Measuring Public Sentiment Towards Nuclear Energy Using Twitter Data

Di Wu, Stephen Gass
University of Rochester
di.wu@simon.rochester.edu, sgass@ur.rochester.edu

Introduction

The public debate over nuclear energy has been smoldering in the United States for many decades. However, increasing concerns over climate change and concerns about energy security and reliability have elevated the profile of nuclear energy in recent years (Bisconti 2020). While one of the most scalable carbon-neutral energy sources, it is not without its downsides. Concerns around nuclear weapons proliferation, environmental contamination, waste disposal and safety are holding 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 nuclear energy. 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 on even relatively niche political debates such as the nuclear energy debate 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 a 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 (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 (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

This project set out to find the key issues that contribute people's opinion of nuclear energy. The age and gender of Twitter users who tweeted about topics related to nuclear energy were inferred from the users' profiles and used to form five major demographic clusters along with profile attributes, such as number of followers. Topic analysis was then applied to these demographic clusters to reveal the differences in the key issues around nuclear energy between different groups. This was followed by three-class sentiment analysis in an effort to measure the overall approval of nuclear energy for the various demographic classifications.

A. Data Collection

Tweets related to nuclear energy from users within the United States were collected from the Twitter API using a series of keyword searches with phrases such as "nuclear energy", "nuclear power", "nuclear plant", etc. Each Tweet returned from the keyword searches included a user's screen name, self-reported "real" name, self-reported location, and profile picture url which were each collected and used for demographic predictions. The API's keyword search function is not comprehensive and only returns Tweets that were created, at most, a week prior to the execution of the search. While the project was able to collect relevant Tweets over a period of about four weeks, this was not enough to meet our expectations for the project. Relevant Tweets were, therefore, also collected from two additional sources. First, an open source dataset containing 3,000 nuclear energy-related Tweets from mid-2020 were incorporated into the primary dataset. Second, a list of all registered pro-nuclear, anti-nuclear, and environmental organizations in the United States was scraped from Wikipedia, and the nuclear energy-related tweets from these accounts were incorporated into the primary dataset as well. The incorporation of Tweets from these three sources generated a dataset comprising over 19,000 nuclear energy-related Tweets from about 5,500 users within the United States.

B. Text Preprocessing

The various models used in analysis each had their own data preprocessing requirements. For the demographic inference of Twitter users, very little preprocessing was required. Emojis were removed from each user's biography, and ".gif" files were removed from the dataset of user profile pictures.

Additionally, simple preprocessing of our Tweet corpus was required for key topic analysis. This included the standardization to lower case as well as the removal of punctuation, new line characters, emojis, and stop words. Additionally, the keywords used for the Twitter API search were removed along with Twitter handles (beginning with '@'), web addressing (beginning with 'http'), rare words that only occur in the corpus a single time, and words that are only one or two letters long. Finally, lemmatization was used to standardize the language used in each Tweet, keeping only the nouns, adjectives, verbs, and adverbs.

A separate and less rigorous preprocessing was performed on the Tweet corpus prior to sentiment analysis. In general, the transformer models that were used in this project do not require text preprocessing, however, words that did not contribute to the sentiment of a Tweet were removed, namely, Twitter handles and web addresses. Emojis, capitalization, and punctuation were left unchanged for sentiment analysis.

C. Latent Dirichlet Allocation

To identify the key concerns that people have regarding nuclear energy, we leveraged Latent Dirichlet Allocation (LDA) to reveal the latent subtopics within our corpus of nuclear-energy related Tweets. LDA is a generative, unsupervised, probabilistic model commonly used to reveal latent topics in text corpora (Blei et al. 2003). 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. Thus, 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.

D. M3-Inference

To generate demographic predictions for users in our dataset, we used the deep learning systems, M3-Inference (Multimodal, Multilingual, and Multi-attribute) to predict the age category, gender, and organizational status of each user account that had tweeted about nuclear energy in our dataset. 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 that was 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. This relatively low performance for age classification indicates that conclusions made regarding age classification should be taken with a grain of salt. This project made no attempt at evaluating the accuracy of the M3-Inference model on our dataset.

E. Transformers for Sentiment Analysis

To identify the sentiment for each Tweet, we used the transformer models that have proven state-of-art performance on natural language processing tasks. For this project, we compared five models that hold the best performance on the SST-2 dataset for text classification including BERT, RoBERTa, XLNet, ELECTRA and BART. We used pretrained models from the *hugglingface* transformers library and fine tuned 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.

In this project, we used BERT as a baseline model and compared its performance with four other models across three annotated datasets that were designed for tweets sentiment classification tasks. The best performing model was 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 were hand labeled and used to evaluate the accuracy of the chosen model.

