Analyzing Sentiments on Medical Negligence Amongst Minority Communities on Reddit Using Computational Social Science

Angel Westbrook

University of Chicago, Division of the Social Sciences

Preceptor: Dr. Damien Bright

Advisor: Dr. Shelly Robinson

June 25th, 2025

Introduction

In everyday conversation and online spaces, different minority groups share similar stories of how their illnesses are not taken seriously, they're misdiagnosed, or are refused treatment for their illnesses. Thus, the negative symptoms that they may be experiencing would, for example, be attributed to mundane things like anxiety or stress by medical practitioners who dismiss other underlying and serious medical issues the patient(s) may be having because of their implicit biases. Because the medical system often neglects certain populations from the focus of medical care and medical education, it discourages and actively opposes a more holistic and nuanced view of medical care related to different populations. Because of this, it forces these minority groups to advocate for themselves by seeing multiple different doctors, self-diagnosing, or resorting to online spaces to voice their grievances about their experiences with seeking treatment, while their symptoms may evolve into something life-threatening.

Concerning online spaces, there are many social forums, videos, posts, and news about minority experiences when it comes to seeking treatment, along with independent 'research' as a means to air their grievances and organize their bottom-up experience through these online spaces. What I've done for my project is examine social forums, specifically Reddit, to analyze and measure the prevalence of how people from minority groups feel about medical bias when seeking medical care. Thus, I lead with the question: **How can we measure sentiments among minority groups regarding medical neglect in online discussion forums, and what sentiments are most prevalent within these communities?**

Literature Review

Key words and definitions: medical bias/neglect (a blanket term I will use to highlight negligence of doctors amongst minority communities). Web-scrapping (A way to collect posts from Reddit), VADER, RoBERTa, Coherence Score, and Topic Modeling (explained in the methods section).

There are many key dimensions within the medical world that a smorgasbord of researchers have taken upon themselves to research and highlight to explain issues related to medical negligence/bias. One dimension focuses the issue on the chronic illnesses that affect women and argues that these illnesses are often misunderstood, misdiagnosed, and mistreated due to their disproportionate impact on women, as one researcher, Dr. Eleanor Krassen, argues in her work. With there being a lack of research funding to help understand these conditions, it creates a lack of understanding within the diagnostic process to properly treat women. *(Krassen, 2022).* Researchers like Krassen also highlight the structural and social barriers that women face in being treated properly, which deepens the issue of medical neglect within the diagnostic process. This issue becomes even more nuanced when the legality of this is added. By examining the legal aspect of diagnostic discrimination, we can see gender bias in medical malpractice lawsuits with reduced financial settlements for women compared to men, which highlights the issue of women often not receiving justice when facing medical issues as a result of diagnostic discrimination. *(Plaza, 2018).*

Another researcher, Chesley Carter, gives a specific case of diagnostic discrimination in her article as she introduces the concept of "health omissions," arguing that systemic anti-Black racism leads to the erasure of Black patients and their suffering from medical knowledge and care which coincides with Krassen's point on the lack of education for women and both comment on the exclusion of these demographics from medical education which leads to diagnostic discrimination. She uses a case study of "Uncle Bobbie," aka Bobbie Moore, a black man from St. Louis whose ALS diagnosis was denied for years and who died as a result. While his family, who were his primary caretakers, insisted to medical practitioners that he may have ALS through their online research, which informed them to make this self-diagnosis, medical practitioners kept denying that this was the case. After experiencing a major health episode, he had to be rushed to Midway Hospital, where he was finally diagnosed, but died as a result of the late-stage effects of ALS. *(Carter, 2021)*.

Medical tourism is another dimension of medical neglect as well, the practice of traveling internationally to seek sufficient medical care when the individual's country's healthcare system has inadequate resources, high costs, or discriminatory practices that prevent them from obtaining the treatment that is needed for the individual. In a case study conducted by Dr. Aren Z. Aizura, she highlights medical tourism by examining the lives of two transgender women who are required to go to Thailand to get gender reassignment surgery (GRS) because Thailand doesn't have as strict regulations for GRS, compared to many Western countries which require psychiatric evaluations and real-life experience (RLE) criteria to be considered for this procedure. Additionally, they note how Thai medical staff are much more respectful and friendlier about their gender identity and don't face discrimination due to their identity as they do in Western countries when seeking treatment from medical practitioners. *(Aizura, 2010)*.

