Meeting Notes

- Collecting Data
    - See if deepseek really has the data or if it can make the data for left and right media
    - If not work out some way to collect large amounts of left and right data

Individual Words:

**Simplified Political Bias Score Formula**

$$\text{BiasScore}(w) = \tanh\left(\lambda \cdot \log_2\left(\frac{\text{rel\_R} + \epsilon}{\text{rel\_L} + \epsilon}\right) \cdot \text{IDF}(w) \cdot (\text{sentiment\_R} - \text{sentiment\_L})\right)$$

**Where**:

- $\text{rel\_L}, \text{rel\_R}$: How often the word appears in left/right texts (normalized by corpus size).
- $\log_2(\dots)$: Measures if the word is more common on the right (+) or left (-).
- $\text{IDF}(w)$: Penalizes boring words (e.g., "the") and boosts rare/political words.
- $\text{sentiment\_R} - \text{sentiment\_L}$:
    - **+value**: Right uses it positively / Left uses it negatively.
    - **-value**: Left uses it positively / Right uses it negatively.
- $\tanh(\dots)$: Squishes the final score between **-1 (left)** and **+1 (right)**.
- $\lambda$: Sensitivity knob (default: $0.1$–$0.3$).

**Thought Process (Plain English)**

1. **Frequency Check**:
    - If "woke" appears **10x more on the right**, the raw count suggests right bias.
    - *But*: Right might use it **negatively** (e.g., "anti-woke"). So we need sentiment!
2. **Sentiment Adjustment**:
    - Left uses "woke" positively (+0.8)? Right uses it negatively (-0.6)?
    - Then: $(\text{sentiment\_R} - \text{sentiment\_L}) = -0.6 - 0.8 = -1.4$.
    - **Negative value flips the bias to the left**, even if right says it more!
3. **IDF Penalty**:

### Why It Works

- **"Woke" Example**:
  - Right says it more (frequency → right), but *negatively* (sentiment → left).
  - Result: Left-leaning score (**-0.68**). Matches reality!
- **"Traditional" Example**:
  - Right says it more (+frequency) *and* positively (+sentiment).
  - Result: Right-leaning score (**+0.22**).
- **"Trump" Example**:
  - Both sides say it a lot (frequency cancels out), but right is more positive.



### When to Tweak Parameters

1. $\lambda$: Increase (e.g., 0.3) for stronger polarization signals.
2. **IDF**: Use stricter $\mathrm{df}(w)$ thresholds to filter common words.
3. **Sentiment**: Use better lexicons (e.g., train on partisan hashtags).

Need scores for other words? Just plug in the numbers!

CUMulative score for the article: