Introduction

Our online deliberation platform, developed jointly with Stanford University, is a video-based chat room where people can discuss various topics such as immigration policy or refining the voting system. This is meant to assist with policy and decision making, aided by an auto moderator that handles the agenda of the discussion. Transcripts of the discussion can also be used to gauge public opinion on various issues.

One of our primary goals is to enhance the automated moderator to strengthen its ability to analyze the current state of the deliberation and address issues. Specifically, if there’s a lull/silence in the discussion, the automated moderator should determine if it’s because:

- Users have fully discussed the topic, in which case it will advance to the next item in the agenda.
- Users haven’t discussed enough, in which case it will nudge users to discuss more.
- Users are only focusing on one side of the debate, in which case it will notify users to discuss the other side more.

Objectives

There are two main objectives for our project:

- **Overall Goal:** Given the entire deliberation up to a certain moment, determine the current state of the conversation, i.e., whether the discussion is sufficient and balanced.
- **Subgoal:** Given a single participant’s speech, determine whether it’s related to the pros or cons of the current topic they are discussing.

- **Front-end integration:** Display the list of pros and cons on the agenda.

**Sentiment Analysis**

- **Approach 1**
  - Use **Sentiment Analysis** to detect positive/negative tone
  - Approximate pros/cons label from sentiment

- **Accuracy:** 42.18%

**Topic Modeling**

- **Detect keywords and topic of pre-made agenda items (pro/con arguments provided by political scientists).**
- **Use Latent Semantic Analysis model to generate sentence similarity scores between users’ speech in deliberations and agenda items. (Assumptions that words that are close in meaning will occur in similar pieces of text)**
- **A dictionary and document term matrix (DTM) is created, then singular value decomposition is used on the DTM to retain most significant dimensions of DTM to find the most prevalent topics in text. We use metrics like cosine similarity to gauge similarity between input words (user’s speech) and dictionary words (agenda items).**
- **Classify a user’s sentence as neutral, pro, or con, a neutral statement based on whether the top match (agenda item with the highest sentence similarity score) is a pro or a con.**

**Accuracy:** 48.36%

**Discussion**

Problems we have identified:

- **Not recognizing synonyms**
- **Certain important words in the topic appear in both pros and cons, which confuses models**
- “Federal”, “illegal immigrants” etc; commonly mentioned in both sides of arguments
- Some pre-made agenda points are not well-phrased and confuses models.

Other challenges we faced:

- Limited data (only 10 deliberations)
- **Inaccurate audio to text transcriptions by Google API**

**Future Steps**

**Text Similarity Models:**
- **Train model using more context, e.g., online articles**
- **Transform input into bigrams instead of individual words**

**Front-end Integration:**
- **Allow human observers to provide true labels on pro/con classification and the actual argument discussed**
- **Enable AI moderator to take actions based on model results**

**Others:**
- **Experiment with RST trees, other similarity models, NLP tools**
- **Combine topic modeling and sentiment for better results**

Acknowledgments

We would like to thank Professor Kamesh Munagala, and our Graduate Mentor Sravya Yandamuri, for their guidance and support, as well as Duke Summer Projects in Computer Science. We also appreciate help from members of the Stanford Social Algorithms Lab: Professor Ashish Goel, Lodewijk Gelauff, and Sukolsak Sakhaewong.