Measuring Public Sentiment Towards Nuclear Energy Using Twitter Data

Abstract

The public debate over nuclear energy has been smoldering in the United States for many decades. In recent years, increasing concerns over climate change and the reliability of renewable energy sources have elevated the profile of nuclear energy as a zero-emissions energy source. This paper seeks to analyze the public discourse around nuclear energy on the social media platform, Twitter. With our analysis, we have obtained not only similar results to those published in survey analyses but also key insights that are not represented in the published survey data. We find that 67 percent of users in our dataset primarily expressed positive sentiments towards nuclear energy, and more interestingly, that men and older people express more positive sentiments than women and younger people, respectively. Additionally, the key concerns regarding nuclear energy are nuclear waste, accidents, and nuclear weapon proliferation. Accounts belonging to organizations are the predominant advocates of nuclear energy as a potential aid in addressing climate change, while accounts belonging to individuals, particularly women and young people, tweeted more about the threats posed by nuclear weapons and accidents. Ultimately, the broadly positive sentiments expressed in our dataset bode well for the future of nuclear energy in the United States.

Introduction

Public opinion of nuclear energy has gone through many cycles of interest and disinterest through the decades (Bisconti 2020). While one of the most scalable zero-emissions energy sources, it is not without its downsides. Concerns around nuclear weapons proliferation, environmental contamination, waste disposal and safety have historically held back the industry (Baron and Herzog 2020, Guptaa et al. 2019). Because of the highly politicized nature of the nuclear energy industry, public support or opposition to nuclear power is often the deciding factor as to whether existing reactors stay online and new nuclear power plants get built.

The availability of massive amounts of social media data provide a new opportunity to rapidly measure public opinion towards politicized topics of this nature. Public opinion survey data regarding nuclear energy have been regularly conducted for decades, but these tend to be relatively costly and suffer from small sample sizes. The social media platform, Twitter, provides a unique opportunity for opinion mining because of the highly political and emotional language used on the platform. Tweets have a short character limit, which prevents detailed, technical discussion and incentivizes highly sentimental language (Pfitzner et al. 2012). Additionally, our analysis focused exclusively on Tweets from users within the United States, and it is common for Twitter users to provide their self-reported location, which we used to filter out Tweets from users outside of the United States.

Related Work

Numerous studies have used social media for measuring public sentiment towards energy related topics. For examples, Park (2019) measured South Korea's public response (both positive and negative) to major media events related to nuclear energy by analyzing seven years' worth of social media posts from four major platforms. In particular, this paper was able to successfully identify South Korea's public response to the Fukushima accident in 2011, as well as positive news events such as research partnerships and major trade deals. Geotagged Twitter data has been used to measure the Alaskan public's opinion of their current and preferred energy options, such as tidal, solar, and nuclear energy (Abdar 2020). Likewise, sentiment and topic analysis have been conducted on Twitter data from the U.K. and Spain to identify the public's energy preferences and attitudes towards climate change in general (Loureiro and Allo 2020).

The inference of the demographics of Twitter users is possible with new deep learning algorithms (Wang et al. 2019). These demographic predictions have been combined with state-of-the-art sentiment and topic analysis models to reveal interesting patterns in large Twitter datasets (e.g. Duong et al. 2020).

Finally, there exists a large literature analyzing survey data regarding public opinions on nuclear energy. Most notably, Bisconti (2020) summarizes 36 years of survey data in the United States to show how support for nuclear energy has changed over the decades. This work is particularly important for understanding attitudes of different demographics towards nuclear energy, as well as the key issues affecting public opinion.

Methodology

We set out to find the key issues that contribute to people's opinion of nuclear energy. The age and gender of Twitter users who tweeted about topics related to nuclear energy are inferred from the users' profiles and used to form five major demographic clusters. Topic analysis is then applied to these demographic clusters to reveal the main priorities of different groups. This is followed by three-class sentiment analysis of the users' Tweets regarding nuclear energy in an effort to measure the overall approval of nuclear energy as well as the approval rate for various demographic classifications.

Data Collection. Tweets related to nuclear energy from users within the United States were collected from the Twitter API [1] using a series of keyword searches with phrases such as "nuclear energy", "nuclear power", "nuclear plant", etc. Each tweet returned from the keyword searches includes a user's screen name, self-reported "real" name, self-reported location, and profile picture url which are each collected and used for demographic predictions. The API's keyword search function returns Tweets from both individual accounts and organizations, so the organizational status of each account is also necessary to predict for accurate demographic analysis. Tweets were collected over a period of several months in the end of 2020, resulting in a dataset comprising over 50,000 nuclear energy-related Tweets from about 25,000 unique users within the United States.