F. FP-Growth

To identify the correlation between users' demographics and their sentiment expressed towards nuclear energy, we mined the most frequent patterns within each sentiment of Tweets based on the labeled results of our chosen classifier. In this project, we used FP-Growth, proposed by (Han et al. 2000) to extract the frequent patterns. FP-Growth is an algorithm that can mine frequent patterns without candidate generation. Their study shows that the FP-growth method is 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.

Experiment

A. Key Topic Analysis

The number of topics used in LDA analysis was obtained 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 1. This figure shows that the coherence score for this dataset essentially levels off between 70 and 80 topics with a *very* slight increase until the edge of the experimental range of 110 topics. This result likely indicates that beyond 70 topics, the major topics within the corpus had solidified and that any additional topics added to the model simply reduced the dissimilarity within the numerous obscure topics. For convenience, we used a model with 110 topics, which resulted in the relatively high coherence score of 0.556. We then labeled the top five topics resulting from this model. Figure 2 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 nuclear research, utility usage, visions of the future, and renewable energy.

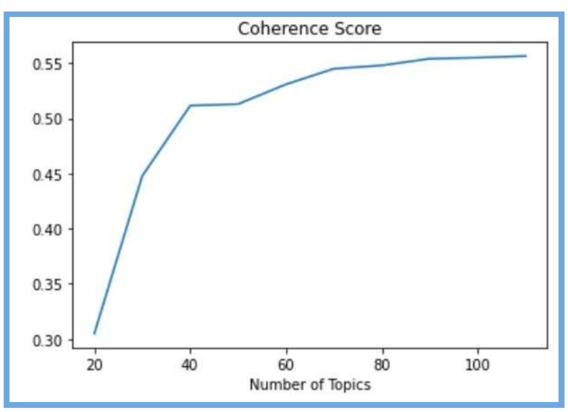

Figure 1: Coherence Scores of the LDA model on for the whole dataset

Surveys of public opinion towards nuclear energy show that negative opinions tend to be much more focused than positive opinions (Bisconti 2020), so it was expected that the most common topic would be a negative one like the discussion of nuclear waste. On the positive side, Topic 5 reflects the desire for clean, renewable energy sources which is consistently one of the primary considerations for people who favor nuclear energy in open-ended survey questions. Interestingly, there is quite a bit of technical discussion, Topics 2 and 3, which reflects the uses 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. Perhaps what's most interesting are the topics that are less frequent in our Tweet corpus. The fear of major accidents, such as Chernoble or Fukushima, is found in surveys as the primary factor for an unfavorable opinion of nuclear energy, but the topic of "Safety" was only the fourteenth most common topic in our corpus. This is likely due to the diversity of language that people use to describe this general concern and highlights an area of this methodology that could be the focus of future work on this subject.

B. Topic Analysis of User Demographic Clusters

Demographic predictions of users in our dataset resulted in a gender distribution that was 67.3% male and 32.7% female which indicates that males tweet about nuclear energy much more than female Twitter users. This matched documented gender demographic of Twitter as a whole, however, 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 (Barbera & Rivero, 2015) (Mislove, et al., 2011). The predicted age distribution was 9.2% "<=18", 20.0% "19-29", 20.1% "30-39", and 50.7% ">=40". Notice that this age distribution is heavily skewed towards older Twitter users, particularly compared to Twitter user's age distribution as a whole (Mellon, 2017). Finally the organizational status prediction indicated that only 14.4% of our users were organization accounts, however, Tweets from organizations accounted for almost a third of all Tweets in our dataset. 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 was found using the elbow method based on within-cluster variance, which resulted in a total of five clusters for our dataset. These five clusters each comprised between 500 and 1,500 users and were given the following names, in no particular order, based on their majority demographic attributes: "Females over 30", "Males over 40 with many followers", "Males over 40 with few followers", "Males under 40", and "Organizations".

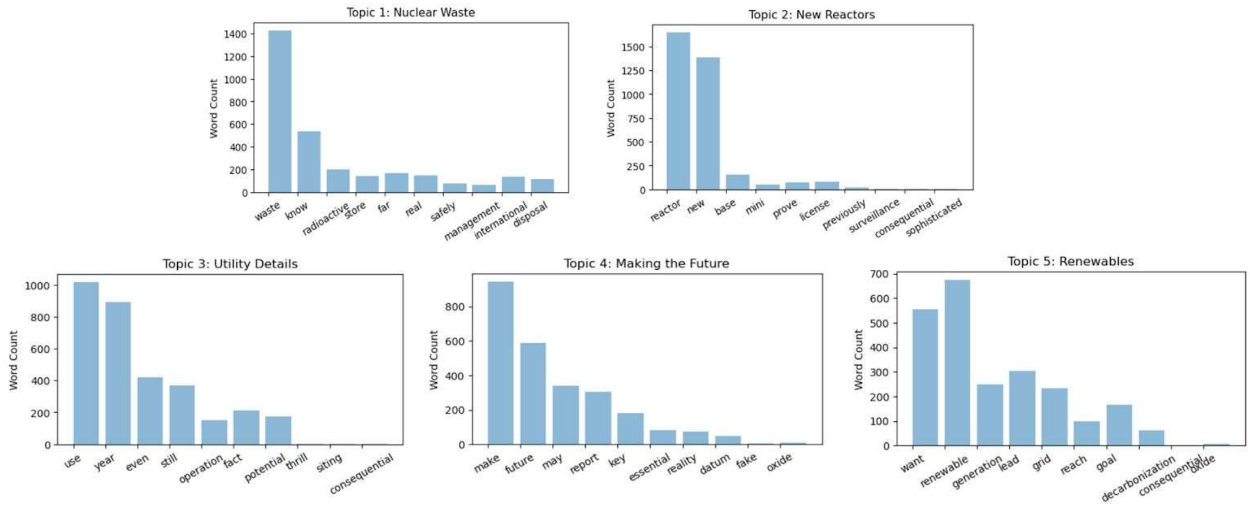

Figure 2: Word Frequency of Top 10 Topic Keywords

We then divided our dataset of nuclear energy-related Tweets according to these user clusters and conducted topic analysis to obtain the top five most popular subtopics for each demographic cluster. This result can be seen in Table 1. There were several interesting patterns that emerged from this analysis. First, the topic of “weapons” does not enter into the top five topics of the dataset as a whole but is, however, a major topic for three of the five demographics, namely, “Females over 30”, “Males over 40 with many followers”, and “Males over 40 with few followers”. This relative obscurity of “weapons” discussions in the dataset as a whole is likely due to the fact that Organizations, who generally don't tweet about weapons, overshadow discussion of nuclear weapons because, as indicated in the previous discussion, organizations have a high representation in the public discussion of nuclear energy by tweeting regularly on the topic. However, organizations don’t have complete control of the public conversation on Twitter, as evidenced by the fact that “Climate Change” was a major topic for the Organizations cluster but only appeared as the thirteenth most popular topic in the dataset overall. The implication here, that individuals don’t generally think of nuclear energy in relation to climate change, is also backed by survey data (Bisconti 2020).

Additionally, the ubiquity of the topic of nuclear waste across four of the five clusters was interesting, although not entirely unexpected, and likely significantly contributed to the overall results of the sentiment analysis discussed in detail below. Finally, an interesting anecdote is that it appears that young Twitter users are much more political and engaged in news events than older users. While it should be reiterated that our age predictions should be taken with a grain of salt, it is an unmistakable result of the topic analysis that the “Males under 40” demographic, which also included many younger female users as well, was almost exclusively focused on detailed debate and arguments surrounding political news events.

Table 1: Topic Analysis by Demographic Clusters

	Whole Dataset	Females (>30)	Males (>=40) many followers	Males (>=40) few followers	Males (<40)	Organizations
Topic 1	Waste	Waste	Potential Problems	Waste	Political Discussion	Reactor Design
Topic 2	New Reactors	Weapons	Policy	Politics	News	Making the Future
Topic 3	Utility Details	Political Discussion	Weapons	New Facilities	Detailed Debate	Climate Change
Topic 4	Making the Future	New Facilities	Waste	Weapons	Renewables	Events
Topic 5	Renewables	National Interests	Political Discussion	Reactor Design	Policy Debate	Waste

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**

Sentiment Analysis

To expand the scope of our analysis, we trained and evaluated 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 computed 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 for each user. Next, we classify the sentiment of a user by comparing a threshold with users' sentiment score. In this case, we set the threshold to 0.

A. 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 trained and evaluated each of our candidate models across three annotated datasets. The model that achieved the best average performance was then used for subsequent analysis and evaluated on a small, hand labeled sample of our nuclear energy dataset.

1. BERT - Pre-training of Deep Bidirectional Transformers for Language Understanding (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 finetuned 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 for SST-2.
2. BART - Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension (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 for SST-2.
3. RoBerta - A Robustly Optimized BERT Pretraining Approach. In a recent study, (Liu et al. 2019) found that BERT was significantly undertrained, and can match or exceed the performance of every model published after it. They propose an improved recipe for training BERT models, called RoBERTa. The large RoBERTa model achieved a 0.964 F1 score for SST-2.
4. XLNet - Generalized Autoregressive Pre-Training for Language Understanding 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. Thus, 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. Also, (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 for SST-2.
5. ELECTRA - Pre-Training Text Encoders as Discriminators Rather Than Generators (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 for SST-2.