With these issues related to different minority groups, when it comes to things such as race, sex, and gender identity as examples, it raises the question of how medical educators teach future practitioners how to treat. In one study conducted by Dr. June Futterman; it examines how medical educators at Georgetown University Medical Center understand and use race in clinical decision-making. The study also identifies a growing awareness of the impact of structural racism on health outcomes, suggesting that medical education needs to address the biological and cultural aspects of race as well as the sociopolitical context. *(Futterman, 2024)*.

These differing issues within the medical world highlight many key elements related to medical negligence that exist within the medical field and how these researchers define these elements through their research. With many different patients receiving inadequate care, it raises an inquiry of how many other people have had their symptoms worsen or have potentially died as a result of medical practitioners not taking their symptoms seriously.

Examining how health literature and practices teach people about people from different demographic backgrounds could also address medical myths within the field, such as the belief that black patients feel less pain than white patients *(Hoffman, 2016)*. This belief isn't unique to the medical field either, and has become such a popular belief that even leading health officials within the US hold these ignorant beliefs while making policy decisions related to health within the US. *(Reed, 2025)*. Harmful beliefs like these cause mistrust amongst patients and influence them to not seek treatment when they have a serious illness, which could be directly linked to voicing their sentiments within social forums like Reddit. *(Murphy, 2024)*.

From this, when examining these social forums, utilizing a computational-based lens will be useful for capturing the prevalence of these sentiments. Within the realm of computational social science, which is an interdisciplinary subfield that uses computational methods in research. Measuring sentiments is something that has been applied in a variety of ways and utilized by various researchers within this emerging subfield. When it comes to sentiment in everyday language, one source states that we can split sentiment into two meanings: (a) a feeling, or something of emotional significance, and (b) a particular (usually subjective) point of view. *(Puschmann and Powell. 2018)*. Within the subfield of computational social science, we collect sentiment, sometimes called opinion mining, to capture the subjective viewpoints on a variety of topics related to things like history, politics, events, and more. As one source states, sentiment analysis techniques can be split into four categories: machine learning, deep learning, lexicon-based analysis, and hybrid approaches that mix the previous techniques. *(Mao et. al. 2024)*. For an applied example, in one market study, sentiment analysis is used to analyze microblogs on Twitter to collect user-submitted reviews on two stores in the UK: Tesco and ASDA over the Christmas holidays in 2014, utilizing R. *(Younis, 2015)*. This resulted in a confirmation that this technique can be used by businesses to make it easier for them to capture customer sentiment about their companies, and could be replicated on other social media platforms such as Facebook, Reddit, TikTok, and more.

Shedding some more light on the issue of medical discrimination, utilizing computational methods creates innovative techniques to analyze how different demographics are treated when they have to receive care from medical practitioners, and what disparities may exist within this. For my scope, I will utilize these computational methods, specifically utilizing sentiment analysis through the machine learning method, which classifies the data I've collected into measurable categories, which I will then analyze qualitatively to capture and measure the sentiments of minority groups from different subreddits within Reddit to answer my question of how can we measure sentiments among minority groups regarding medical neglect in online discussion forums, and what sentiments are most prevalent within these communities.

To do this, I've created a script within Python that can give me the prevalence of sentiment, and have then analyzed the sentiment in order to see how the data correlates and what inferences can be made when studying the data. The main deliverables of this project will be the data visualizations, my analysis of the data, my presentation, and the Jupyter script that I've created.

By addressing this issue, we can see a layout of what potential policies, practices, and ethics need to be considered to inform researchers, advocates, NGO's, that can utilize this knowledge to understand sentiments about medical negligence and its adverse effects on minority communities and challenge it in a more informed way.

Methods

The main goal of this project is to create a script that can A). Web scrape from Reddit with specific communities and keywords to collect 8,000+ posts; B). Categorize the posts from web scraping into “neutral”, “negative”, or “positive” categories. And C). Visualize the data and its measurements with graphs in order to perform an analysis of the data to derive what can be determined from it.

