Rhetorical Polarization on C-SPAN

PURDUE UNIVERSITY | The Data Mine

CCSE Center for C-SPAN Scholarship & Engagement

Neandrea Lee, Dustin Santoso, Ciarra Ashley, Jax Mader, Harry Jarvis, Isaac Lugo, Brandon Miramontes, Abraham Nunez, Marcella Thomas, Sreeram Kondapalli, Elizabeth Adamson, Tatum Beaugh, Angeluis Sayavongsa-Vaughn, Alexander Brenham

Detecting rhetorical polarization in the United States Congress using a multimodal approach.
Political polarization refers to the growing divide in political attitudes, where individuals or groups adopt increasingly extreme positions, leading to a significant gap between opposing viewpoints. This phenomenon often results in the clustering of opinions at two distant and antagonistic poles, reducing common ground and complicating consensus-building (Jones, 2021). Often times, this phenomenon is reflected in the types of speeches we see on C-SPAN (Dietrich, 2019), making present study important to understanding the evolution of polarization on Capitol Hill.

C-SPAN Data
We utilized a corpus of floor speeches from the 2009 US Congress, sourced from C-SPAN's archives. Our analysis encompassed 8,505 speeches averaging around 2 minutes in length.

Sentiment labels:
This project focused on sentiment classification (positive, negative, neutral). Our research team developed a codebook through initial labeling, sorting 250 speeches to create a training foundation after establishing interrater reliability.

Positive	Neutral	Negative
"…is actually a source of pride for him and for the– for his constituents in Missouri…."	"Amendment Number 6, printed in House Report Number 111-36, offered by Mr. Ovestar of Minnesota"	"…this bankruptcy amendment will only lead to a worse situation… is no solution at all."

Figure 1: Examples of Positive, Neutral, and Negative Sentiment

Methodology: Our team developed a deep learning model for sentiment classification using a **recurrent neural network (RNN)** with a categorical cross entropy loss function:

- We processed transcriptions by tokenizing and padding text sequences before splitting into training and testing sets.
- Our RNN architecture used the following:
  - Word embeddings (100D)
  - Dual LSTM layers (64/32 units),
  - Dropout regularization (30%)
  - Softmax output layer for three-class prediction.
  - Trained with Adam optimization over 50 epochs

We also used a **convolutional neural network (CNN)** to classify sentiment using frames from the training videos. CNNs are a recommended tool for image classification. Our CNN comprises of the following layers:

- RELU (Rectified Linear Unit) – a nonlinear activation function layer that returns zero if it receives any negative input, but for any positive value, it returns that value.
- DROPOUT and POOLING – reduces spatial dimensions and reduces the likelihood of over fitting.

To classify sentiment using audio, we used a **recurrent neural network (RNN) with Mel Frequency Coefficients** (MFCCs). These were acquired after audio was extracted from the videos using FFmpeg.

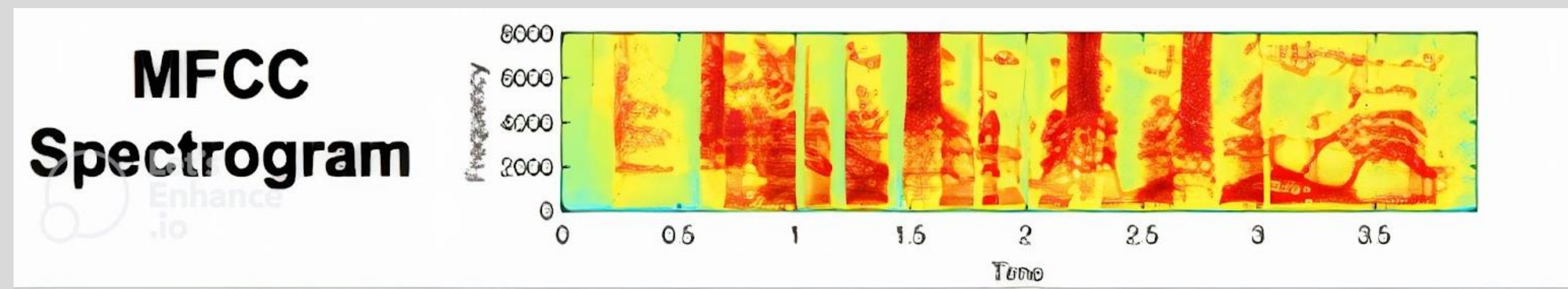

Figure 2: Example of a Mel Frequency Coefficient (MFCC) Spectrogram similar to what was used in this study.