Text Preprocessing. To analyze the tweets' text, we develop a text preprocessing pipeline similar to that of Baziotis et al. (2017). This pipeline includes performing sentiment-aware tokenization, spell correction, word normalization, word segmentation (for splitting hashtags) and word annotation. The challenging part of tokenization is to correctly split the word without separating expressions or words that should be kept intact. This is particularly important in texts from social networks, with "creative" writing and expressions like emoticons, hashtags and so on. Thus, we implement a social tokenizer geared specifically towards Twitter. This tokenizer is capable of recognizing Twitter markup, emoticons, emojis, expressions such as dates, times, currencies, acronyms, censored words (e.g. f**k), words with emphasis (e.g. a *great* time) and more. After tokenization, we use the Viterbi algorithm to perform spell correction and word segmentation with word statistics (unigrams and bigrams) computed from our unlabeled dataset to obtain word probabilities. Finally, the text is converted to lowercase, and all URLs, emails and mentioned usernames are removed to retain the natural language elements from the text data.

Latent Dirichlet Allocation. We employ Latent Dirichlet Allocation (LDA), a generative, unsupervised, probabilistic model that is commonly used to reveal latent topics in text corpora (Blei et al. 2003), to identify the key concerns that people have regarding nuclear energy. Each Tweet in the text corpus is modeled as a bag of words with an occurrence count for each token. In our analysis, we not only look at individual tokens but also highly correlated groups of words. Therefore, bigrams and trigrams are also added to our corpus. These documents and occurrence counts are then used to generate a probability distribution over a pre-specified number of topics. Each topic, therefore, is the probability distribution over each token in the corpus. The highest probability tokens within each topic can then be used to give the topic a human-interpretable name.

M3-Inference. To generate demographic predictions for users in our dataset, we use the deep learning systems, M3-Inference (Multimodal, Multilingual, and Multi-attribute) to predict the age category, gender, and organizational status of each user account. These three attributes, while not comprehensive of a user's demographics, are extremely informative, particularly with respect to social and political opinions.

The M3-Inference model is a state-of-the-art, multilingual, multi-attribute system trained on a massive Twitter dataset to be able to take the user's biography, screen name, user name, and profile image to make demographic classifications (Wang et al. 2019). The four age categories are "<= 18", "19-29", "30-39", and ">= 40"; gender classifications are "male" and "female"; and organizational statuses are "isorg" and "notorg", the latter of which designates that the account is for an individual rather than an organization. For demographic classifications in English, this model has a macro-F1 score for age, gender, and organizational status of 0.42, 0.92, and 0.90, respectively. The relatively low performance for age classification indicates that conclusions made regarding age classification should be taken with a grain of salt. Please note that this paper makes no attempt at evaluating the accuracy of the M3-Inference model on our dataset.

Transformers for Sentiment Analysis. To identify the sentiment for each Tweet, we use the transformer models that have shown state-of-the-art performance on natural language processing tasks. This paper compares five models that hold the best performance on the SST-2 dataset for text classification including BERT, RoBERTa, XLNet, ELECTRA and BART. We use pretrained models from the huggingface transformers library [2] and fine tune them by adding an additional output layer to the pretrained models. A more detailed introduction for each model will be presented in the sentiment analysis section below.

We use BERT as a baseline model and compare its performance with four other models across three annotated datasets designed for tweets sentiment classification tasks. The best performing model is then used in analysis of sentiment of nuclear energy-related Tweets. To validate the performance of our chosen model on our nuclear energy dataset, a random sampling of 100 Tweets is hand labeled and used to evaluate the accuracy of the chosen model.

FP-Growth. Connections between users' demographics and their sentiment towards nuclear energy are mined using frequent patterns mining. We divide our users into two classes based on the sentiment classification of their Tweets and then mine the user demographic inferences within each class of users. We use FP-Growth, proposed by Han et al. (2000) to extract these frequent patterns. This algorithm has