To examine sentiments about medical neglect on Reddit, I will utilize Python to create a machine learning script that will autonomously analyze and measure the sentiment of 8,708 Reddit posts from medical subreddits. To do this, I have coded a list of subreddits to loop through, which tells the script to search through the list of subreddits. Additionally, I’ve created a separate keywords list to find specific information about medical discrimination within each Reddit post from these specific communities. For the subreddits, web scraped from this list: r/medcine, r/health, r/disability, r/nursing, r/womenshealth, r/askdocs, r/chronicillness, and r/disability. And for the keywords, I will be using: discrimination, bias, racist, racial bias, oppression, prejudice, medicine, advocating, unequal, neglect, transphobia, homophobia, and sexism. Since my focus is on minority sentiments and inequity within medical service, I’ve used these terms that I think would typically be used when talking about discrimination to find posts that talk about this. With this setup, I’ve looped through each Reddit community for their top 1000 posts with each keyword I’ve told Python to use. To accomplish this, I’ve used many library imports. I used *PRAW,* a Python library import (sort of like an app), to create a personal use script that could connect me to the Reddit API to do the scraping, basically creating a key for me to unlock the ability to scrape from Reddit. For the actual scraping, I used *time,* another import, to create a 0.2-second delay between post searches since Reddit tries to limit scraping if you try and scrape instantaneously.

After web scraping, I then used *VADER* to perform a basic sentiment analysis on whole posts that were found in the communities I've listed and keywords I've input. After collection, I've put the posts into the categories: neutral, negative, and positive. As a result, it gives me a list of posts and *VADER's* scoring system that uses a sentiment scale with a sentiment score ranging from **-1 to 1** analyze the polarity of each post, where scores close to 1 indicate strong positive sentiment, scores close to -1 indicate strong negative sentiment, and scores close to 0 indicate neutral sentiment. I then took this list and imported *seaborn*, *matplotlib*, and *pandas* to create a simple bar graph that shows the quantity of posts for each category. After performing this basic analysis for sentiment, I then compared the sentiment scores with RoBERTa and Go Emotions to get a more robust categorization for the collection of the data for analysis. With the RoBERTa and Go Emotions imports, instead of just having the categories: neutral, negative, and positive. RoBERTa places posts collected into 28 emotion categories (29 including neutral). With its full list of emotions being: 'admiration', 'amusement', 'anger', 'annoyance', 'approval', 'caring', 'confusion', 'curiosity', 'desire', 'disappointment', 'disapproval', 'disgust', 'embarrassment', 'excitement', 'fear', 'gratitude', 'grief', 'joy', 'love', 'nervousness', 'optimism', 'pride', 'realization', 'relief', 'remorse', 'sadness', 'surprise', and 'neutral', to apply a class label to each Reddit post collected and providing a confidence score with each post to place it into one of the listed categories.

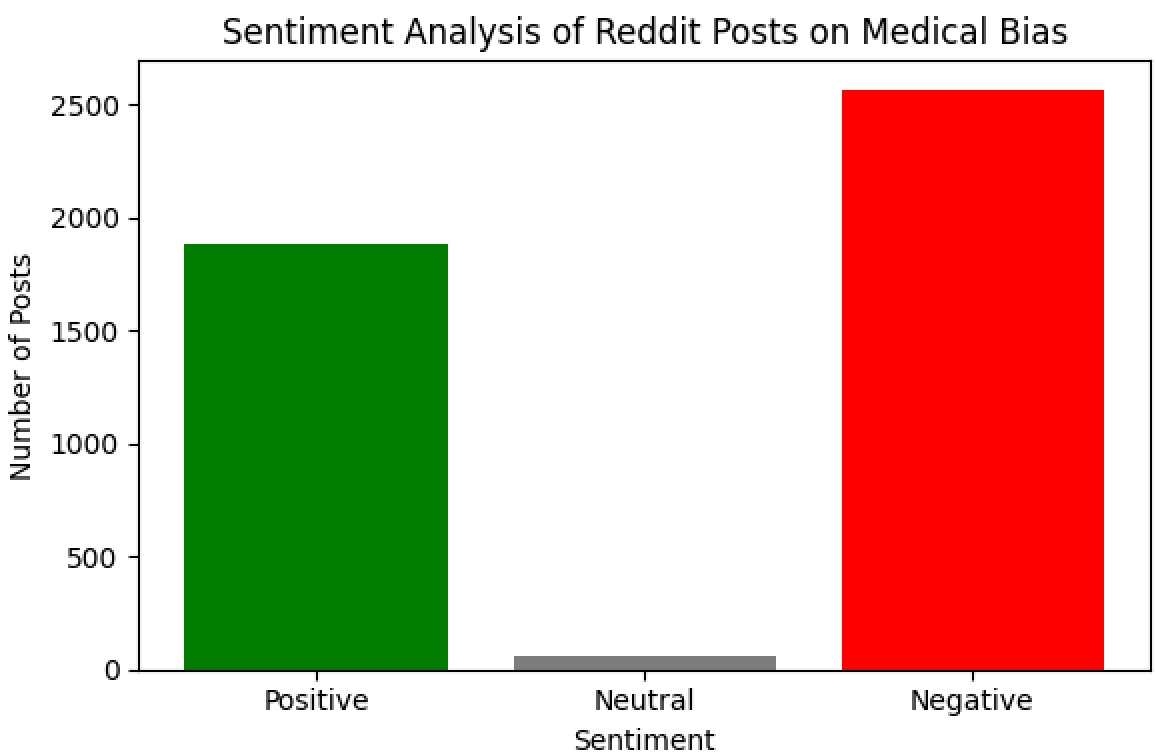
After collecting these posts, I will save them to a .csv file and clean the data to remove missing and duplicate values, while RoBERTa already pre-cleans data when running sentiment analysis, VADER has a harder time doing this; therefore, basic cleaning has been done to improve data efficiency. After the data is cleaned, I will use language data analysis to discover recurring themes within Reddit posts about medical discrimination to see if it matches the search terms I've input and help with my analysis to find major themes within the data, especially if it can find new themes/topics I didn't account for within these Reddit posts, and see the prevalence of topics the LDA model finds for analysis.

To do this, I have given the script the number of topics I want it to analyze; these topics are not predefined, and in the context of this study, are the number of word groups that are produced. I have then interpreted these word groups to see what I think the topic would be for each group. However, if I give the script too many topics, it could be accounting for too many outliers and noise, thus returning irrelevant word groups. Therefore, I've measured the number of optimal topics with a coherence score measurement from 1 to 0, where 1 has a high coherence and 0 has a very low coherence. After producing the optimal number of topics, I also included a weight to each topic to see which topic would be the most prevalent for analysis, along with word clouds to examine the dominant words within each topic.

When looking at these word clouds, I need to draw upon my own knowledge, feelings, and inferences as to what the words in each topic group mean. As an example, if one topic has the words: "children", "grades", "teacher", I would infer that this group could more than likely be about how teachers grade the children they teach in K-12. Thus, making a qualitative analysis of the meaning of each topic, which I will address more in the discussion section of this study.

