motivated by political or economic gain, or they may themselves be victims of misinformation disseminated by social media.

Research shows that polarization can be attributed to influencers, whose views are often amplified over others in a network (Centola, 2020). In political discourse on Twitter, such influencers can be major politicians: a study about politicians in the 2016 U.S. election found that tweets with “more emotive and moral words” get retweeted more often (Brick, Linden, & De-Wit, 2019). Thus, in a race to gain more popularity, politicians often intensify their rhetoric, whether they believe or disbelieve in climate change. Moreover, Ezra Klein from Vox Media suggests that politicians’ dependence on social media to read the will of the people may polarize politicians’ opinions, leading to more unsatisfying choices when the public is faced with polarizing public policy that aims to tackle climate change (Newton, 2020).

Climate change is a serious, pressing issue. Consensus among the public about its legitimacy is vital to allow for concrete action against it. In our project, we want to study the intensity of the rhetoric between opinions for and against climate change by doing sentiment analysis on tweets. We are hoping to determine the extent to which the debate surrounding climate change has been polarized, primarily by comparing opinion-based tweets with neutral fact-based tweets. Thus, our research question is as follows:

How polarized is the debate surrounding climate change on Twitter?

Dataset Description

We are using one dataset, a `.csv` file from Kaggle by Edward Qian, a student from the University of Waterloo. Each row contains a tweet, the id of the tweet, and a sentiment value for that tweet. The sentiment value for each tweet is one of the 4 values listed below.

-1 : The contents of the tweet do not support the idea of man-made climate change

0 : The contents of the tweet are neutral on idea of man-made climate change.

1 : The contents of the tweet support the idea of man-made-climate change.

2 : The contents of the tweet link to factual news on man-made climate change.

The dataset has 43943 tweets on climate change, collected from April 27, 2015 to February 21, 2018. It has been approved by 3 reviewers.

Example row:

```
1, '‘Fabulous! Leonardo #DiCaprio’s film on #climate change is brilliant!!! Do watch.  
https://t.co/7rV6BrmxjW via @youtube’’, 793124402388832000
```

Computational Plan

New Library

We intend to use `vaderSentiment` as the new library in this project. It is a sentiment analysis library that measures the positivity, neutrality and negativity of each string. It also gives a compound value based on the these three values.

Analysis

Tweets and their sentiment values will be read from the `.csv` file and stored in a list of lists, with each inner list containing the three elements corresponding to the description of the tweets above. We will first count the frequencies of each sentiment value.

We will then use the `vaderSentiment` library to find the compound values of all the neutral and news-based tweets. Using these values, we will determine a range of compound values that is considered ‘normal’ for a climate change tweet. We are interested in seeing how this range compares to the suggested range of neutral strings, which is between -0.05 and 0.05.

Having done that, we then plan to find the compound values for the opinion-based tweets. The range determined in the previous part will serve as a cutoff of what is a neutral tweet. We want to see what the frequency of the positive, neutral and negative tweets are for both the ‘deniers’ and the ‘supporters’ of climate change.

We will also store the actual compound values calculated for each tweet in each sentiment value. The frequency of positive, neutral and negative tweets will serve as a rough understanding of the situation but our main interest lies in the actual compound value of each tweet. We wish to see the extent of positivity, neutrality or negativity in opinion-based tweets.

Presentation

Using `plotly`, we will plot multiple graphs to display all our conclusions. We will have a bar graph for the number of tweets per sentiment value and the number of positive, negative, and neutral tweets for sentiment values -1 and 1. We will then use a histogram to show the distribution of the compound values of the opinion-based tweets and the fact-based tweets. The distribution of the compound value of impartial tweets will aid us in picking an appropriate range of neutrality for climate change tweets.

These graphs will be put in a `pygame` application which displays the graph for the sentiment values. The user can click on a sentiment value and it will bring them to the graphs corresponding to that sentiment value.

References

- Brick, C., Linden, S. V. D., & De-Wit, L. (2019, January 16). *Are Social Media Driving Political Polarization?* https://greatergood.berkeley.edu/article/item/is_social_media_driving_political_polarization
- Centola, D. (2020, October 15). *Why Social Media Makes Us More Polarized and How to Fix It*. Scientific American. <https://www.scientificamerican.com/article/why-social-media-makes-us-more-polarized-and-how-to-fix-it/>
- Cook, J., Nuccitelli, N., Green, S. A., et al. (2013, 15 May). Quantifying the consensus on anthropogenic global warming in the scientific literature. *Environmental Research Letters*, 8(2). <https://iopscience.iop.org/article/10.1088/1748-9326/8/2/024024>
- Hutto, C. J. (2020). *VADER-Sentiment-Analysis*. <https://github.com/cjhutto/vaderSentiment>
- Maslin, M. (2019, 30 November). Here Are Five of The Main Reasons People Continue to Deny Climate Change. *Science Alert*. <https://www.sciencealert.com/the-five-corrupt-pillars-of-climate-change-denial>
- NASA. (2020). *Scientific Consensus: Earth's Climate is Warming*. Global Climate Change: Vital Signs of the Planet. <https://climate.nasa.gov/scientific-consensus/>
- Newton, C. (2020, February 28). Why we can't blame social networks for our polarized politics. *The Interface*. <https://www.theverge.com/interface/2020/2/28/21153060/social-network-polarization-ezra-klein-why-were-polarized-q-a>
- Qian, E. (2019). *Twitter Climate Change Sentiment Dataset*. <https://www.kaggle.com/edqian/twitter-climate-change-sentiment-dataset>
- Weart, S. (2017, August 17). *The Discovery of Global Warming [Excerpt]*. Scientific American. <https://www.scientificamerican.com/article/discovery-of-global-warming/>