[1] 1https://git.io/JvAjh

[2] https://git.io/JfUEh

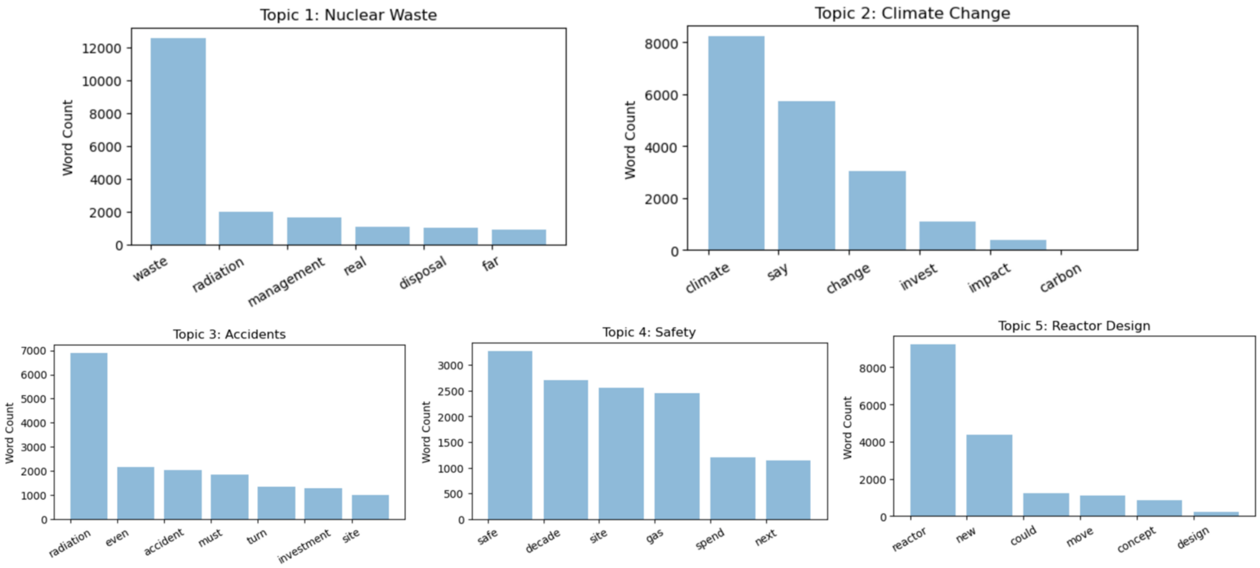

Figure 1: Word Frequency of Top Ten Topic Keywords.

been shown to be efficient and scalable for mining both long and short frequent patterns, and is about an order of magnitude faster than the Apriori algorithm and also faster than some recently reported new frequent pattern mining methods. Our results are obtained using a support confidence of 15 percent.

Experiments

Key Topic Analysis

The number of topics used in LDA analysis is determined by optimizing for the coherence score of the LDA model on our nuclear energy-related Tweet dataset. The coherence scores for various numbers of topics is plotted in Figure 2. This figure shows that the coherence score for this dataset reaches a maximum at 70 topics, which resulted in the relatively high coherence score of 0.57. We then labeled the top five topics resulting from this model. Figure 1 presents each of these topics with their top ten key words and corresponding word counts. This figure shows that the most common subtopic in discussions of nuclear energy on Twitter is nuclear waste. This is followed by discussions of Climate Change, Accidents, Safety, and Reactor Design.

Surveys of public opinion towards nuclear energy show that negative opinions tend to be much more focused than positive opinions, so it was expected that the most common topic would be a negative one like the discussion of nuclear waste. This analysis also matches survey results that show accidents, waste, and safety to be the general public's primary areas of concern around nuclear energy (Bisconti 2020).

Interestingly, there is quite a bit of technical discussion in our dataset as evidenced by Topic 5. This reflects the nature of the public discourse around nuclear energy as a highly technical and nuanced topic of decision. This also illustrates the value of Twitter as a social media platform and the ease with which users can share news articles about the latest research and energy policy reports. In addition to these top five topics, other popular topics include renewable energy (Topic 6), building the future (Topic 7), and national security (Topic 12).

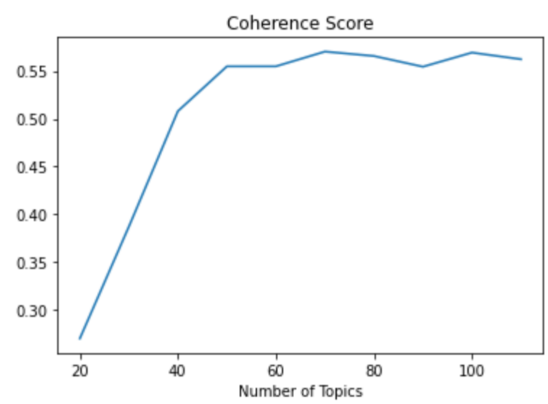

Figure 2: Coherence Scores of the LDA model on nuclear-energy related Tweets.

Topic Analysis of User Demographic Clusters

Demographic predictions of users in our dataset resulted in a gender distribution that is 69.1 percent male and 30.9 percent female which indicates that males tweet about nuclear energy much more than female Twitter users. While Twitter as a whole has been documented to have somewhat similar gender disparities, it is difficult to quantify just how much more skewed the topic of nuclear energy is towards male Twitter users than Twitter as a whole when taking into account the accuracy of these predictions (Barbera

Table 1: Topic Analysis by Demographic Clusters.

	Whole Dataset	Men ≥ 40	Women ≥ 40	Men and Women 30-39	Men 19-29	Organizations
Topic 1	Waste	Weapons	Trump	Accidents	Weapons	Reactor Design
Topic 2	Climate Change	Radiation	Weapons	Potential Problems	Utility Details	Making the Future
Topic 3	Accidents	Reactor Design	Energy Needs	Reactor Design	Energy Needs	Climate Change
Topic 4	Safety	Waste	Security	Weapons	Renewables	Electricity Generation
Topic 5	Reactor Design	Safety	Waste	Renewables	Reactor Design	Waste

& Rivero, 2015) (Mislove, et al., 2011). The inferred age distribution was 9.5 percent "$\leq$ 18", 19.7 percent "19-29", 19.7 percent "30-39", and 51.1 percent "$\geq$ 40". Notice that this age distribution is heavily skewed towards older Twitter users, particularly compared to Twitter users' age distribution as a whole (Mellon, 2017). Finally, the organizational status prediction indicated that only 11.7 percent of our users were organization accounts, however, Tweets from organizations accounted for roughly 20 percent of all Tweets in our dataset, suggesting that organizations have disproportionate influence over the public discourse around nuclear energy on Twitter. This is not particularly surprising because one would expect an organization dedicated to nuclear energy or environmental issues to be more vocal than the average Twitter user regarding nuclear energy.

These demographic predictions were combined with other user attributes, such as number of followers, to form demographic clusters using K-prototype. The optimal number of clusters is determined using the elbow method based on within-cluster variance, which results in a total of five clusters for our dataset. These five clusters each comprise between roughly 3,000 and 7,500 users and are given the following names, in no particular order, based on their majority demographic attributes: "Men over 40", "Women over 40", "Men and Women 30-39", "Men 19-29", and "Organizations".

We then divide our dataset of nuclear energy-related Tweets according to these user clusters and conduct topic analysis to obtain the top five most popular subtopics for each demographic cluster. This result can be seen in Table 1. There are several interesting patterns emerging from this analysis. First, the topic of "weapons" does not enter into the top five topics of the dataset as a whole, but it is, however, a major topic for four of the five demographics. This relative obscurity of "weapons" discussions in the dataset as a whole is due to the fact that Organizations tend not to discuss issues around nuclear weapons and, instead, tweet about nuclear energy in relation to climate change. Notice that while the topic of "climate change" is the second most common topic overall, it does not appear in the top five topics for any of the demographic clusters except Organizations. The implication here, that the general public does not broadly think of nuclear energy in relation to climate change, is also backed by survey data (Bisconti 2020). Our results particularly demonstrate the disproportionate influence that organizations have in shaping the public discourse on social media.

Additionally, the topic of "nuclear waste" appears to be a major topic among three of the five demographic clusters. Interestingly, nuclear waste does not appear as a major topic for younger users under 39 years old. Instead, these users tend to tweet about nuclear energy in relation to accidents and weapons. Because Fukashima (2011) is the only major nuclear accident that has occurred during the lifetime of these users, this result seems to indicate the weight with which that single accident continues to affect public opinion today. Finally, the focus of Women over 40 on the topic of "Trump" illustrates how topical the public discourse on social media can be. This demonstrates one of the main benefits of this approach to measuring public sentiment which is its rapidity and ability to measure public opinion in relation to current events.

Sentiment Analysis

To expand the scope of our analysis, we train and evaluate a range of sentiment analysis models to select the best performing model to conduct a three-class classification of the sentiment (positive, neutral, and negative) of each Tweet in our dataset. We then compute the sentiment score for each user who tweeted about nuclear energy based on the prediction results of our classifier. In our dataset, each sentiment is presented by a number (Positive: 1, Neutral: 0, Negative: -1). The sentiment score for each user is calculated by averaging the sentiment of all the Tweets belonging to each user. A positive classification is given to users with a positive sentiment score and vice versa.

Transformer models used for sentiment classification. In this section, we compare five pretrained transformer models that have achieved state-of-art performance in natural language processing tasks across multiple datasets. They have achieved similar performance on SST-2, a Stanford Sentiment Treebank (SST) dataset with binary labels (positive and negative). To find the model that performs best for Twitter text, we train and evaluate each of our candidate models across three annotated datasets. The model that achieves the best average performance is then used for sub-

Table 2: Performance Comparison Between Classifiers.

	Twitter Airline		Apple Sentiment		SentEval2017		Average Score	
	Accuracy	F1	Accuracy	F1	Accuracy	F1	Accuracy	F1
BERT	0.805	0.814	0.752	0.758	0.745	0.745	0.767	0.772
ROBERTA	0.816	0.827	0.756	0.769	0.749	0.752	0.774	0.783
XLNET	0.836	0.838	0.758	0.774	0.737	0.741	0.777	0.784
BART	0.831	0.835	**0.799**	**0.801**	0.725	0.725	0.786	0.787
ELECTRA	**0.852**	**0.856**	0.781	0.782	**0.791**	**0.793**	**0.808**	**0.810**

sequent analysis and evaluated on a small, hand labeled sample of our nuclear energy dataset.

BERT (Devlin et al. 2019) was designed to pretrain deep bidirectional representations of tokens from unlabeled text by jointly conditioning on both left and right context in all layers. As a result, the pre-trained BERT model can be fine-tuned with just one additional output layer to create state-of-the-art models for a wide range of tasks, such as question answering and language inference, without substantial task-specific modifications to its architecture. The large version of Bert achieved a 0.932 F1 score on SST-2.

BART (Lewis et al. 2019) was trained by corrupting text with an arbitrary noising function, and learning a model to reconstruct the original text. BART uses a standard Transformer-based neural machine translation architecture which, despite its simplicity, can be seen as generalizing BERT (due to the bidirectional encoder), GPT (with the left-to-right decoder). The major advantage of this set up is the flexibility of the arbitrary transformations that can be applied to the original text, including changing its length. As a result, they reported a 0.966 F1 score of their large version of BART on SST-2.

