# 2022 REO Blueprint Competition:
# Utilizing Multimodal Sentiment Analysis to Eliminate Climate Change Deniers and Distribute Educational Materials Over Social Media

Isaac Bahler
Russell Frost
Dalhousie University - Bachelor of Science
Lakehead University - Department of Software Engineering

***Abstract*—While extensive research using sentiment analysis has been conducted using the vast amounts of data generated by social media, little research has been conducted into how multi-modal analysis can be used to address climate change misinformation and learn demographic details regarding who is effected. The following research proposal aims to amalgamate existing studies that detect climate change deniers along with effective online climate research delivery to combat the misinformation that exists in these select communities. By analysing what the response is to different initiatives the aim is more effectively support and educate those who have fallen prey to misinformation online and to limit its spread. Multi modal sentiment analysis is the ideal tool for conducting such a study as by using more than one media we gain a more effective understanding of how people perceive the problem. The benefit of such a system would be enormous it would allow for more effective education, action and iteration to ensure that the online ecosystem is one that breeds positive activity rather than undermines the global push to save our planet.**

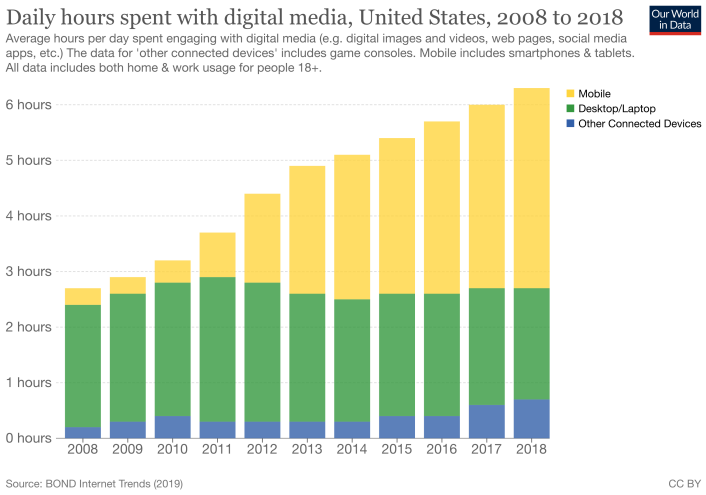


Fig. 1. Open Source Our World in Data Diagram demonstrating the increase in hours spent on digital media

## I. Introduction

Since Dr. Moore's observation in his 1965 paper of the rapidly increasing number of components that could be put on a chip(1), humans have been entering the digital age and the medium through which ideas had traditionally spread is beginning to change. The rise of social media platforms and online communication was exponential, from 2008 to 2018 the average time spent online daily for adults doubled and the primary modality for information acquisition became social media and online sources over traditional print media (2). Information, opinions, ideas; all travel exponentially faster through social media whether they be verifiable or not (3). Given the extraordinary reach and influence that these platforms have worldwide it is unsurprising that they are often a vector to motivate action on a number of issues. Unfortunately with the speed and ease of social media misinformation and bad actors seeking to weaponize it have become equally commonly for social media to become a source of deniers and the misinformed (4). On no contemporary is this perhaps a greater problem than that of climate change.

The literature is undeniable in this area in this area, global climate change is evident around the world and the scientific community has widespread ideas about how to slow it down and reverse it. This idea's; things like carbon capture or emission credit cannot be put in place without the widespread support and demands of citizens. The prevalence of climate change deniers or those misinformed on the subject is one source of opposition to the ideas mentioned above (4). Creating a system that can rapidly detect misinformation, determine the intent of the poster and iterate different means of educating them online then becomes a problem of massive importance.

## II. Topic Background

Sentiment analysis has been effectively used to find climate change denial sentiment on twitter(5); these tools struggle with sarcasm or double negation, however, these issues may in part be due to small data size. Methods of data cleaning and collection such as looking at specific hashtags on twitter could fix issues of sarcasm or double negation as well, provided a large enough sample size in that hashtag(5). On a deeper level multimodal (more than one type of input) sentiment

analysis has been used for the continuous prediction of a persons emotions from a feed of audio, visual, and biological signals(6); multimodal sentiment analysis of text, image and video data can also be used for social media posts(7). The multi-faceted nature of such an analysis tool makes it ideal for analyzing users motivation and for evaluating the effectiveness of educational materials then presented to them.

## III. Hypothesis/Design

### A. Hypothesis

We believe that a supervised multimodal sentiment analysis can be used analyze the intent of those posting about the climate information and to provide deniers or the misinformed with educational information.

### B. Design

The hope is to create a neural network that can identify sentiment of climate informed, denying, and confused individuals; this will allow for us to discover where climate change denial and confusion sentiments originate, and gauge the effectiveness of educational posts in those areas based on the sentiment differences between a posts response and the community as a whole.

## IV. Materials/Methods

### A. Materials

Materials include a computer, a database of social media posts, and a machine learning library such as TensorFlow in Python. Samples of climate change information or action that need to be conveyed to the public in later portions of the proposal may also be beneficial.

### B. Methods

*1) Preparation of Materials and Methods:* The first step of preparation is cleaning the data, this may involve; removing articles (i.e. the, a, an), which can help increase learning speed without reducing accuracy; removing the likelihood of ironic or sarcastic posts, which can be done in the data collection stage (i.e. by looking at specific hashtags on twitter; and the correction of spelling mistakes.

After the data is cleaned, adding noise to the data can help increase the data size. Noise can added to text by swapping words, deleting adjectives, and removing punctuation.

The final step of data preparation is categorization and separating out the test set. Initially all the data being used to train the AI must be categorized into the desired sentiment categories, this is so that we can get an accuracy score from the AI as it is being trained; for our purposes these categories would be 'climate change confused', 'climate change denial', and 'climate change informed'.

*2) Neural Network Parameters:* After the data is prepared focus can shift to the training of the neural network; for these purposes utilizing a layered network with Naive Bayes classification for text and and a support vector machine for classification of images seems to be most effective(8). The training can use k-fold cross validation for the creation of test sets; in this case the data is split into 'k' sections and each section is tested against all the others. Finally the complexity of the network itself can be tweaked, changing the number of input and output nodes to find a balance that isn't over-trained or under-trained.

If we are correct then our sentiment analysis tool should be able to detect the global warming sentiment of posters, allowing for the targeting of those who are confused about global warming to be targeted by educational content without also attracting communities of global warming deniers. From here we can categorize datasets as having global warming informed, confused, or denial leanings.

*3) Experiment Methodology:* To discover if our neural network can be effectively used to target those who are confused about the climate it would be helpful to look at interactions between a post with global warming confused hashtags, a global warming informed hashtag, and a global warming denial hashtag. Using a post with no hashtag as a control, we could then test if the interactions in each hashtag have the desired effect; for example an educational post in a global warming confused environment would hope to see more global warming informed responses as the confused are educated, or a post in a global warming denial environment would hope to see more global warming confused (in the process of becoming informed) and informed replies as deniers are educated.

*4) Complementary Work:* In the future using a modified system research could be done comparing the effectiveness of climate education on differing social media platforms such as TikTok contrasted with Instagram. Tangentially this system could also be used to evaluate the effectiveness of advertisement campaigns using the same iterative approach that presents and evaluates informative posts. The system could even be used in the identification of different online knowledge groups or virtual echo chambers with the aim of preventing dangerous outcomes or information. With slight alterations the different uses for the analysis tool proposed here are near limitless, motivating action an creating support for different public initiatives would be markedly easier than at present.

