Computational Analysis of the Discourse around Roe v. Wade

Sujan Dutta, **Harshit Rajgarhia**, **Shrinivas Sampath Muthupalaniyappan**
Rochester Institute of Technology
{sd2516, hr9397, sm3533}@rit.edu

1 Introduction

On June 24, 2022, the US Supreme Court overturned Roe v. Wade, the landmark piece of legislation that made access to abortion a federal right in the United States. The decision dismantled 50 years of legal protection and paved the way for individual states to curtail or outright ban abortion rights.

A US presidential committee in 1981, after years of deliberations involving professionals across every relevant field, proposed a scientifically derived line between life and death [for the Study of Ethical Problems in Medicine *et al.*1981]. The committee's criteria have served as a foundation for laws in most states adopting a standard for legal death. However, due to the constant change in laws regarding abortion and reproductive rights and shifts in powers between federal and state, there has been an ever-growing disagreement in public opinion when it comes to answering the question *When does life begin?*. Many political surveys indicate the shifts in public opinion regarding Abortion laws [KFF, Arney and Trescher1976]. Politicians and polarized news media channels could play a potential role in shaping public opinions around these topics and such opinions could potentially be harnessed by exploiting NLP techniques using social media interactions. More specifically, we treat this stance mining problem as an interval estimation problem. At two extremes we have *life begins at conception* and *life begins at birth*. Using language models, we try to estimate where the majority of public opinion lies.

The goals of this collaborative work are: a.) Building contemporary text corpora using user comments (related to abortion laws) collected from the official YouTube Channels of major US cable networks b.) Leverage state-of-the-art NLP algorithms and language models to aggregate public opinion on topics surrounding laws and rights to abortion c.) Using language models to mine the public stance on *When does life begin?* and profiling them according to their ideologies and political proclivity

Although this proposed work will majorly focus on the topics related to abortion and reproductive rights, the techniques used could apply to mine public stance on other debatable issues like *What should be the minimum wage?*, *What should be the social security benefit amount?* etc. Also, our methods could potentially be applied to issues in other countries and other languages.

2 Dataset Description

We consider the dataset of YouTube comments provided by the course instructors. This dataset contains more than 92 million comments on 241,112 videos posted between 2014, January 1 to 2022, Aug 27 by the official YouTube handles of three major US news networks: CNN, Fox News, and MSNBC. We use a keyword-based approach to extract the relevant comments for our analysis. More specifically we first build a set of 20 keywords relevant to the discussion of Roe v. Wade and extracted the comments that contain at least one of the keywords. Below we list the most frequent five keywords:

1. `roe v. wade`
2. `abortion`
3. `pro-life`
4. `pro-choice`
5. `unborn`

Using this filtering technique, we create a large-scale corpus of relevant comments for each network denoted by $\mathcal{D}_{cnn}$, $\mathcal{D}_{fox}$, and $\mathcal{D}_{msnbc}$. Table 1 presents the distribution of comments in each corpus.

Table 1: Network-wise distribution of comments

News network	#comments
CNN	$108,715$
FOX News	$109,298$
MSNBC	$98,195$

3 Exploring Qualitative Differences Across Different Networks

The motivation behind this section is to use simple NLP tools to understand the qualitative differences that might exist in the comments posted on different news channels. This exploratory analysis also helped us design more sophisticated experiments that we discuss in the later sections.

3.1 Experimental setup

We choose bi-gram frequency as a quick tool to identify the differentiating features in the data. Instead of finding popular bi-grams in the entire dataset, we focus on some relevant phrases and look into their right context. The first phrase we target is `abortion is` (S_1). So our experiment is simply finding the popular bi-grams that people use after S_1 across different channels. But if we consider all the comments where S_1 is mentioned, we run into the following problem - both the phrases *"abortion is woman right"* and *"abortion is not woman right"* contributes to the bi-gram `woman-right` which might result in incorrect analysis. To tackle this, for each network, we create a subset of comments after removing comments that contain negations (i.e. `not`). We denote these subsets as $\mathcal{D}'_{cnn}$, $\mathcal{D}'_{fox}$, and $\mathcal{D}'_{msnbc}$. We run our bi-gram analysis both on the original datasets and these subsets. Besides S_1, we also consider the phrase `I am`. The goal is to find out how people might self-identify.

3.2 Results

In order to interpret the bi-grams, we create wordclouds of different datasets. Figure 1 shows the common bi-grams that come after `abortion is`. In $\mathcal{D}_{fox}$ we note a very surprising term `kkkible` which is created by replacing the 'B' in Bible with 'kkk'. Upon further investigation, we found that there are some 50 users who repeatedly used this term in their comments. `constitutional-right` is prominent in all the datasets except $\mathcal{D}'_{msnbc}$. We note that the relative importance of `woman-right` is significantly more in $\mathcal{D}_{cnn}$ than other datasets. Where as in $\mathcal{D}_{msnbc}$ and $\mathcal{D}'_{msnbc}$ `violent murder` and `premedicated violent` are very prominent. Our analysis reveals that while the constitutional right is discussed by users on every channel, some channel viewers emphasize woman's rights, and some on violent murder.

Figure 2 presents the popular bi-grams occurring after `I am`. We found that `pro-life` is most popular in $\mathcal{D}_{fox}$ and $\mathcal{D}'_{fox}$ followed by `pro-choice`. But in $\mathcal{D}_{msnbc}$ and $\mathcal{D}'_{msnbc}$ the ranking flips. Again in $\mathcal{D}_{cnn}$ and $\mathcal{D}'_{cnn}$ we see `pro-life` at the top. Besides these, people also wrote `pro-abortion` and about their voting preferences (`voting-trumppence`). These results motivated us to look into the users who self-declared their stance as either pro-life or pro-choice. Table 2 shows the network-wise distribution of users who explicitly declared their stance in the comments. We also aggregate the results and found that a total of 916 unique users declared themselves as pro-choice and 856 users as pro-life.

Table 2: Distribution of users who explicitly declared their stance across networks

-	I am pro-life	I am pro-choice
CNN	377	447
FOX News	440	348
MSNBC	230	334

4 Exploring the User Behaviour at Comment Level

Understanding the overall idea in user comments in any digital media corpus is important. Analyzing the comments from the users will reveal the latent topology behind a corpus level or at a group level. The following corpus is used $\mathcal{D}_{cnn}$, $\mathcal{D}_{fox}$, and $\mathcal{D}_{msnbc}$ throughout our two experiments. Moreover, the analysis obtained from these experiments helps design our future experiments accordingly.

4.1 Topic Modeling

Interpreting core themes associated with a document collection is a fundamental task in today's information era. Topic Modeling is the new revolution in text mining. Especially using vector space models for topic modeling can help uncover the latent structure in the corpus. Also known as an unsupervised way of identifying topics by aggregating similar comments which have semantically the same structure. In order to identify the topics in an unsupervised way, there are three most important steps, Figure 3 shows the approach we used for Topic Modeling. The first step in topic modeling is to convert the corpus to embedding, we used a sentence transformer from hugging face called SBERT [Reimers and Gurevych2019] to convert the documents to embedding. Then in our next step, the vector representation is reduced by certain dimensions so the clustering algorithm can capture the local clusters. We used UMAP [McInnes *et al.*2018] a dimensionality reduction algorithm to reduce the dimensions. After reducing the vectors to certain dimensions a clustering algorithm called HDBSCAN [McInnes *et al.*2017] is used to cluster the documents.

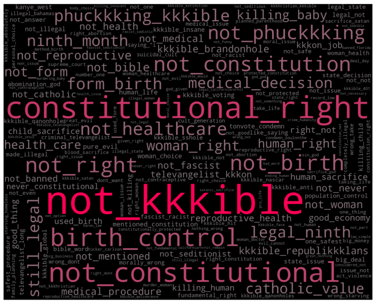
In this experiment we are trying to obtain answers to the following questions i) Identifying prominent topics in user comments at the channel level and ii) Exploring the topics involved in pro-choice and pro-life at the channel level.

4.1.1 Channel Level Topic Exploration

In this experiment we consider the three following corpus $\mathcal{D}_{cnn}$, $\mathcal{D}_{fox}$, and $\mathcal{D}_{msnbc}$. To avoid bias while clustering, we removed the keyword "abortion" before applying the Topic Modeling approach (in Figure 3) to all three news channels. Hyperparameter tuning is an important step to capture the global cluster instead of the local cluster. The following three parameters are set to a certain value as per Table 3 to capture the global cluster.

Table 3: Topic Modeling Parameters

Parameters	#value
n_neighbours (UMAP)	200
min_cluster_size (HDBSCAN)	5000
min_sample_size (HDBSCAN)	2500



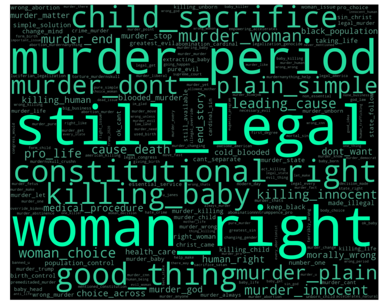
(a) $\mathcal{D}_{fox}$

(b) $\mathcal{D}_{msnbc}$

(c) $\mathcal{D}_{cnn}$

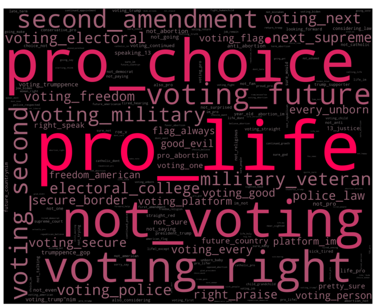
(d) $\mathcal{D}'_{fox}$

(e) $\mathcal{D}'_{msnbc}$

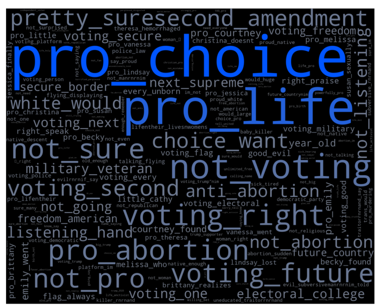


(f) $\mathcal{D}'_{cnn}$

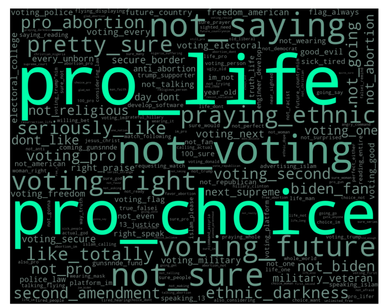
Figure 1: Popular bi-grams after the phrase `abortion is` in different datasets



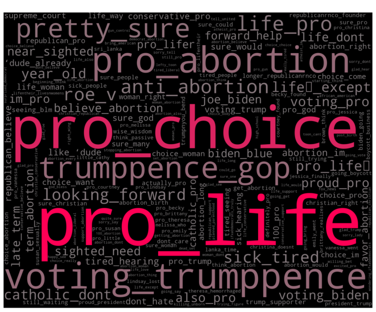
(a) $\mathcal{D}_{fox}$



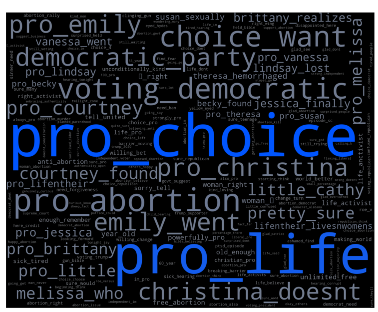
(b) $\mathcal{D}_{msnbc}$



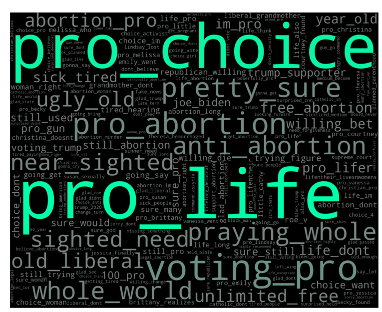
(c) $\mathcal{D}_{cnn}$



(d) $\mathcal{D}'_{fox}$



(e) $\mathcal{D}'_{msnbc}$



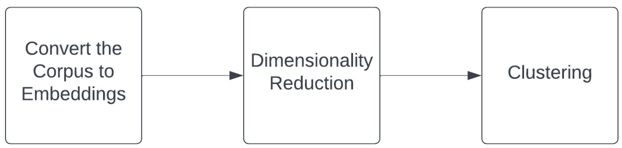
(f) $\mathcal{D}'_{cnn}$

Figure 2: Popular bi-grams after the phrase `I am` in different datasets

4.1.2 Results

The results of the topic modeling are shown in the form of wordcloud in Figure 4. The most prominent topics in all three

Figure 3: Topic Modeling Approach



channels are i) row_v_wade ii) constitutional rights iii) covid vaccine.

4.1.3 Investigating Pro Choice vs Pro Life

In order to understand the topics involved in the pro-choice and the pro-life debates in user comments. We tried to filter out the comments containing the word pro-choice and pro-life from each channel. This resulted in two new corpora for each channel. The corpus size is mentioned in Table 4. For each corpus mentioned in Table 4, we applied the topic modeling approach from section 7.1

Table 4: Pro Choice and Pro Life Corpora Size

Channels	**#Pro-choice**	**#Pro-life**
CNN	1615	2363
FOX	1874	1234
MSNBC	1264	1777

4.1.4 Results

This experiment resulted in interesting topics being found in each channel for their subcategories pro-choice and pro-life. Table 5 shows the topics about criticizing democrats were more prominent in the pro-choice column in all the channels. Interestingly, in the pro-life column, the CNN channel contained topics supporting pro-choice only in the case of rape, while the MSNBC channel contained topics supporting the Republican party.

Table 5: Pro Choice vs Pro Life Topic Modeling

Channels	**#Pro-choice**	**#Pro-life**
CNN	Criticize Democrats Criticize Women Criticize CNN Using Religion	Rape - (Pro-Choice) Criticize Biden Roe v Wade Overturn
FOX	Criticize Democrats	Criticize Republican Support Republican
MSNBC	Criticize Democrats	Criticize Republican Support Republican

4.2 Analyzing Before and After Roe v Wade Overturn

In this section, there will be two experiments carried out to find out any differences between before and after the roe v wade overturn. The first experiment focuses on the temporal distribution of the data extracted from the original corpus provided by the instructor. The focus of the second experiment is to calculate the unigram frequency of the words before and after the overturning of Roe v. Wade and examine any changes.

4.2.1 Temporal Distribution

In this experiment, we will filter the data provided by the instructor by searching for comments containing the keyword "abortion". The total data spans from 01-Jan-2014 to 09-Sep-2022. The size of the newly fetched $\mathcal{D}_{cnn}$, $\mathcal{D}_{fox}$, and $\mathcal{D}_{msnbc}$ are listed in the Table 6. To calculate the Temporal distribution for the entire corpus all three channels are added together and made into one. The size of the newly made corpus is 2,93,022 comments.

Table 6: Temporal Distribution network-wise comments fetched only using the keyword "abortion"

News network	**#comments**
CNN	$112,376$
FOX News	$95,734$
MSNBC	$84,912$

4.2.2 Results

Roe v Wade was overturned on 24^{th} June 2022, from Figure 5 after the year 2022 there is a clear spike indicating the user's distribution of the comments is much higher during that period. This shows further experiments need to be carried out to understand user behavior before and after Roe v Wade was overturned.

4.2.3 Analyzing Unigram Frequency Before and After Roe v Wade Overturn

In this experiment, we analyzed the frequency of individual words before and after the overturning of Roe v. Wade. We subtracted the frequency of similar words and normalized the data by dividing them by each other. The results of the experiment are discussed in the results section.

4.2.4 Results

The results of the experiment are shown in Table 7. It appears that the words "blue" and "scotus" (an abbreviation for the Supreme Court of the United States) were used more frequently after the overturning of Roe v. Wade. Similarly, "pro-life" and "democracy" also showed increased usage after the decision. Further analysis of the most commonly used words, such as "blue" and "republican," may provide insight into potential changes in voting behavior.