RoBerta (Liu et al. 2019) is an improved recipe for training BERT models. The large RoBERTa model achieved a 0.964 F1 score on SST-2.

XLNet is a permuted language model. In their study, Yang et al. (2019) pointed out that BERT neglects dependency between the masked positions and suffers from a pretrain-finetune discrepancy. Therefore, they proposed XLNet, a generalized autoregressive pretraining method that enables learning bidirectional contexts by maximizing the expected likelihood over all permutations of the factorization order and overcomes the limitations of BERT thanks to its autoregressive formulation. In addition, Li et al. (2020) observed the target-position-aware representation and relative position encoding features of XLNet, leading to a better benchmark score at the cost of 1.2× arithmetic operations and 1.5× execution time on a modern CPU. As a result, the large XLNet model achieved a 0.97 F1 score on SST-2.

ELECTRA (Clark et al. 2020), like XLNet, was also designed to solve the input noise limitation of BERT. Instead of masking the input, they proposed an alternative approach that corrupts it by replacing some tokens with plausible alternatives sampled from a small generator network. Then, instead of training a model that predicts the original identities of the corrupted tokens, they trained a discriminative model that predicts whether each token in the corrupted input was replaced by a generator sample or not. Thorough experiments demonstrate this new pre-training task is more efficient than MLM because the task is defined over all input tokens rather than just the small subset that was masked out. As a result, ELECTRA achieved a 0.971 F1 score on SST-2.

Training and Evaluation. To fine tune these models for sentiment classification, we add an additional output layer on top of the pretrained models' output. Since we use the large version of each of the models, we add a linear layer with dimensions of [1024, 256, 3], and the activation function, ReLU, is added to the last layer of each model. The AdamW optimizer (Loshchilov et al. 2017) is used to optimize the loss function. Because this is a multi-class classification task, we use the cross-entropy as the metrics to optimize for loss function. Moreover, we apply advanced fine tuning techniques (Discriminative fine tuning, Slanted triangular learning rate, Gradual unfreezing) proposed by Howard et al. (2018) for task-specific features (i.e., learning word embeddings, learning context embeddings, and learning sentiment outputs). As suggested by Howard's work, the combination of all these techniques allow our classifier to achieve state-of-art performance for our sentiment classification task.

We train our models on three different datasets (SemEval-2017 Task 4A, Twitter US Airline Sentiment and Apple Computers Twitter Sentiment), each of which is annotated with one of three sentiments: negative, neutral, and positive. After fine tuning our classifier, all of our models outperform the top performers of SemEval-2017 and Twitter US Airline Sentiment (Leaderboard of Apple Computer Twitter Sentiment is not available). These results are shown in Table 2. In particular, the pretrained ELECTRA model achieves an impressive weighted macro F1 of 0.81 score across the three datasets.

Next, we train the ELECTRA model on 20,000 randomly selected samples from the combination of all three datasets. The model achieves 0.79 accuracy on the combined dataset. The confusion matrix is shown in Figure 3. Lastly, we hand label 100 samples from our dataset to evaluate the fine-tuned ELECTRA model, which is able to correctly predict 74 percent of our samples. Since we are making predictions of datasets containing three classes, the lower accuracy compared to the results from (Clark et al. 2020) is reasonable. The confusion matrix for our ELECTRA model's performance on the hand labeled samples is given in Figure 4.

Sentiment results. We observe broad public approval expressed from Tweets related to nuclear energy. Out of 25,212 users, 67 percent of the users expressed positive attitudes to-

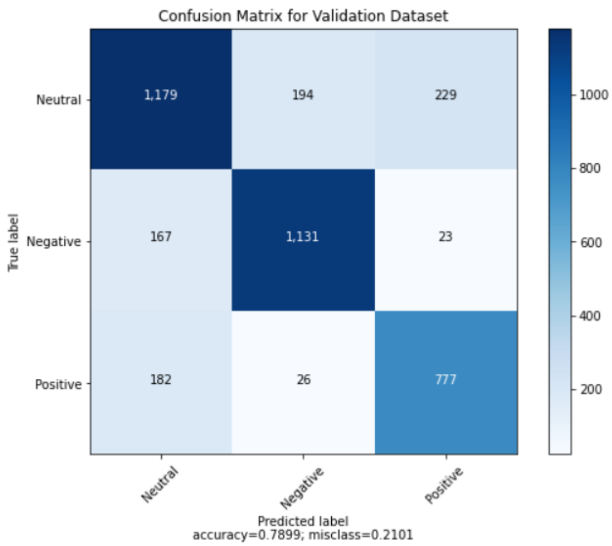

Figure 3: Confusion matrix for the Combined Dataset.

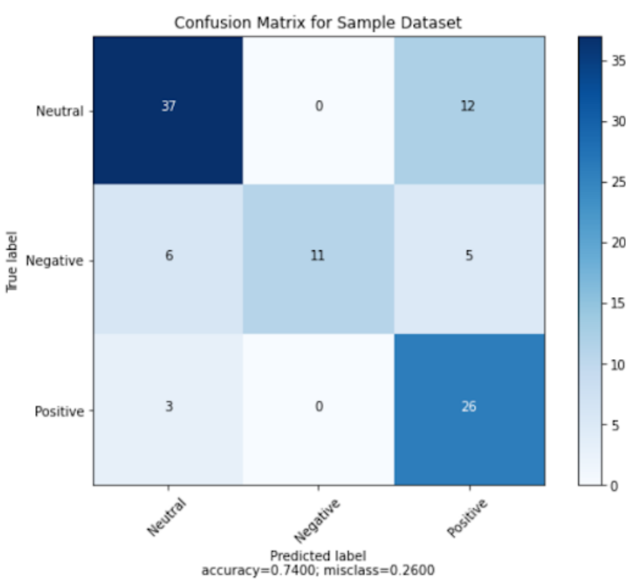

Figure 4: Confusion matrix for the Sampled Dataset.

wards nuclear energy, compared to just 25 percent of users who predominately expressed negativity towards nuclear power. This result reflects published survey data showing that those who favored nuclear energy outnumbered those who opposed by 61 percent to 26 percent (Bisconti 2020). Moreover, people who posted positive Tweets are more outspoken compared to negative users. On average, positive users Tweeted about nuclear energy 20 percent more than negative users, suggesting that positive opinions of nuclear energy are more strongly held than negative ones.

Another important insight from our study is that different user groups express very different sentiments towards nuclear power. The results of FP-Growth, shown in Figure 5, reveal that older male users, organizations and accounts with relatively many followers are the most likely demographics to express positive sentiment when tweeting about nuclear energy. In contrast, women and younger users are more likely to express negative attitudes towards nuclear energy.

Among positive users, {'male', '$\geq$40'} is the most frequent 2-itemset with a frequency of 5376 occurrences which means about 30 percent of positive users are male who are older than 40 years old. Based on the frequent patterns that we mined, the number of positive Tweets from {'male', '$\geq$40'} are three times more than the number of negative Tweets from this demographic. More interestingly, out of roughly 16,200 positive users, 2,350 are organizations, accounting for 15 percent of the positive users in our dataset. In contrast, out of 6,200 negative users, only 470 are from organization accounts (about 7 percent). This means organizations appear roughly twice as frequently among positive users than negative users, and pro-nuclear organizations are more numerous overall. Lastly, follower count appears to be heavily correlated with positive attitudes towards nuclear energy with follower counts in the top qaurtile often appearing among frequent patterns for positive users and follower counts in the bottom quartile appearing among the frequent patterns of negative users. This interesting pattern suggests that high-profile Twitter users tend to be more positive towards nuclear energy.

```
[[{'male', 'pos'}, 10171],
 [{'female', 'pos'}, 3820],
 [{'>=40', 'male', 'pos'}, 5376],
 [{'follower_count_top25%', 'pos'}, 4696],
 [{'follower_count_50%-75%', 'pos'}, 4070]],
[{'follower_count_bottom25%', 'neg'}, 1640],
[{'follower_count_25%-50%', 'neg'}, 1617],
[{'follower_count_50%-75%', 'neg'}, 1548],
[{'follower_count_top25%', 'neg'}, 1319],
[{'follower_count_bottom25%', 'pos'}, 3927],
[{'follower_count_25%-50%', 'pos'}, 4114],
[{'follower_count_50%-75%', 'pos'}, 3995],
[{' pos', 'follower_count_top25%'}, 4305],
[{' pos', 'isorg'}, 2350],
[{'isorg', 'neg'}, 471])
```

Figure 5: Partial Frequent patterns from FP-Growth.

To further illustrate this finding, we plot the ratio of positive to negative users for each demographic class in Figure 6. This reaffirms the gender and age incongruity in attitudes towards nuclear energy in our dataset and matches published analyses of survey data reporting that men are more likely to have a favorable opinion of nuclear energy than women and that favorability of nuclear energy increases with the age of the participant (Bisconti 2020).

Our analysis also reveals a large disparity between how outspoken pro-nuclear versus anti-nuclear organizations are on social media. This result manifests as a positive to negative Tweet ratio, shown in Figure 6, greater than 1.0 for organizations. As an anecdote to this finding, the largest anti-nuclear organization in the United States, Greenpeace USA, did not tweet about nuclear energy even once during the pe-

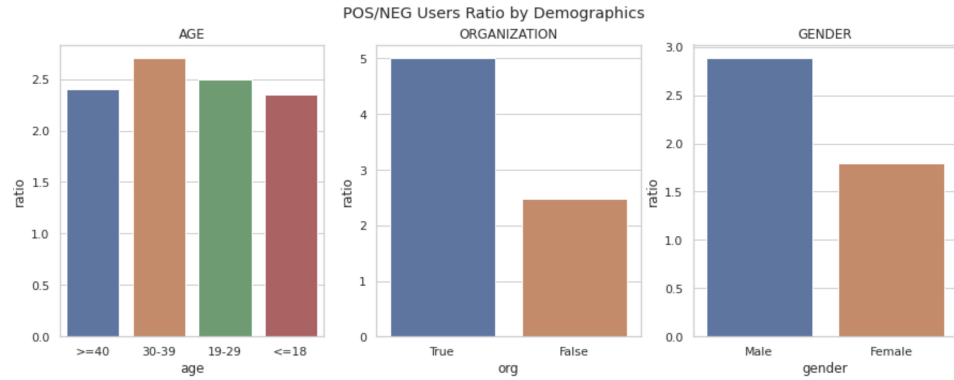

Figure 6: Ratio between pos/neg users by demographics.

riod in which data was collected for this project. Perhaps this is a sign that they and other anti-nuclear organizations are reconsidering their long-held stance that nuclear energy is not "green" in the face of the increasing threat posed by climate change.

Finally, our analysis reveals that California is a highly "pro-nuclear" state. Users reported to be living in California had a positive to negative sentiment ratio of 2.53, which is a significant majority. This finding sheds light on the direction of the current debate over whether to close the Diablo Canyon nuclear plant. In 2016, regulators voted to shut down the nuclear plan due to a desire to incorporate more renewable energy into the electric grid. However, this decision could be reversed if legislatures choose to classify nuclear energy as "green energy" (Shellenberger 2019). With California's growing population and a flourishing, high-tech economy, the demand for electricity has been steadily rising for years especially after the two energy shortage crises in 2001 and 2020. In addition, California has some of the most ambitious renewable energy goals in the country, aiming to get 60 percent of its electricity from renewables by 2030 and raising that to 100 percent by 2045. The motion to close down Diablo Canyon, the last operational nuclear plant in California which single-handedly generated 9 percent of California's electricity and 20 percent of its clean, carbon-free electricity, has generated a vibrant political discussion around the role of nuclear energy in addressing climate change (Nikolewski 2018). According to the results of our analysis, the political desire within California to reconsider Diablo Canyon's close-out is strong.

Conclusion and Future Work

The motivation of this study is to overcome the limitations of public opinion surveys due to their small sample size and high cost. We aim to identify the key issues affecting overall public sentiment towards nuclear energy using nothing more than publicly available Twitter data. In doing this, we are able to not only reproduce several key results from large survey analyses but also produce several new finding not available in surveys data. In particular, our results demonstrate the affect that organizations have on shaping public discourse around political issues. For example, it is largely the Twitter accounts of organizations that are advocating nuclear energy as a potential zero-emissions energy source to combat climate change while individual users are more concerned about nuclear waste, weapons, and accidents. In addition, our sentiment analysis showes a disparity between pro-nuclear and anti-nuclear organizations in which pro-nuclear organizations are much more vocal in their advocacy of nuclear energy as a potential aid in addressing climate change.

We then find that ELECTRA had the best performance for sentiment classification of Tweets and measured an overall approval rate similar to published results. Based on these sentiment predictions, we mine the frequent patterns of predicted demographic attributes within each sentiment class. The results of this analysis match published survey results indicating that older males have a more positive opinion towards nuclear energy, while female and younger users are more likely to be less approving of nuclear energy. While this paper mainly focuses on latent topics of subgroups based on demographic information, future research in this area could be extended to incorporate occupational inference (Hu et al. 2017) and the users' level of knowledge about the topic, which has been shown to correlate with favorable opinions of nuclear energy (Bisconti 2018).

One geographic trend that this paper is able to highlight is the high rate of positive attitudes that Californians show towards nuclear energy, which is particularly relevant given the current debate over the closure of the Diablo Canyon nuclear plant. Our results show that the California public is much more favorable towards nuclear energy than the regulators controlling the destiny of the power plant, suggesting a potential reversal of fortune for the plant scheduled to close by 2025. Future improvements of this work can be done by collecting geo-tagged Tweets as well as collecting Tweets over a longer time period to provide a more representative sample of sentiments from the population. In particular, a more representative dataset could show more trends across geographic areas and across time.

The success of our analysis in matching published survey results for a political topic that is as niche and technical as the debate over nuclear energy validates the strength of our methodology and social media's relevance in measuring the political opinions of the broader public. The methodology presented in this paper is general enough to be extended to any political topic of comparable importance and can provide vital insight into issues that are not well represented in the survey literature. This paper emphasizes the abiding role that social media plays in understanding the public discourse and political sentiment.

References

Abdar, M.; Basiri, M. E.; Yin, J.; Habibnezhad, M.; Chi, G.; Nemati, S.; and Asadi, S. 2020. Energy choices in Alaska: Mining people's perception and attitudes from geotagged tweets. In *Renewable and Sustainable Energy Reviews,* 124:109781

Barbera, P.; and Rivero, G. 2015. Understanding the Political Representativeness of Twitter Users.In *Social Science Computer Review.* 33:712–729.

Baron, J. and Herzog, S. 2020. Public opinion on nuclear energy and nuclear weapons: The attitudinal nexus in the United States. In *Energy Research Social Science.* 68:101567. issn: 2214-6296.