## V. Discussion

When considering the feasibility and impact of any study it is essential to consider whether the analysis methods match the objectives. Given the desire of this proposal to directly identify how people feel about a given post or piece of information, sentiment analysis then becomes the only viable option. Conducting multimodal sentiment analysis extends our proposed study in manner which, to our knowledge no other groups have conducted research in the climate action field; making our report unique and innovative. The research method

was also chosen as unlike other forms of data collection such as surveys collecting the needed video, images or text requires no additional effort, payment or specialized methods as it is already readily available. This means that it could be easily implemented in a research setting with the appropriate support while still providing a direct contribution to the field. The research method is limited in that the neural network can mistakenly identify climate change deniers or draw incorrect conclusion regarding education effectiveness. This would be an example of the network using false parameters and would require correction in future iterations.

We feel our research is important as sentiment analysis is an exploding field, showing utility in areas such as determining customer satisfaction on products(9), analysis of sentiment on MeToo tweets(10), and due to the misinformation issue that is now prevalent(4), we believe that it is important to fight against climate change denial in particular. We hope to fight climate change denial through education delivered through a lens of climate change sentiment analysis in internet groups (such as forums, hashtags or communities).

## VI. Conclusion

Given the prevalent nature of the issue at hand and the pressing need for solutions to the climate crisis in all fields, we believe that multimodal sentiment analysis could be an excellent stopgap measure in increasing public knowledge about climate change and motivating support for climate change initiatives. The existing research in the field has provided the groundwork of most major components in our design, integrating them and providing a tangible tool to motivate change is the main initiate. Given the requirements of the challenge to provide a proposal that could be developed easily in a lab setting this tool is an excellent candidate. Taken together this proposal can increase scientific literacy and fight the spread of misinformation motivating users online to support legalisation and actions which directly combat climate change. The urgency of the problem is extensive but by improving the online space we all increasingly inhabit we can inspire vast changes around the world.

## References


[1] Mollick E. Establishing Moore's Law. IEEE Annals of the History of Computing. 2006;28(3):62–75.

[2] Ortiz-Ospina E. The rise of social media. Our World in Data. 2019 Sep. Available from: https://ourworldindata.org/rise-of-social-media#:~:text=The20percentage20of20US20adults.

[3] Kaligotla C, Yücesan E, Chick SE. Diffusion of competing rumours on social media. Journal of Simulation. 2020 Aug;16(3):1–21.

[4] Suarez-Lledo V, Alvarez-Galvez J, et al. Prevalence of health misinformation on social media: systematic review. Journal of medical Internet research. 2021;23(1):e17187.

[5] Garg S, Panwar DS, Gupta A, Katarya R. A Literature Review On Sentiment Analysis Techniques Involving Social Media Platforms. In: 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC); 2020. p. 254–259.

[6] Stappen L, Baird A, Christ L, Schumann L, Sertolli B, Me
ssner EM, et al. The MuSe 2021 Multimodal Sentiment Analysis Challenge: Sentiment, Emotion, Physiological-Emotion, and Stress. In: Proceedings of the 2nd on Multimodal Sentiment Analysis Challenge. MuSe '21. New York, NY, USA: Association for Computing Machinery; 2021. p. 5–14. Available from: https://doi.org/10.1145/3475957.3484450.

[7] Basu P, Tiwari S, Mohanty J, Karmakar S. Multimodal Sentiment Analysis of MeToo Tweets using Focal Loss (Grand Challenge). In: 2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM); 2020. p. 461–465.

[8] Thuseethan S, Janarthan S, Rajasegarar S, Kumari P, Yearwood J. Multimodal Deep Learning Framework for Sentiment Analysis from Text-Image Web Data. In: 2020 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT); 2020. p. 267–274.

[9] Rahul, Raj V, Monika. Sentiment Analysis on Product Reviews. In: 2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS); 2019. p. 5–9.

[10] Basu P, Tiwari S, Mohanty J, Karmakar S. Multimodal Sentiment Analysis of MeToo Tweets using Focal Loss (Grand Challenge). In: 2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM); 2020. p. 461–465.