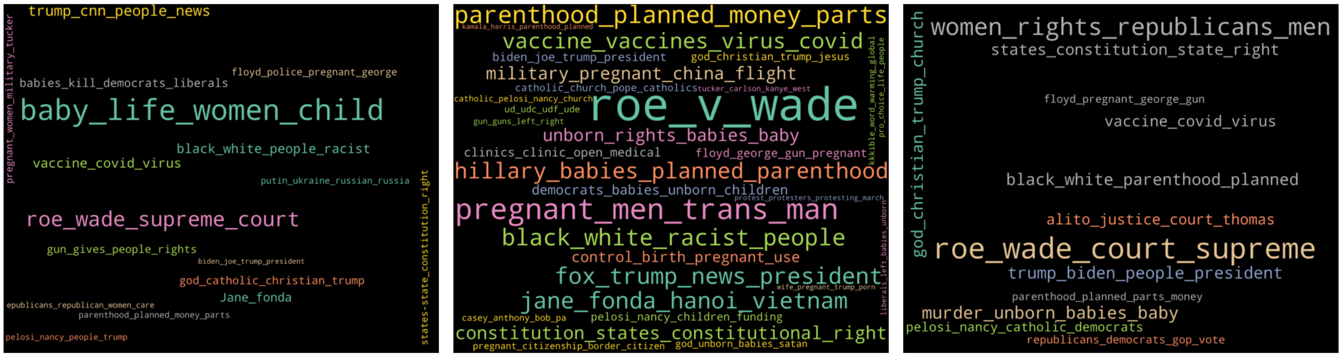


Figure 4: Topic Modeling performed for the following channels (left) CNN, (middle) FOX & (right) MSNBC

Figure 5: Temporal Distribution

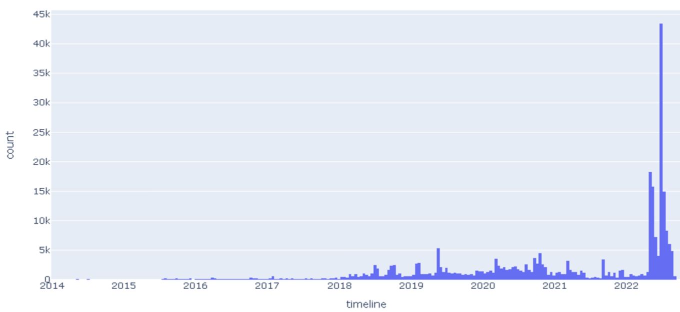


Table 7: Unigram Differentiation

Unigram	#Normalized Difference
scotus	0.155
blue	0.037
rape	-0.164
contraceptive	-0.228
overturn	-0.254
democracy	-0.429
prolife	-0.431

5 Analyzing Sub-topics Through User Behavior

So far, we have presented our analysis only at the comment level. Here, we study the behavior of different groups of users. The research question we try to answer in this section is - what does the user behavior tell us about the interrelationship among the sub-topics?

5.1 Experimental Setup

In order to consider all the users who commented at least once in any one of the channels, we club together $\mathcal{D}_{cnn}$, $\mathcal{D}_{fox}$, and $\mathcal{D}_{msnbc}$. We then create 12 groups of users based on their participation in discussions related to different sub-topics. These sub-topics are chosen in a way to capture most of the issues concerning Roe v. Wade and a few other politically significant topics (e.g. voting, gun law, racism). We also find these sub-topics as prominent features in our topic modeling experiments and bi-gram analysis. Below is the list of sub-topics:

1. `gun law`
2. `science`
3. `bodily autonomy`
4. `religion`
5. `constitutional right`
6. `human right`
7. `pro-choice`
8. `pro-life`
9. `racism`
10. `reproductive right`
11. `women's right`
12. `voting`

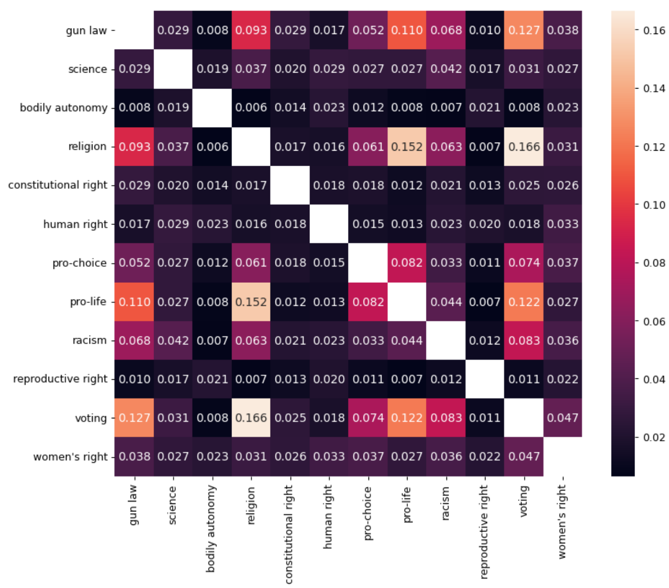
A simple keyword-based method is used to create user groups from these topics. If a user mentions any topic-related keyword in one of their comments in the merged dataset, we add that user to the corresponding set. After creating the 12 user sets we are ready to analyze the user engagement patterns among the sub-topics. We use the Jaccard index (Eq. 1) to compute the user overlap between the sub-topics. Measuring the user overlap might reveal the issues that are discussed together by similar users and the issues that are discussed in isolation.

$$J(\text{group}_A, \text{group}_B) = \frac{\text{group}_A \cap \text{group}_B}{\text{group}_A \cup \text{group}_B} \tag{1}$$

5.2 Results

We create a user co-occurrence matrix (Figure 6) to study the amount of user overlaps between every pair of sub-topics. Some of the sub-topic pairs with very high user overlap are - $\langle$`voting, religion`$\rangle$, $\langle$`pro-life, religion`$\rangle$, $\langle$`voting, gun law`$\rangle$, $\langle$`voting, pro-life`$\rangle$, $\langle$`pro-life, gun law`$\rangle$, $\langle$`religion, gun law`$\rangle$, and $\langle$`racism, voting`$\rangle$. On the other hand, we found `bodily autonomy` and `reproductive right` have significant overlap with only a few sub-topics (`human right, women's right`). This matrix reveals the user overlap for the pair $\langle$`voting, pro-life`$\rangle$ is significantly higher than $\langle$`voting, pro-choice`$\rangle$. The same pattern is found for the pairs ⟨`religion, pro-life`⟩ and ⟨`religion, pro-choice`⟩. We also note some interesting pairs with significant user overlap like - ⟨`racism, science`⟩, ⟨`racism, religion`⟩.

Figure 6: User Co-occurrence Matrix



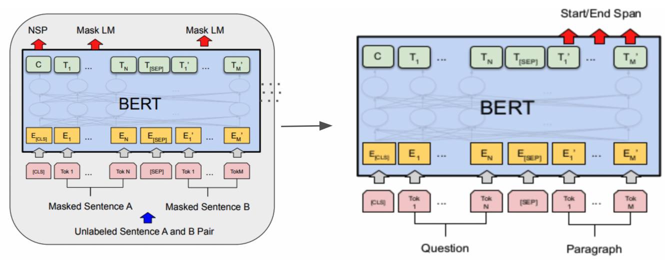
6 Leveraging Question Answering System in Mining Public Opinion

6.1 BERT Question Answering

Language model pre-training has been shown to be effective for improving many natural language processing token-level tasks such as question answering. Pre-trained BERT Question Answering model is used in this experiment. In BERT [Devlin *et al.*2018], there are two steps in the framework: pre-training and fine-tuning. During pre-training, the model is trained on unlabeled data over different pre-training tasks. For fine-tuning, the BERT model is first initialized with the pre-trained parameters, and all of the parameters are fine-tuned using labeled data from the downstream tasks. Each downstream task has separate fine-tuned models, even though they are initialized with the same pre-trained parameters. One of the Downstream tasks of BERT is a question answering whose pre-training and the fine-tuning procedure is shown the Figure 7 below.

In the question-answering task, the input question and passage are represented as a single packed sequence, with the question using the A embedding and the passage using the B embedding. Only a start vector $S \in \mathbb{R}^H$ and an end vector $E \in \mathbb{R}^H$ are introduced during fine-tuning. The probability of word i being the start of the answer span is computed as a dot product between T_i and S followed by a softmax over all of the words in the paragraph: $P_i = \frac{e^{S \cdot T_i}}{\sum_j e^{S \cdot T_j}}$. The analogous formula is used for the end of the answer span. The score of a candidate span from position i to position j is defined as

Figure 7: BERT Question Answering

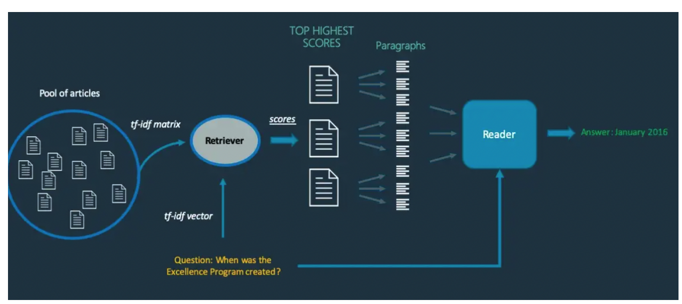


$S \cdot T_i + E \cdot T_j$, and the maximum scoring span where $j \geq i$ is used as a prediction. The training objective is the sum of the log-likelihoods of the correct start and end positions.

6.2 Closed Domain Question Answering

The Closed Domain Question Answering aka cdQA [Félix MIKAELIAN] architecture is based on two main components: the Retriever and the Reader. You can see below in Figure 8 a schema of the system mechanism. When a question is sent to the system, the Retriever selects a list of documents in the database that are the most likely to contain the answer. The Retriever creates TF-IDF features based on uni-grams and bi-grams and computes the cosine similarity between the question sentence and each database document. After selecting the most probable documents, the system divides each document into paragraphs and sends them with the question to the Reader, which is basically a pre-trained BERT-based Question Answering model as discussed in Section 6.1. Then, the Reader outputs the most probable answer it can find in each paragraph. After the Reader, there is a final layer in the system that compares the answers using an internal score function and outputs the most likely one according to the scores.

