



Public Utilities Commissions Letters

Anthony Bishop-Gylys, Danny Harmetz, Rebecca Waterson



Introduction

Public utility commissions (PUCs) regulate state power grids. During their proceedings, state residents can provide opinions that PUCs must consider; however, **despite overwhelming and bipartisan support** in these public comments for the right of consumers to sell power from solar panels back to the grid (called "net metering"), there are cases where **PUCs have ruled to the contrary**. This is **illegal, undemocratic, and hinders the renewable energy transition**, but it is difficult to evaluate the sentiment of public comments by hand in real time due to their quantity (thousands per proceeding).

Goal

Can we create a model that can accurately classify the sentiment of public comments toward net metering rights? Also, what correlations exist between sentiment and the rhetorical framing of the comments?

Methodology

- Scraped public comment text from one solar energy-related proceeding on the **Arizona PUC website**
- Annotated each comment for **sentiment toward consumer solar rights**, as well as the rhetorical framing of the contained opinion, including **whether the comment made an environmental, economic, anti-utility, justice-based, and/or liberty-based argument**
- Fine-tuned **BERT** to predict sentiment of a given comment toward the practice of **net metering energy produced from solar panels**
- Analyzed correlation trends within different comments based on rhetorical categories

Results

Sentiment Analysis:

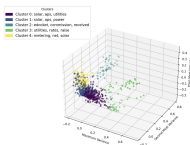
The fine-tuned BERT model obtained an average accuracy of 94.3% for 5-fold cross-validation for classifying public comments as pro- or anti-net metering. The model was cross-validated using 1009 examples of public comments with a skew of 77.2% pro-solar. This accuracy is comparable to the ~97% accuracy of current state-of-the-art BERT-based models for sentiment analysis.

Chi-Squared Statistical Analyses:

A chi-squared test was conducted with two hypotheses in mind: (1) Do public comments in support of solar (pro_solar) correlate with comments in support of the environment (enviro)? (2) Do public comments with pro_solar rhetoric correlate with comments making economic arguments (econ)? (3) Do public comments with enviro rhetoric correlate with comments containing econ rhetoric?

	Chi Sq Test	Critical Val	P Val
pro_solar vs. enviro	85.089	3.8414	5.0182e-14
Reject H0, accept HA; there IS a correlation between pro_solar and enviro			
pro_solar vs. econ	52.680	3.84145	8.5370e-08
Reject H0, accept HA; there IS a correlation between pro_solar and econ			
enviro vs. econ	11.3971	3.84145	0.32742

Accept H0, reject HA; there is NO correlation between enviro and econ



Understanding and Representing the Rhetoric Correlation Matrix:

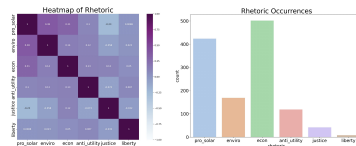
We have columns of 1s and 0s representing the presence or absence of certain rhetoric and want to find out whether there are correlations between them. The heat-map simplifies comparison of the correlations between the categorical variables through the dimension of color.

Results

K-Means Clustering:

The comments have inherent groupings based on the frequency and weight of words.

Cluster 0 represents technical discussions around utility companies. Cluster 1 focuses on renewable energy and metering. Cluster 2 deals with procedural aspects of public comment submission or regulatory processes. Cluster 3 focuses on financial aspects of utilities. Cluster 4 addresses issues or opinions regarding solar power and associated utility management.



Conclusions

The fine-tuned BERT model performed well, especially considering the limited quantity of examples used for training. Thus, such a model is likely to be useful for researchers intending to classify the sentiment of public comments as proceedings are occurring, and even more so if the quantity used to train the model is increased.

In analyzing the public comments from Arizona, we found that there are correlations between comments with pro-solar sentiment and environmental rhetoric as well as pro-solar sentiment and economic rhetoric. With the p-values being significantly below the threshold of 0.05, there is confidence in rejecting the null hypothesis.

One factor that affected our findings is that our chi-squared analyses, clustering, correlation matrix, and sentiment analysis model are based off of comments from Arizona only, so an analysis of other states may provide different results.

Challenges

Annotating enough data to fine-tune BERT was time-consuming and we only were able to annotate a minimal number of examples for training. Ultimately, however, we would like to be able to train our model on data from and apply it to proceedings from more than just one state. We would especially like to explore the applicability of the model to proceedings in a state that was not represented in the training data. We also would like to work with political scientists to synthesize our quantitative findings with a qualitative analysis of how PUCs took their public comments into account.