MFFC Spectrogram - MFCCs convert sounds into numbers by analyzing frequencies, filtering important sounds, and simplifying data for speech recognition and audio analysis. A spectrogram is shown above which highlights frequency peaks on the y-axis across the time dimension on the x-axis. The color changes are created based on decibel levels.

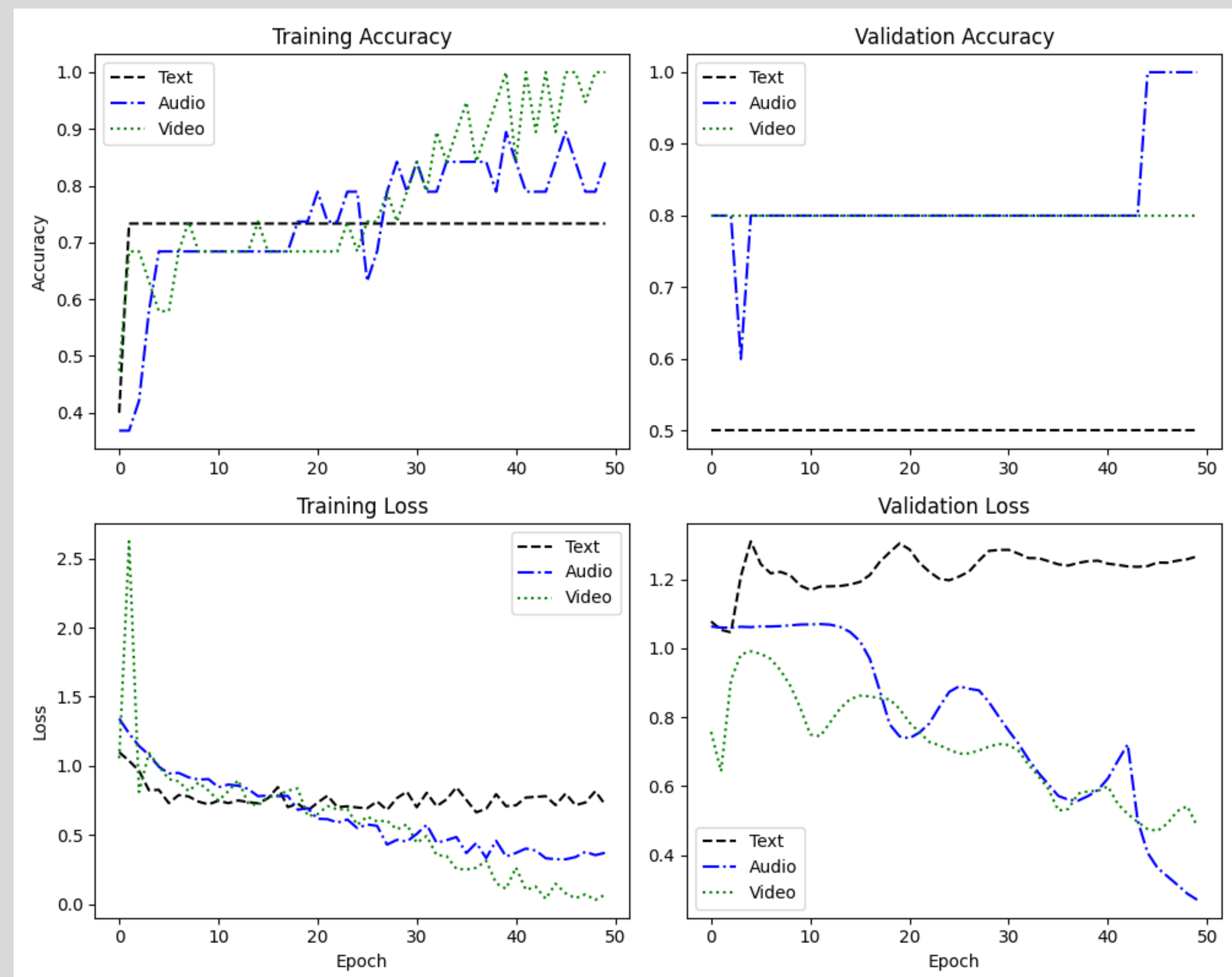

Figure 3: Training and Validation Performance for Text, Audio, and Video Neural Networks

Training and Validation Performance (See Figure 3):

- Text Analysis - Training and Validation accuracy change rapidly then stagnate, indicating potential overfitting. Validation and training loss stabilize to 1.27 and 0.72 respectively.
- Audio Analysis – Validation accuracy increases sharply while validation stagnates, indicating potential overfitting. Final validation and training loss are 0.27 and 0.37, respectively.
- Image Analysis - Validation accuracy remains largely unchanged, indicating potential overfitting. Final validation and training loss are 0.48 and 0.07, respectively.