Figure 8: Mechanism of cdQA pipleine



6.3 Mining Public Opinion

6.3.1 Experimental Setup

Data set described in Section 2 which is the large-scale corpus of relevant comments for each network denoted by $\mathcal{D}_{cnn}$, $\mathcal{D}_{fox}$, and $\mathcal{D}_{msnbc}$ is further filtered using Regular Expressions to capture viewer's comments which are target points for *life begins..* discussion. Table 8 represents the skeleton of the staged data before passing to the cdQA pipeline.

Table 8: Staged Data for cdQA pipeline

News network	Video Title	Set of Comments
CNN	Video Title$_i$	Viewer Comments
Fox News	Video Title$_i$	Viewer Comments
MSNBC	Video Title$_i$	Viewer Comments

6.3.2 Predicting Top Answers from cdQA pipeline

Each entry in the dataset described in Table 8 becomes a separate corpus entity on its own where for a news network, a specific video level data with all its viewer's comments which are target points for *life begins* discussion is passed to the cdQA pipeline with a query Q. The cdQA pipeline outputs the top n answers *(n=2)* with its associated metadata including the viewer comment, video title, and News Network. Query Q *When does life begin?* was used against every video-level corpus. Table 9 represents the skeleton of the output data results from the cdQA pipeline.

6.3.3 Result Aggregation and Analysis

Top n answers *(n=2)* were aggregated at a news network level to analyze the popular public opinion to the Q *When does life begin?*. Figure 9 represents the distribution of most popular public opinion for the News Channel FOX News, MSNBC and CNN respectively. It is evident from the distribution of answers that across every US Cable News Network *Conception* and *at Conception* is the most popular answer to the question *When does life begin?* Also, there is a relatively higher proportion of pro-choice answers like *first breath* amongst the viewers of MSNBC compared to FOX and CNN news channels.

7 Analysing shifts in Public Opinion on Overturning of Roe vs. Wade Decision

On June 24, 2022, the US Supreme Court overturned Roe v. Wade. Using the similar experimental setup described in Sub Section 6.3, video Level Corpus was divided per US cable news network into two temporal sets with one having videos posted before the day Roe vs. Wade overturn and another after Roe vs. Wade overturned. Temporally divided video corpus was passed to the cdQA pipeline with a query Q. The cdQA pipeline outputs the top n answers *(n=2)* with its associated metadata including the viewer comment, video title, and News Network. Query Q used for this experiment was *When does life begin?*. Top n answers *(n=2)* were aggregated at a news network level to analyze shifts in public perception for the question *When does life begin?* before and after the overturning of the Roe vs Wade Decision. Figure 10 represents the distribution of most popular public opinion amongst the viewers of News Channels like FOX News, MSNBC, and CNN respectively for both these timelines. From Figure 10, which is the distribution of top answers, one can observe that the overturning of the Roe v Wade decision didn't have a significant impact on deviating viewers' opinion across every News Network.

8 Analysing Viewers Stance using Natural Language Inference

8.1 Natural Language inference

Given a premise $\mathcal{P}$ and a hypothesis $\mathcal{H}$, the natural language inference (NLI) task, also known as Recognizing Textual Entailment (RTE), involves predicting entailment, contradiction, or semantic irrelevance (i.e., neither entailing nor contradicting) [MacCartney and Manning2008]. Textual entailment is much more relaxed than pure logical entailment and can be viewed as a human reading $\mathcal{P}$ would infer most likely $\mathcal{H}$ is true. For instance, the hypothesis some men are playing a sport is entailed by the premise of a soccer game with multiple males playing [Bowman *et al.*2015]. Recently textual entailment has found application in mining stance on politically significant issues [Dutta *et al.*2022].

8.2 Aggregating Viewer's stance on `Life begins at Conception`

From the QA experiment, it was evident that *Life begins at Conception* is the most popular perception talked about amongst viewers of every news network. However, advanced NLP models like Question Answering seldom fail to capture the real intent of the viewer's opinion. On deeply analyzing the viewers' comments which resulted in the answer *Conception* to the question *When does life begin?*, it could be observed that even though viewers say *Life begins at Conception*, their real intent may vary across several dimensions. On manually analyzing the viewer's comments across every News network, viewers' intent may vary when they talk about *Life begins at Conception*. The figure 11 below highlights that for every news network even if viewers say *Life begins at Conception*, their stance could be either in the direction of *support, contradiction, or neutral*.

To study the distribution of the viewers' intent on *Life begins at Conception* across US cable News networks, NLI-based Text entailment techniques were leveraged. The stance of every Viewer's comment treated as Premise ($\mathcal{P}$) which answered *Conception* to the question *When does life begin* was determined using the Text entailment technique against the Hypothesis ($\mathcal{H}$) *Life begins at Conception*. Pre-trained RoBERTa [Liu *et al.*2019] model which was further fine-tuned on the Natural language Inference task was used to predict the stance for this experiment The predicted stance was aggregated to the News Network level to analyze the distribution. Figure 12 below represents the stance distribution across every news network. From the Distribution of the stance of viewers', it can be seen that the viewers of MSNBC have a relatively more contradictory stance on *Life begins at Conception* when compared to the viewers of FOX and CNN.