Baziotis, C.; Pelekis, N.; and Doulkeridis, C. 2017. DataStories at SemEval2017 task 4: Deep LSTM with attention for message-level and topicbased sentiment analysis. In *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017),* Vancouver, Canada: Association for Computational Linguistics, pp. 747–754.

Bisconti, A. S. 2018. Changing public attitudes toward nuclear energy. In Progress in *Nuclear Energy. Shippingport 60th Anniversary: A Time to Take Stock of Nuclear Energy's Status,* 102:103–113. issn: 0149-1970.

Bisconti, A. S. 2020. Factors Affecting Public Opinion of Nuclear Energy in the United States. In *Reference Module in Earth Systems and Environmental Sciences.* isbn: 978-0-12-409548-9.

Blei, D. M.; Ng, A. Y.; Jordan, M. I.; and Lafferty, J. 2003. Latent dirichlet allocation. In *Journal of Machine Learning Research.* 3:2003.

Clark, K.; Luong, M.-T.; Le, Q. V.; AND Manning, C. D.; 2020. ELECTRA: Pretraining text encoders as discriminators rather than generators. In *International Conference on Learning Representations.*

Devlin, J.; Chang, M. W.; Lee, K.; and Toutanova, K. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Minneapolis,* Minnesota: Association for Computational Linguistics, 1:4171–4186.

Duong, V.; Pham, P.; Yang, T.; Wang, Y.; and Luo, J. 2020. The Ivory Tower Lost: How College Students Respond Differently than the General Public to the COVID-19 Pandemic. arXiv preprint. arXiv:2004.09968 [cs.SI]. Ithaca, NY: Cornell University Library.

Guptaa, K.; Nowlinb, M.; Ripbergera, J.; Jenkins-Smitha, H.; and L.Silvaa, C. 2019. Tracking the nuclear 'mood' in the United States: Introducing a long term measure of public opinion about nuclear energy using aggregate survey data. In *Energy Policy.* 133:110888. issn: 0301-4215.

Han, J.; Pei, J.; and Yin, Y. 2000. Mining frequent patterns without candidate generation. In *SIGMOD'00,* Dallas, TX. pp. 1– 12.

Howard, J.; and Ruder, S. 2018. Universal language model fine-tuning for text classification. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers),* Melbourne, Australia: Association for Computational Linguistics, pp. 328–339.

Hu, T.; Xiao, H.; Nguyen, T. T.; and Luo, J. 2017. What the Language You Tweet Says About Your Occupation. arXiv preprint. arXiv:1701.06233. Ithaca, NY: Cornell University Library

Lewis, M.; Liu, Y.; Goyal, N.; Ghazvininejad, M.; Mohamed, A.; Levy, O.; Stoyanov, V.; and Zettlemoyer, L. 2019. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. arXiv preprint. arXiv:1910.13461. [cs.CL]. Ithaca, NY: Cornell University Library.

Li, H.; Choi, J.; Lee, S.; and Ahn, J. H. 2020. Comparing BERT and XLNet from the Perspective of Computational Characteristics. 2020. In *AInternational Conference on Electronics, Information, and Communication (ICEIC),* Barcelona, Spain, pp. 1-4, doi: 10.1109/ICEIC49074.2020.9051081.

Liu, Y.; Ott, M.; Goyal, N.; Du, J.; Joshi, M.; Chen, D.; Levy, O.; Lewis, M.; Zettlemoyer, L.; and Stoyanov, V. 2019. Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692. Ithaca, NY: Cornell University Library.

Loshchilov, I.; and Hutter, F. 2017. Decoupled weight decay regularization," arXiv preprint arXiv:1711.05101. Ithaca, NY: Cornell University Library

Loureiro, M. L.; and Allo, M. 2020. Sensing climate change and energy issues: Sentiment and emotion analysis with social media in the U.K. and Spain. In *Energy Policy*. 143:111490. issn: 0301-4215.

Nikolewski, R. 2018. Regulators vote to shut down Diablo Canyon, California's last nuclear power plant. 2018. *LA Times*.

Park, E. 2019. Positive or negative? Public perceptions of nuclear energy in South Korea: Evidence from Big Data. In *Nuclear Engineering and Technology.* 51:626–630.

Pfitzner, R.; Garas, A.; and Schweitzer, F. 2012. Emotional divergence influences information spreading. In *AAAI ICWSM 2012,* PP. 2-5.

Shellenberger, M. 2019. Why California May Go Nuclear. In *Forbes*.

Wang, Z.; Hale, Scott; A., David I.; Grabowicz, P.; Hartman, T.; Flöck, F; and Jurgens, D. 2019. Demographic Inference and Representative Population Estimates from Multilingual Social Media Data. In *The World Wide Web Conference.*

Yang, Z.; Dai, Z.; Yang, Y.; Carbonell, J.; Salakhutdinov, R.; and Le, Q. V. 2019. XLNet: Generalized Autoregressive Pretraining for Language Understanding. arXiv preprint arXiv:1906.08237. Ithaca, NY: Cornell University Library.