B. *Training and Evaluation*

To fine tune these models for sentiment classification, we added an additional output layer on top of the pretrained models' output. Since we used the large version of all the models, we added a linear layer with [1024, 256, 3] dimension, and the activation function, ReLU, was applied to these layers. The AdamW optimizer (Loshchilov et al. 2017) was used to optimize the loss function. Because this was a multi-class classification task, we used the cross-entropy loss function. We trained our models on three different datasets (SemEval-2017 Task 4A, Twitter US Airline Sentiment and Apple Computers Twitter Sentiment), each of which are 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 model ELECTRA pre-training achieved an impressive average weighted macro F1 of 0.81 score across the three datasets.

Next, we trained the ELECTRA model on 20000 samples from the combination of all three datasets. The model achieved 0.79 accuracy on the combined dataset. The confusion matrix is shown in Figure 4. Lastly, we hand labeled 100 samples from our dataset to evaluate the fine tuned ELECTRA model. The model was 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 within expectation. The confusion matrix for our ELECTRA model for our hand labeled sample is given in Figure 5.

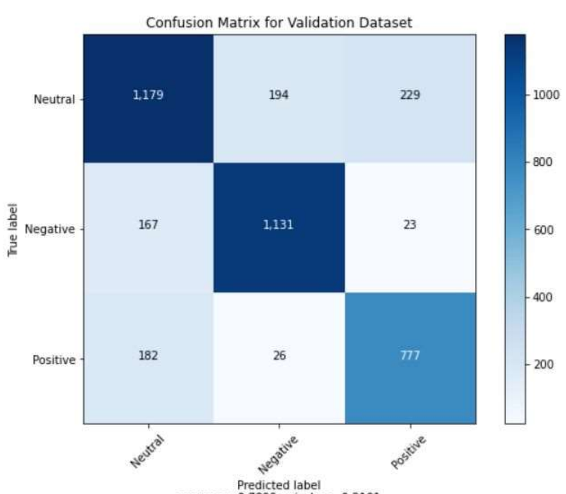

Figure 4. Confusion matrix for the Combined Dataset

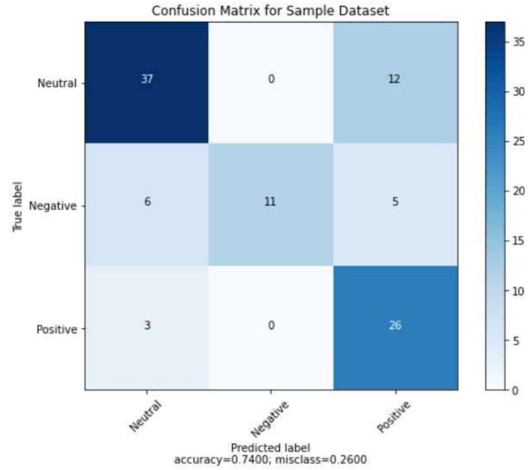

Figure 5. Confusion matrix for Sampled Dataset

C. *Frequent Patterns within user subgroup*

To distinguish different attitudes towards nuclear energy among user subgroups, we mined the frequent patterns within sentiment for both users and organizations. To do this, the user demographic predictions obtained with the M3-Inference model were combined with user data returned from the Twitter API such as the user's number of followers and whether their account is verified as an official account. All the frequent patterns are mined based on a minimum of 15% support confidence. In these results, users were put into quartiles of follower count based on the distribution of followers in our dataset (e.g. an attribute value of "top 25%" was given to users who had a follower count in the upper quartile of the users in our distribution).

D. *Analysis of Results*

a). *A positive outlook of Nuclear power.* Based on the results of our analysis, we observed large and strong positivity expressed from tweets related to nuclear energy. Out of 5477 users, 65 percent of the users expressed positive attitudes towards nuclear energy, compared to the 20 percent of users who 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.

b). *Organizations and older males are more positive towards Nuclear Power.* One important insight from our study is that different user groups express very different sentiments towards nuclear power. The results of FP-Growth revealed 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 Tweets, {'male','>=40'} is the most frequent 2-itemsets with a frequency over 3000 occurrences. Based on the frequent patterns we mined, the number of positive tweets from {'male','>=40'} are three times more than the number of negative Tweets from them. More interestingly, 64 percent of the positive Tweets are coming from organizations. In contrast, out of 3854 negative tweets, 2507 are from non-organization accounts, and only 1447 negative tweets are from organizations. This means organizations have 30% more positive Tweets than negative tweets. Lastly, 1347 positive Tweets in our dataset came from accounts with a number of followers among the "top 25%" group, which is about 2.4 times higher than that of negative Tweets for this group, suggesting that high-profile Twitter users tend to be more positive towards nuclear energy.

To further illustrate this finding, we plotted the ratio of positive to negative Tweets for each demographic class in Figure 6. This reaffirms the gender and age disparity in attitudes towards nuclear energy which we found to match 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).

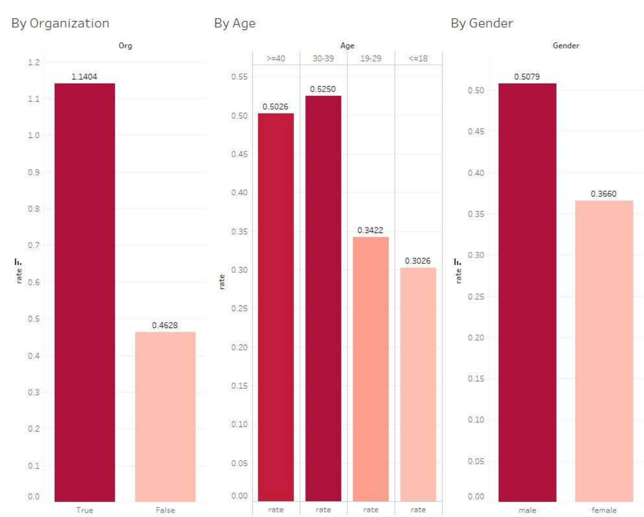

Figure 6. Ratio between pos/neg users by demographics

c). *Slight Disparity between pro-nuclear and anti-nuclear organizations.* Our analysis also reveals a disparity between how outspoken pro-nuclear versus anti-nuclear organizations are on social media. This result manifests as a positive to negative Tweet ratio greater than 1.0 for organizations. As one anecdote to this point, the largest anti-nuclear organization in the United States, Greenpeace USA, did not tweet about nuclear energy even once during the period 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 on nuclear energy in the face of the increasing threat posed by climate change.

d). *California is a strong "pro-nuclear" state.* Our analysis revealed that California is a highly "pro-nuclear" state. The ratio between the number of positive to negative Tweets in California is about 3.2, compared to the average of 2.9 across our entire dataset, which suggests users from California have favorable opinions of nuclear energy. The findings from our analysis shed lights on the emerging 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 concerns over nuclear waste and increasing electricity prices. However, this decision could be reserved if nuclear energy can be classified as "renewable energy" (Shellenberger 2019). With California's growing population and a flourishing economy majoring in high-technology, 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 raise 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 topic modeling analysis and the positive attitude of California citizens, it is possible that the decision to close out the last nuclear plant in California will be reconsidered.

Conclusion and Future Work

The motivation of this project was to overcome the limitations of public opinion surveys due to their small sample size and high cost. Additionally, we aimed to identify the key issues affecting overall public sentiment towards nuclear energy. To do this, we performed topic analysis on the major demographic clusters of users who tweeted about nuclear energy. This analysis matched published survey results suggesting that the general public does not typically think of nuclear energy in relation to climate change (Bisconti 2020). We find that it is largely the Twitter accounts of organizations and that are advocating nuclear energy as a potential zero-emissions energy source that can combat climate change.

We then found that ELECTRA had the best performance for sentiment classification of Tweets. Based on the prediction results of ELECTRA, we mined the frequent patterns of 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 negative towards nuclear energy. This research 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 user's level of knowledge about the topic, which has been shown to correlate with approval (Bisconti 2018).

One geographic trend that this project was able to highlight was the high rate of positivity that Californians showed towards nuclear energy. This 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 our work can be done by collecting Tweets over a longer time period to provide a more representative sample of sentiments from the US population. In particular, a more representative dataset could be extended to show trends across more geographic areas.

Bibliography

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.

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.

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

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.

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 International 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.

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

Nikolewski, R. 2018. Regulators vote to shut down Diablo Canyon, California's last nuclear power plant. In *LA Times*. https://www.latimes.com/business/la-fi-diablo-canyon-nuclear-20180111-story.html

Shellenberger, M. 2019. Why California May Go Nuclear. In *Forbes*.https://www.forbes.com/sites/michaelshellenberger/2019/09/03/why-california-may-go-nuclear/?sh=64067e85ae02

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.