8.3 Aggregating Viewer's stance on `Abortion is a Constitutional Right`

It was observed from bi-gram and topic modeling analyses that `constitutional-rights` bi-gram is heavily paired in the discourse of *abortion*. To study the distribution of the viewers' stance on *Abortion is a Constitutional Right* across US cable News networks, the NLI-based Text entailment technique is used again. Pre-trained RoBERTa [Liu

Table 9: Result Data from cdQA pipeline

Query	Rank	Answer	Viewer Comment	Video Title	News network
When does life begin	1	Answer$_i$	Viewer Comment$_i$	Video Title$_i$	CNN
When does life begin	1	Answer$_j$	Viewer Comment$_j$	Video Title$_j$	Fox News
When does life begin	2	Answer$_k$	Viewer Comment$_k$	Video Title$_k$	MSNBC

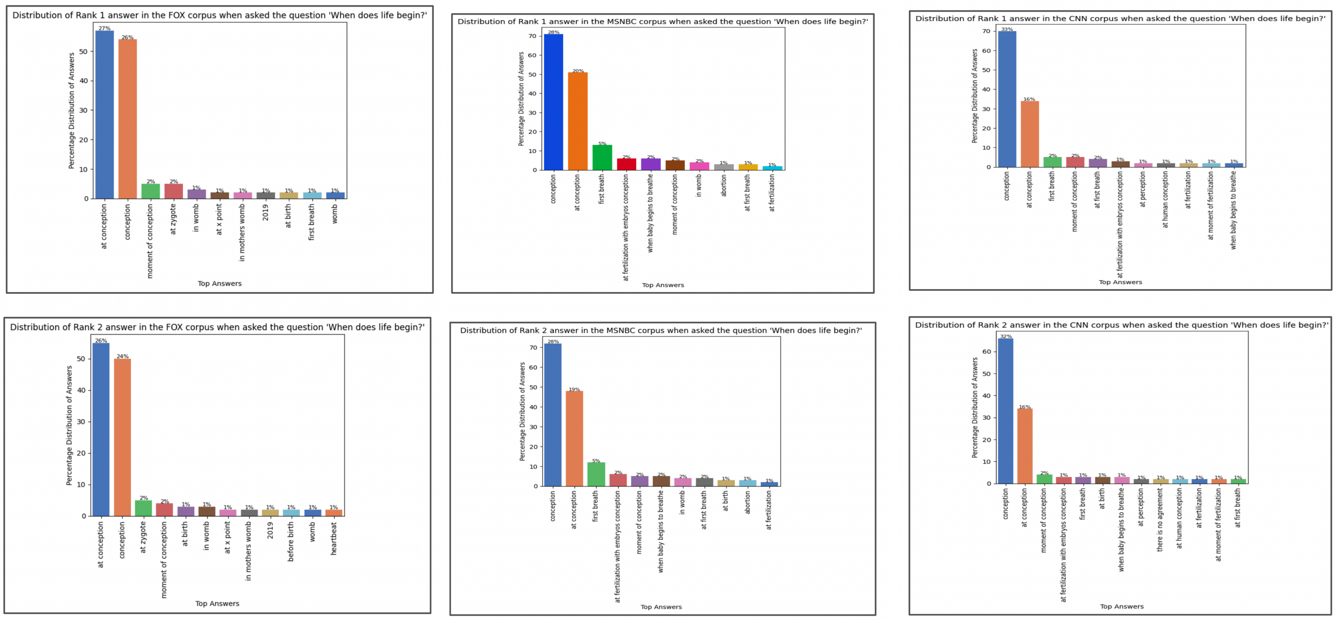


Figure 9: Distribution of Top 2 Answers to question *When does life begin* in (left) Fox News (middle) MSNBC and (right) CNN

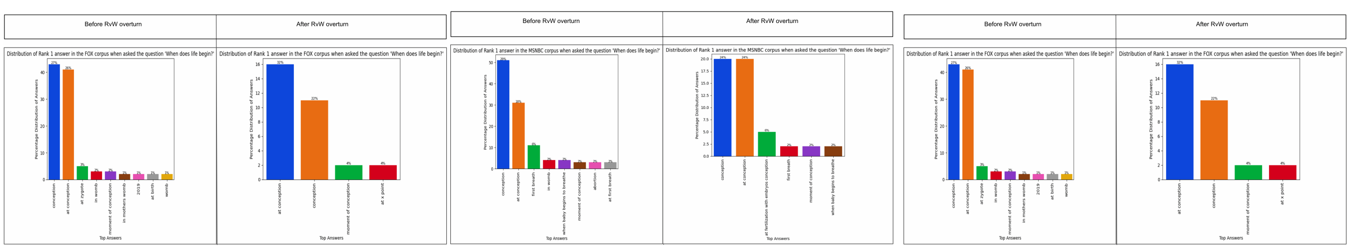


Figure 10: Distribution of Top 2 Answers to question *When does life begin* in (left) Fox News (middle) MSNBC and (right) CNN before and after the Overturn of Roe vs. Wade Decision

*et al.*2019] model fine-tuned on the Natural language Inference task was used to aggregate viewers' stances across every US cable news network. A corpus was created per US cable News network where user comments treated as Premise ($\mathcal{P}$) were filtered(using Regex) around *abortion rights../ constitutional rights../ woman rights..* based abortion discussion. The stance of every Viewer's comment was determined using the Text entailment technique against the Hypothesis ($\mathcal{H}$) *Abortion is a Constitutional Right* and was aggregated to analyze its distribution at every US cable News network. Figure 13 represents the stance distribution of viewers across every news network. From the distribution of the stance of viewers', it can be seen that the viewers of FOX have a relatively more contradictory stance on *Abortion is a Constitutional Right* compared to the viewers of MSNBC and CNN.

