

Assessing the Veracity of Climate Change Information on Twitter

Luyao(Richard) Wang, Yunlou(Jake) Teng, Hyunuk Kim, PhD, Irena Vodenska, PhD
 Department of Administrative Sciences, Metropolitan College, Boston University,
 1010 Commonwealth Avenue, Boston, MA 02215

Abstract

This research investigates the dissemination of false and misleading information related to climate change on social media, while examining key contributors and influencers responsible for propagating such deceptive content. Disinformation on social media platforms poses a challenge as it is often presented in a complex and convoluted manner, making it difficult to discern. Furthermore, entities with intentions to deceive may collaborate with other organizations to amplify their messages. Our study analyzes climate change-related tweets from 2008 to 2023, utilizing machine learning algorithms to identify and categorize misleading claims based on the CARDS system of climate change misinformation super-claims proposed by Coan et al., 2021. We employ a two-way coding method, comparing cosine similarity with human coding, and achieve approximately 75% overlap between the techniques. We also examine the Twitter activity of organizations funded by ExxonMobil and scrutinize their messaging to identify narratives. Additionally, we pinpoint prominent Twitter accounts that frequently post misinformation and investigate the accounts that propagate this content via retweets, employing a network analysis approach.

Background

In today's digitally connected world, social media platforms have become integral to the dissemination of information, shaping public opinion and discourse on various subjects, including climate change. However, the vast and unregulated nature of these platforms has led to a proliferation of misinformation, with climate change being no exception. As this global crisis intensifies, the spread of misleading and false information on social media has the potential to undermine scientific consensus, hinder public understanding, and stall necessary policy action. It is crucial to identify and combat the sources and spread of climate change misinformation on social media to foster informed decision-making and encourage meaningful action. This research aims to investigate the dynamics of misinformation dissemination on social media, scrutinize the key contributors and influencers perpetuating such falsehoods, and provide insights into the narratives and strategies employed to mislead the public on the issue of climate change.

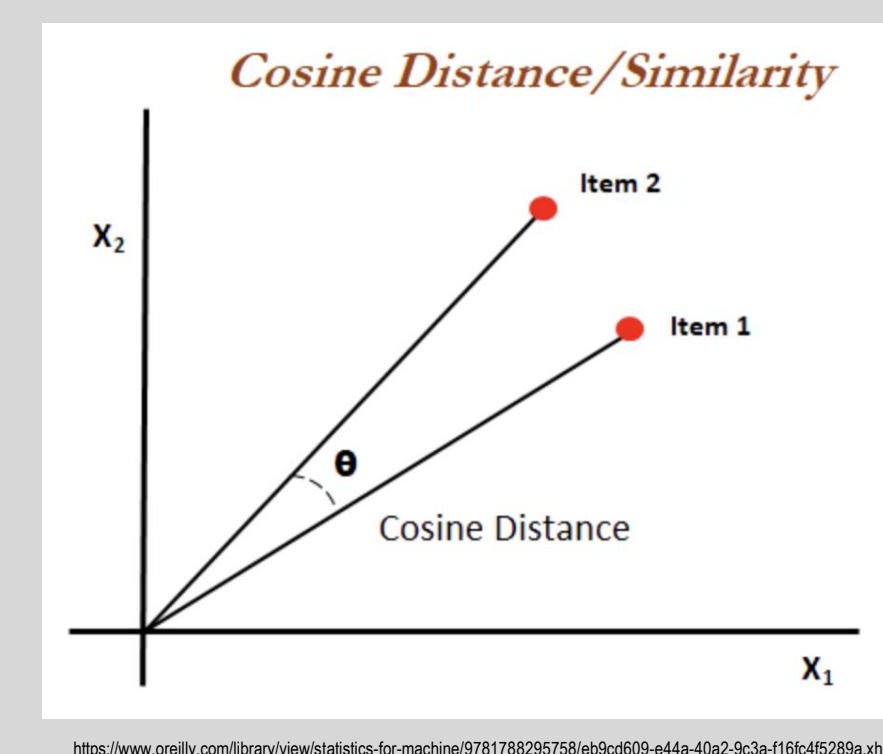


John Ruddick
 @JohnRuddick2

Global warming is not happening.

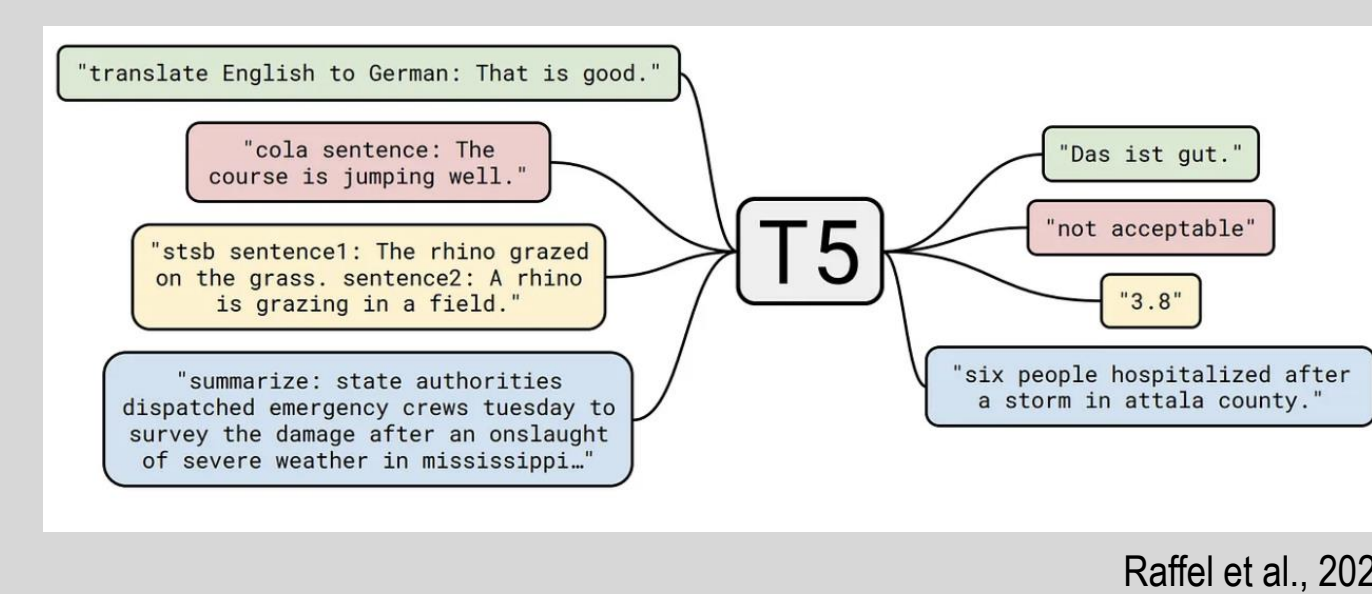
Track 1: Text Similarity Approach

We hypothesized that tweets of which semantics are highly similar to a misinformation claim likely contain the claim in texts. To measure semantic similarity between tweets and claims, we converted all tweets and simplified claims into embedding vectors by using Google's T5 model. This model was pre-trained on the extensive C4 dataset, comprising over 750GB of cleaned and filtered web text from the Common Crawl archive, which allowed T5 to develop a rich understanding of language patterns and structures. The model was subsequently fine-tuned on various downstream NLP tasks, resulting in state-of-the-art performance across numerous benchmarks (Raffel et al., 2020).



Google T5 sentence embeddings

The Google T5 model is a versatile and powerful NLP model based on the transformer architecture, excelling in various tasks like translation, summarization, question-answering, and sentiment analysis. By using a unified text-to-text input-output format and leveraging the C4 dataset for pre-training, T5 achieves state-of-the-art results across multiple benchmarks, making it an effective choice for language understanding and generation challenges.



Results

we computed the cosine similarity between the embeddings of simplified claims and tweets. We selected the top 20 tweets with the highest cosine similarity scores in each subclaim to create an initial version of the survey. For example:

Superclaim	subclaims	Keywords for collecting the tweets	anchors	tweets	Cosine similarity
Global warming is not happening	Heading into ice age	global warming (head OR heading) ice age	heading into ice age	Somebody said something about global warming. Feels like ice age is coming #FreezingCold https://t.co/ITV3yP6JfL	0.8942
Climate solutions won't work	Politics are harmful	climate (policy OR policies) (harm OR harmful)	policies are harmful	Some policies to fight climate change have done more harm than good http://t.co/HFBuVXNDrz	0.8760

Track2: Machine Learning Approach with CARDS

The CARDS (Climate Advocacy, Research, and Discourse Systems) model is a novel approach for systematically categorizing and analyzing climate change discourse based on a hierarchical taxonomy of contrarian claims and counterclaims. By combining

machine learning with expert annotators, the model classifies over 32,000 paragraphs of text from various sources, such as contrarian think tank reports, into super-claims and sub-claims. This comprehensive and fine-grained analysis of climate discourse allows researchers and policymakers to better understand the dynamics of the climate debate and develop effective communication strategies.

Claim Code	Claim label
Global warming is not happening	
1.1	Ice isn't melting
1.2	Heading into ice age
1.3	Weather is cold
1	
1.4	Haban in warming
1.5	Diseases are cooling
1.6	Sea level rise is exaggerated
1.7	Extremes aren't increasing
1.8	Changed the name
Human Greenhouse Gases are not causing global warming	
2.1	It's natural cycles
2.2	Non-Greenhouse Gas forcings
2.3	No evidence for Greenhouse Effect
2.4	CO ₂ not rising
2.5	Emissions not raising CO ₂ levels
Climate impacts are not bad	
3.1	Sensitivity is low
3.2	No species impact
3.3	Not a pollutant
3.4	Only a few degrees
3.5	No link to conflict
3.6	No health impacts
Climate solutions won't work	
4.1	Policies are harmful
4.2	Policies are ineffective
4.3	Too hard
4.4	Clean energy won't work
4.5	We need energy
Climate movement/science is unreliable	
5.1	Science is unreliable
5.2	Movement is unreliable
5.3	Climate is conspiracy

Coan et al., 2021

How do we leverage CARDS?

