Reddit's Reaction to Fire: A Deep Learning Sentiment Analysis of Prescribed Burns and Wildfires

Joey Cartwright and Dr. Lucy Owen
University of Montana

Overview

Prescribed Burns are essential for wildfire management but often misunderstood. Public perception is important for fire management and policy making decisions.

Reddit is a social media platform featuring community conversations about prescribed burns and wildfires.

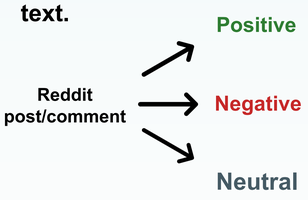
Sentiment Analysis is an algorithm that calculates the mood/sentiment of text (Example: positive, negative, or neutral).

Goal: Apply sentiment analysis + geospatial data to observe public sentiment trends of prescribed burns and wildfires with respect to time and region.

Methods

1. **Gather text data through Reddit's API.**

2. **Use a sentiment analysis model to calculate tone of text.**



3. **Verify model accuracy**

Prolific

4. **Visualize sentiment trends over time, focusing on periods of known prescribed burns and wildfires.**

First Model Selection: VADER

VADER is a rule-based sentiment analysis model that stores words from –4 (negative) to +4 (positive).

{'awesome' : +3.8} {'terrible' : -3.6}

VADER analyzes sentences and assigns each text input an overall sentiment score from –1 (negative) to +1 (positive).

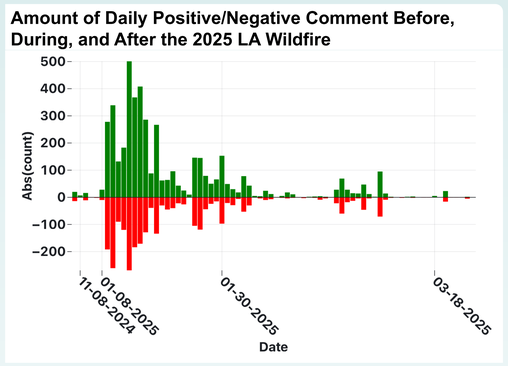


Fig 1. Visualizing sentiment trends from the Los Angeles subreddit. This data was taken from comments posted in January of 2025. During that time, Los Angeles got impacted by a historical wildfire. Green bars above the x axis represent positive comments, and red bars below the x axis represent negative comments. These counts were evaluated using the VADER model.

VADER Strengths	VADER Limitations
• No Training Required. • Evaluates Quickly. • Can understand some social media slang.	• Cannot understand complicated sentiment such as sarcasm. • Only evaluates text for three labels: positive, negative, and neutral. • Struggles to understand conversations where context is needed.

Final Model Selection: BART Zero Shot Classifier

Bart is a neural network model trained on understanding written language.

Zero Shot Classification allows BART's understanding of language to predict whether a label *entails* or *contradicts* a sentence.

Although BART has the skills to classify text for any given label, it still needs to be fine tuned to better predict our label set.

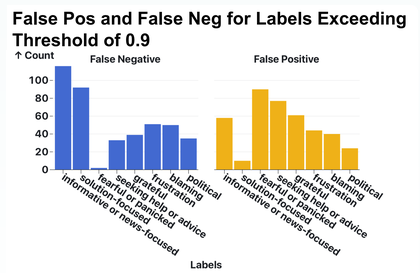


Fig 2. BART's prediction errors on Prolific test data using BART's zero shot classifier.

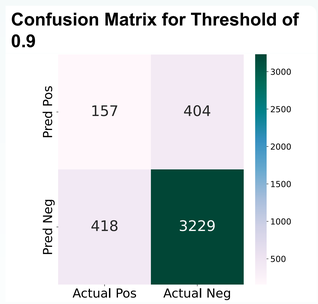


Fig 3. BART's confusion matrix for assigning labels with a confidence score of .90 on Prolific test data. This gave an overall prediction accuracy of .805.

Final Label Set

Grateful	**Political**	**Informative/News**	**Blaming**	**Prescribed Burn**	**Fire Management**
Frustration	**Solution Focused**	**Fearful or Panicked**	**Seeking Help or Advice**	**Wildfire**	

Current Stage: Fine Tuning Using Synthetic Data

1. Prompt AI for text.

2. Organize training data.

Text	Hypothesis	Label
"This sucks"	This text comes off as frustrated	entailment
"This sucks"	This text comes off as grateful	contradiction

3. Distribute training data for more efficient generalization.

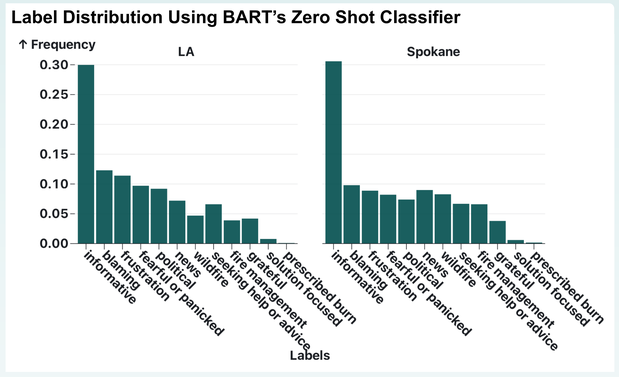


Fig 4. Text from the LA subreddit and Spokane subreddit were evaluated using BERT's zero shot classifier. Labels were given based off BERT's confidence. Any label exceeding a confidence threshold of .95 was assigned to its respective text. This data does not completely, nor accurately represent the true distribution. However, it does help estimate the general distribution.

Fig 5. Both figures above represent the most frequent words found from querying text from the LA and Spokane subreddits. All comments analyzed did not come from any specific time and were randomly selected. Since the words come from random comment queries, it gives me the best general estimate for what topics are most often discussed on Reddit.

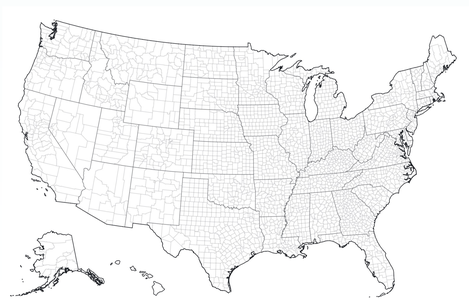
4. Gather test data.

Prolific

By collecting labeled responses through Prolific, I can create a test dataset that enables my model to validate its accuracy.

Direction: What this Data and Model can Accomplish

Plot several sentiment trends across the US on a time scale.



Use the fine-tuned model for disaster relief.

Use the fine-tuned model for geospatial policy feedback.

Acknowledgment
I would like to thank Dr. Lucy Owen for her guidance and feedback throughout this research. I also thank Montana NSF EPSCoR for supporting this work through funding, and my lab at the University of Montana for their support and feedback.

References
- Endsley, K. A., & McCarty, J. L. (2013). Mapping prescribed burns and wildfires from Twitter with natural language processing and information retrieval techniques. Proceedings of the International Smoke Symposium 2013. Retrieved from: https://digitalcommons.mtu.edu/mtri_p/165