- **Do Some Issues Increase Rhetorical Polarization (See Figure 4)?**
- Research Question – In 2009, healthcare and the economy were hot button issues. Does positive/negative sentiment differ across these issues?
- Approach – We used the text, audio, and image neural networks trained using our sample data to produce predicted sentiment for all 2009 speeches, then subset the data using healthcare and economy keywords.
- Healthcare keywords - "reform," "medicine," "insurance," "healthcare," "medicaid," "medicare," "doctors," "hospitals," "insured," and "premiums"
- Economy keywords - "stimulus," "bailout," "tarp," "unemployment," "recovery," "foreclosure," "jobs," "stocks," "economy," and "economic"

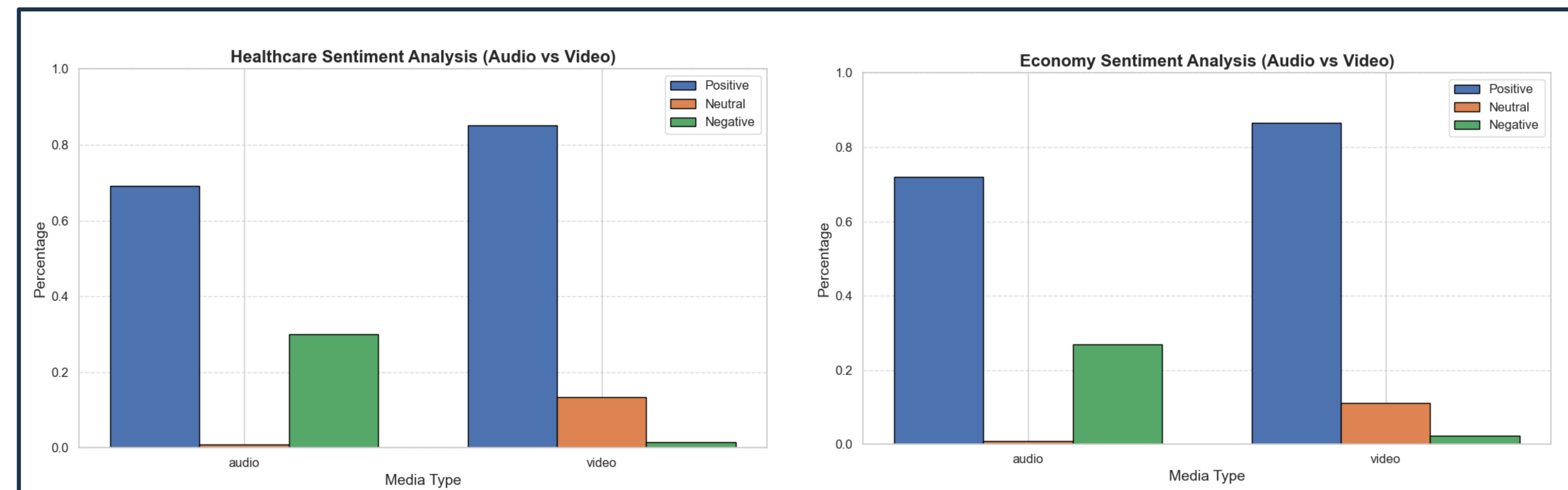

Figure 4: Assessing Sentiment of 2009 Congressional Speeches Withing Healthcare and Economic Debates

- We found the text model tended to produce the same result, so it was exclude from Figure 4.
- In Figure 4, the audio model tended to produce more negative labels than the video model.
- No significant difference between healthcare and economic sentiment ($X^2 = 0.004$, df = 1, p = 0.95)

References

Conclusion: Ultimately, we found audio and video classification worked best when assessing sentiment. We also found no significant difference between healthcare and economic sentiment.
Future Directions: Expand time series. Observe different key words in speeches and attempt a combined multimodal classification model after overfitting issues with the base models are addressed. We would also like to explore how rhetorical polarization changes in response to an external shock, like the Covid-19 pandemic or the recent tariff crisis.