9 Future Work

In our analysis, we showed that there exist quantifiable differences in the discussion regarding Roe v. Wade across the YouTube comment section of different news channels. Our experiments revealed the aggregate audience stance on the questions *"When does life begin?"* and *"Is abortion a constitutional right?"*. Although we designed multiple experiments utilizing advanced NLP techniques like QnA system and transformer-based topic modeling there are a few routes that we would like to explore next. First, we would like to collect other data sources which could be prospective in capturing Public discussion on abortion such as subreddits like

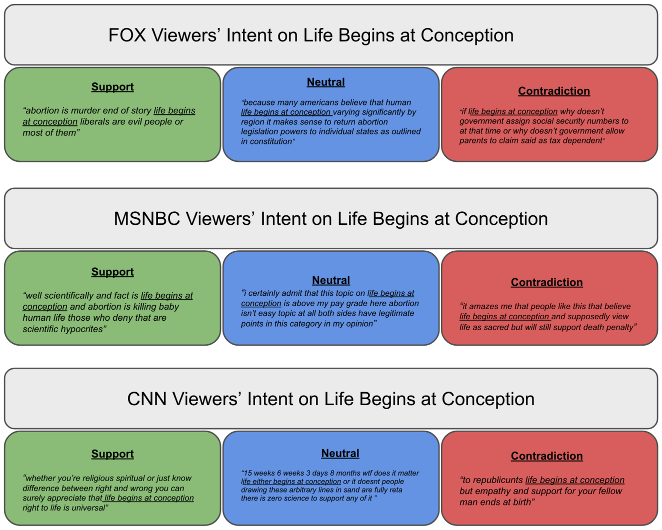


Figure 11: Viewers' Intent on *Life begins at Conception*

r/abortion and r/abortiondebate. We also showed that there exists a significant number of users who self-declared their stance on abortion (i.e. pro-life, pro-choice). However, we did not get a chance to analyze these users' comments from the entire YouTube dataset. We think this could be an interesting research direction. Thirdly, in this work, off-the-shelf models were used that might not be ideal for all the experiments. Hence, we want to fine-tune the QnA system and the Text entailment model in the future.

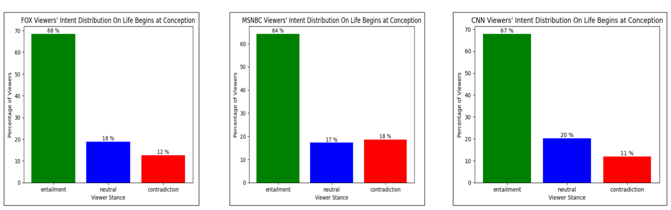


Figure 12: Viewers' Stance Distribution on *Life begins at Conception*

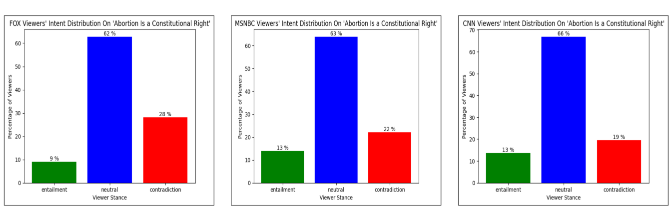


Figure 13: Viewers' Stance Distribution on *Abortion is a Constitutional Right*

References

[Arney and Trescher, 1976] William Ray Arney and William H Trescher. Trends in attitudes toward abortion, 1972-1975. *Family Planning Perspectives*, pages 117–124, 1976.

[Bowman *et al.*, 2015] Samuel R Bowman, Gabor Angeli, Christopher Potts, and Christopher D Manning. A large annotated corpus for learning natural language inference. *arXiv preprint arXiv:1508.05326*, 2015.

[Devlin *et al.*, 2018] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

[Dutta *et al.*, 2022] Sujan Dutta, Beibei Li, Daniel S Nagin, and Ashiqur R KhudaBukhsh. A murder and protests, the capitol riot, and the chauvin trial: Estimating disparate news media stance. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*, pages 5059–5065, 2022.

[for the Study of Ethical Problems in Medicine *et al.*, 1981] United States. President's Commission for the Study of Ethical Problems in Medicine, Biomedical, and Behavioral Research. *Defining Death: A Report on the Medical, Legal and Ethical Issues in the Determination of Death*. The Commission, 1981.

[Félix MIKAELIAN,] André FARIAS Félix MIKAELIAN. Closed domain question answering. https://github.com/cdqa-suite/cdQA.

[KFF,] KFF. Poll: Two-thirds of americans don't want the supreme court to overturn roe v. wade. www.kff.org/health-reform/press-release/.

[Liu *et al.*, 2019] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*, 2019.

[MacCartney and Manning, 2008] Bill MacCartney and Christopher D Manning. Modeling semantic containment and exclusion in natural language inference. In *Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008)*, pages 521–528, 2008.

[McInnes *et al.*, 2017] Leland McInnes, John Healy, and Steve Astels. hdbscan: Hierarchical density based clustering. *The Journal of Open Source Software*, 2(11):205, 2017.

[McInnes *et al.*, 2018] Leland McInnes, John Healy, Nathaniel Saul, and Lukas Grossberger. Umap: Uniform manifold approximation and projection. *The Journal of Open Source Software*, 3(29):861, 2018.

[Reimers and Gurevych, 2019] Nils Reimers and Iryna Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 11 2019.