We applied the CARDS model to a dataset of 680k cleaned tweets from 2007 to 2023, categorizing them into 17 sub-claims. Using the Twitter Academic API and search queries, we gathered 1.44 million tweets and leveraged the original CARDS training set to rebuild the model and assess its performance on this social media dataset.

```
# super claim 1
claim1_1: 'global warming (happen OR happening);
claim1_1_1: 'global warming ice (melt OR melting);
claim1_1_2: 'global warming (head OR heading) ice age;
claim1_1_3: 'global warming weather cold;
claim1_1_4: 'global warming habun;
claim1_1_5: 'global warming (ocean OR oceans) cool OR cooling;
claim1_1_6: 'global warming sea level rise (exaggerate OR exaggerated);
claim1_1_7: 'global warming (extreme OR extremes) (increase or increasing);
claim1_1_8: 'global warming (change OR changed) name;
# super claim 2
claim2: 'human greenhouse (gas OR gases) (cause OR causing) global warming;
claim2_1: 'global warming natural (cycle OR cycles);
claim2_2: 'greenhouse (gas OR gases) (harm OR harmful);
claim2_3: 'evidence greenhouse effect;
claim2_4: 'no2 rise OR rising;
claim2_5: 'emission OR emissions (raise OR raising) co2 (level OR levels);
# super claim 3
claim3: 'climate (impact OR impacts) bad;
claim3_1: 'climate sensitivity low;
claim3_2: 'climate species impact;
claim3_3: 'climate pollutant;
claim3_4: 'climate (degree OR degrees);
claim3_5: 'climate conflict;
claim3_6: 'climate health (impact OR impacts);
# super claim 4
claim4: 'climate (solution OR solutions) work;
claim4_1: 'climate (policy OR policies) (harm OR harmful);
claim4_2: 'climate (policy OR policies) (effective OR ineffective);
claim4_3: 'climate (policy OR policies) hard;
claim4_4: 'clean energy work;
claim4_5: 'need energy (fuel OR nuclear power);
# super claim 5
claim5: 'climate (move OR movement OR science) unreliable;
claim5_1: 'climate science unreliable; # included in the super claim 5
claim5_2: 'climate movement unreliable; # included in the super claim 5
claim5_3: 'climate conspiracy
```

Initial Results

To assess the prediction performance, we chose a sample of 30 tweets from each group, resulting in 540 tweets. We employed the majority rule to label our data, assigning three labels per tweet. So far, we have successfully manually labeled 150 tweets, achieving approximately 60% accuracy.

Next Step

For the Track 1: Upon obtaining our initial dataset, which comprises tweets exhibiting high cosine similarity to the anchor tweets (i.e., potentially containing misinformation), we plan to engage the broader community of Boston University for assistance in the labeling process. By transforming our dataset into a survey, we aim to leverage the collective expertise of the community in identifying tweets that our algorithm deems likely to be misinformation related to climate change. We will initially launch the survey within Metropolitan College, followed by a broader rollout to the entire Boston University community. The outcomes of this survey will offer valuable insights into the quality and characteristics of our dataset, ultimately aiding our understanding of climate change misinformation on social media platforms.

For the Track 2: In order to evaluate the performance of the CARDS model and develop a high-quality dataset, we will persist in undertaking the labeling task. Subsequently, our objective is to construct a classifier capable of detecting misinformation in tweets. To achieve this, we will explore various transformer architectures and employ advanced techniques in order to enhance accuracy and performance.

Reference

Coan, Boussalis, C., Cook, J., & Nanko, M. O. (2021). Computer-assisted classification of contrarian claims about climate change. *Scientific Reports*, 11(1), 22320–22320. <https://doi.org/10.1038/s41598-021-01714-4>

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., ... & Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. *The Journal of Machine Learning Research*, 21(1), 5485–5551.

A., Lee, K., Narang, S., Matena, M., ... & Liu, P. J. (2020). Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. *arXiv preprint arXiv:1910.10683*. Retrieved from <https://arxiv.org/abs/1910.10683>

Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *arXiv preprint arXiv:1908.10084*. Retrieved from <https://arxiv.org/abs/1908.10084>

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Agarwal, S. (2021). Making Pre-trained Language Models Better Few-shot Learners. *arXiv preprint arXiv:2101.06840*. Retrieved from <https://arxiv.org/abs/2101.06840>

Contact Information

Irena Vodenska, PhD, Professor
 Email: vodenska@bu.edu
Hyunuk Kim, PhD, Assistant Professor
 Email: uk@bu.edu

Department of Administrative Sciences
 Website: <https://www.bu.edu/adminsc/>