Modeling Grant Decision Patterns Through Natural Language Processing

Georgia Tech

Team 30: Lauren Mullininx, Krisha Shetty, Kavya Majety, Anna Bajszar, and Kristin Keith

Motivation & Introduction

The problem

U.S. research grants are sometimes canceled unexpectedly, even after initially being approved. It isn't clear why some grants get terminated while others continue. We seek to use Natural Language Processing (NLP) to determine whether the words and framing used in grant proposals influence this decision. Understanding these linguistic signals can help make government funding more transparent and potentially reduce bias in how public research dollars are allocated.

Why is it important? Why do we care?

Grant cancellations waste money, delay research, and unfairly affect certain fields or groups of researchers. If language patterns play a role, uncovering them can help make funding decisions more transparent, fair, and equitable. This data can help researchers write stronger proposals.

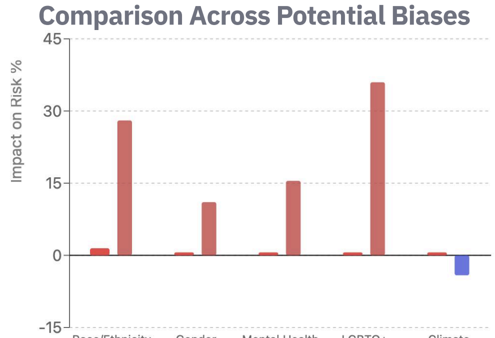
Termination Word Cloud

Our Data

- **Source:** NIH RePORTER API + HHS termination records
- **Method:** API queries and manual data scraping
- **Size:** 4,757 grants (1,850 terminated, 2,907 active)
- **Fields:** Abstract text, recipient institution, funding amount, funding agencies, etc.
- **Type:** Text + metadata (2025, temporal, labeled)

Active grants were randomly sampled from the 2025 NIH dataset, and terminated grants were selected from HHS records and Grant Witness.

Average Delta and Strongest Delta Comparison Across Potential Biases



How does our method compare to other methods?

- There are no models that we know of that uses terminated and non terminated grant data like this
- One especially interesting thing about our model is that it allows us to perform sociological bias testing in the form of counterfactual augmentation and generic term bias testing, which gives our evaluation another dimension

Our Approach

We built two main components:

1. An NLP-based grant termination classifier
2. An interactive visualization dashboard that helps users explore termination risk, keywords, fairness metrics, and demographic patterns

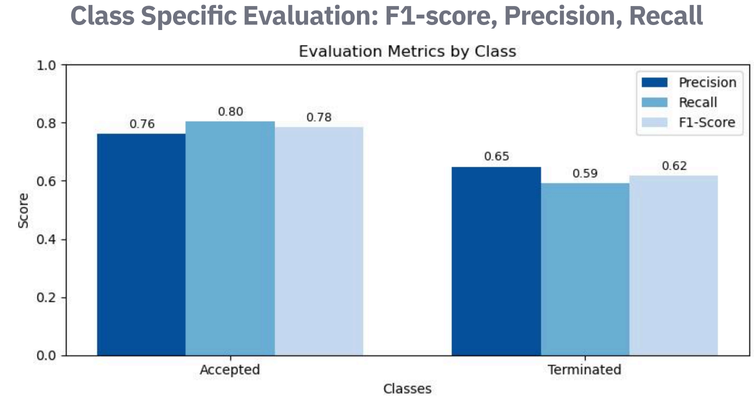
How do they work?

We first cleaned ~11,000 NIH grant abstracts, performed techniques to help mitigate data imbalance, and converted them to a TF-IDF feature matrix. Using logistic regression, we trained a classifier where terminated = 1 and accepted = 0; data was split into 80% training, 20% testing. The model was evaluated with accuracy, F1-score, confusion matrix, and ROC-AUC. We then connected model outputs to graphics: word cloud, bias bar charts, filters, and grant details.

Why our approach is relevant

We believe that language can reflect priorities, clarity, and perceived value in grant proposals. ML models can identify subtle patterns that humans might overlook, while visualizations make these patterns understandable to non-technical users. Past studies have either analyzed solely text or fairness; our project is the first to combine linguistic modeling, outcome prediction, and interactive bias visualization to interpret U.S. grant cancellations.

Class Specific Evaluation: F1-score, Precision, Recall



Experiments & Results

General evaluation metrics

- Accuracy, Precision, Recall, F1-score (macro and per class)
- Confusion matrix (absolute and normalized)
- ROC curve and ROC-AUC score (0.78)

Bias and fairness evaluation

- Counterfactual augmentation was done by creating a duplicate TF-IDF vector and editing zeroing/replacing terms relating to bias.
- By calculating the difference of the probability predicted from the TF-IDF vectors, we gain an understanding of bias.

Results

Metric	Score
Accuracy	72.38%
ROC-AUC	0.7819
Macro Precision	0.706
Macro Recall	0.698
Macro F1-Score	0.701

Bias Category	Strongest Delta	Avg Delta
Climate Change	0.041433	-0.000136
Race/Ethnicity	-0.280540	-0.014777
Mental Health	-0.154946	-0.001670
LGBTQ	-0.359854	-0.005631
Gender	-0.110595	-0.000459