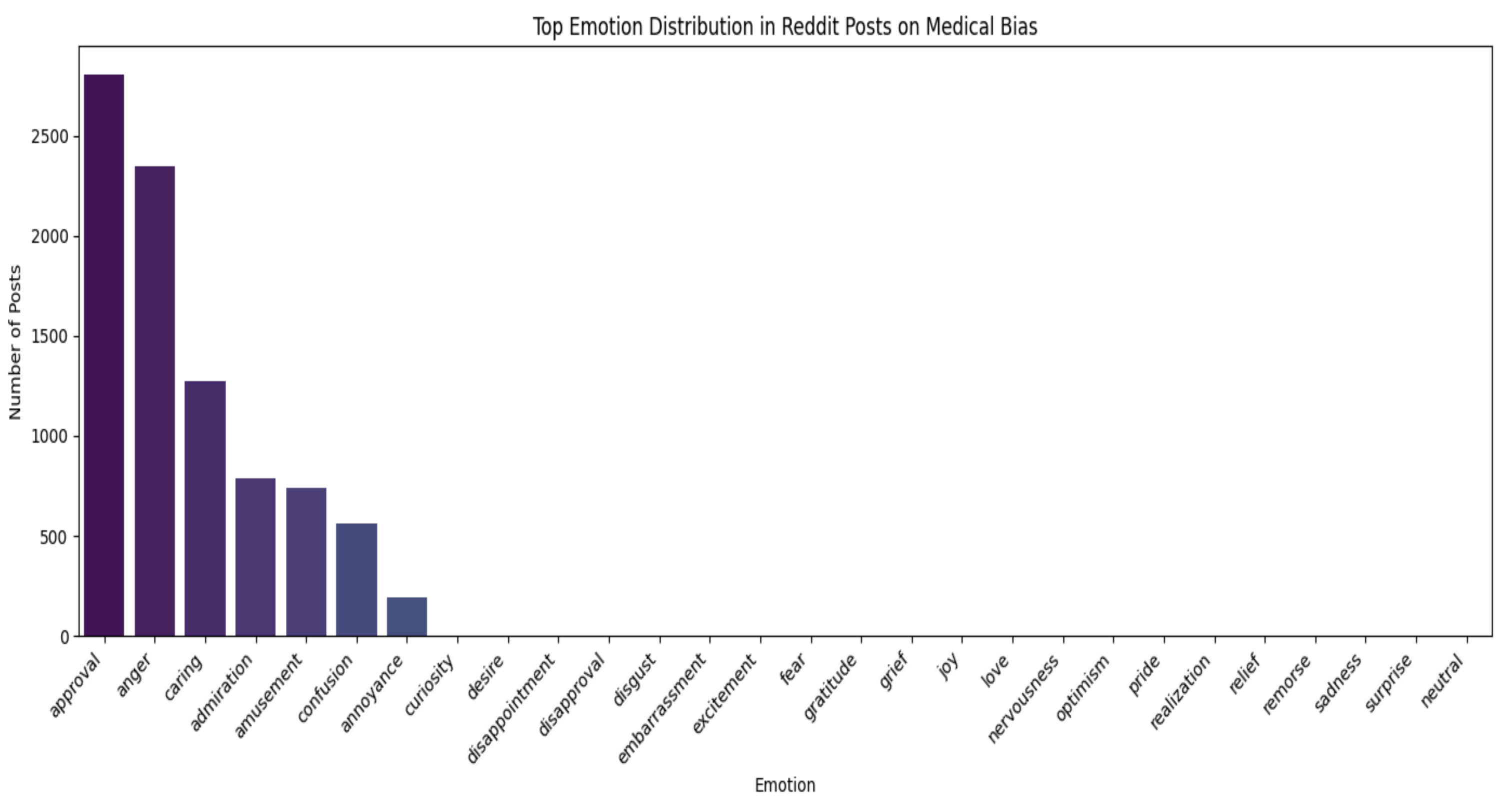
Results

For my results and analysis, I will go in order of my stated methods. After web scraping/cleaning, I ran VADER to analyze 8708 posts; it returned 6883 posts based on the criteria I've set for it, and produced this:



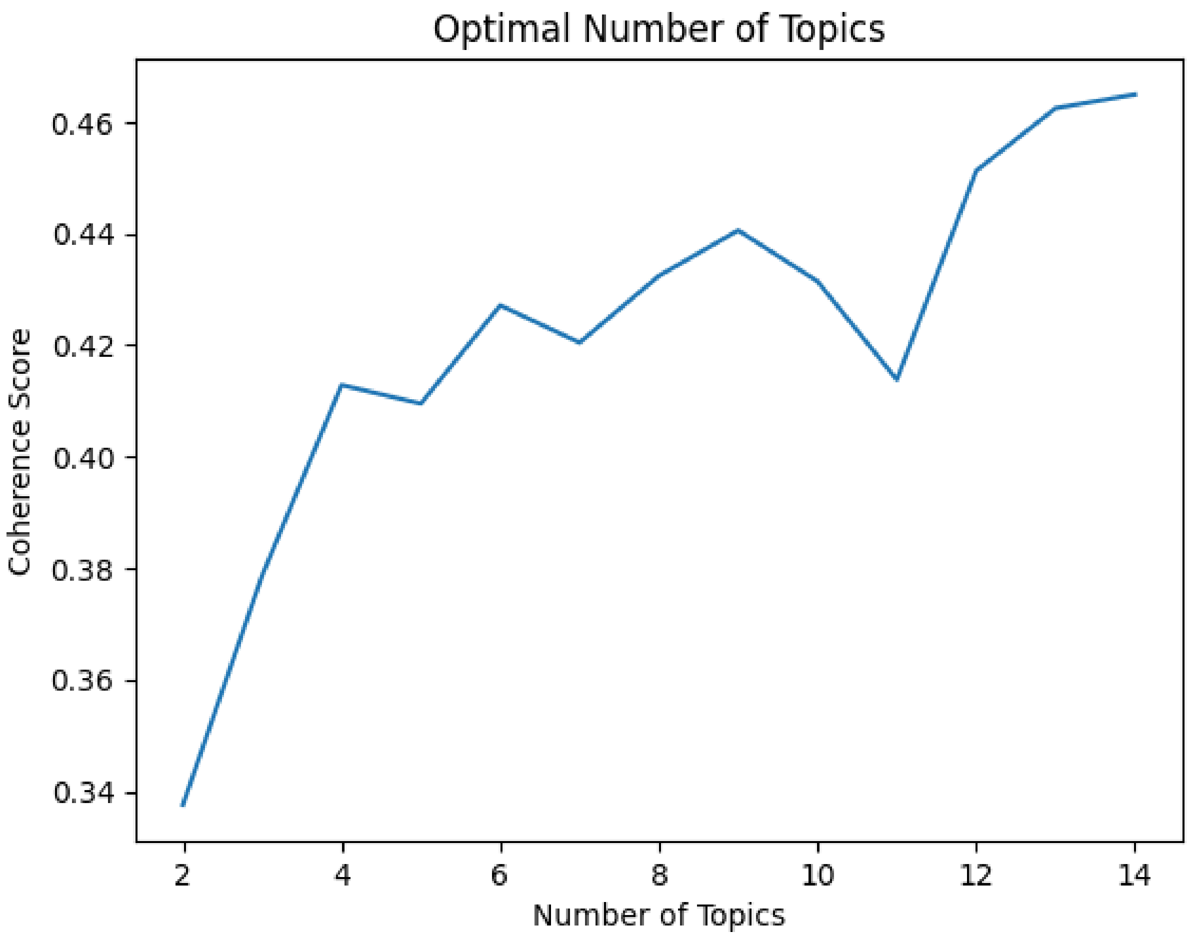
With this figure, when performing a sentiment analysis of the posts collected, we can see a strong negative sentiment when it comes to conversations around medical bias amongst or about minorities, which was the expected result given that this is a topic about an overall negative experience people have. However, there is a significant number of posts that also cover positive sentiments.

However, while this gives a good visualized representation of basic sentiment, I needed to delve deeper into this to have a more in-depth look at what kinds of emotions are most prevalent within the posts. Therefore, after I performed a second analysis with RoBERTa, I produced this table:



From this table, after analyzing 8708 posts, we can see that the dominant sentiment category is approval, with anger following as the second dominant sentiment. These posts that produce the approval category could exist because of potentially good policies, news, and practices that prevent medical bias from happening; however, this doesn't account for why the VADER analysis produces an overall negative sentiment, thus leaving the correlation unclear.

After producing these graphs, I then calculated the coherence, utilizing the elbow method to examine the line chart below:



The Elbow Method helps with the selection of the number of meaningful topics, aka word groups, that I should include that represent my data. Within this graph, it gives me an estimation of how many word groups I should have. When reading the graph, I am looking at the point that looks like an elbow between 10 and 12, and I am concluding that I need 11 word groups in order to properly represent my data. With these word groups, we can add a qualitative angle to this project, to see what word groups appear, and what words are most prevalent when medical negligence is spoken about on Reddit. *(Bobbit, 2023).*

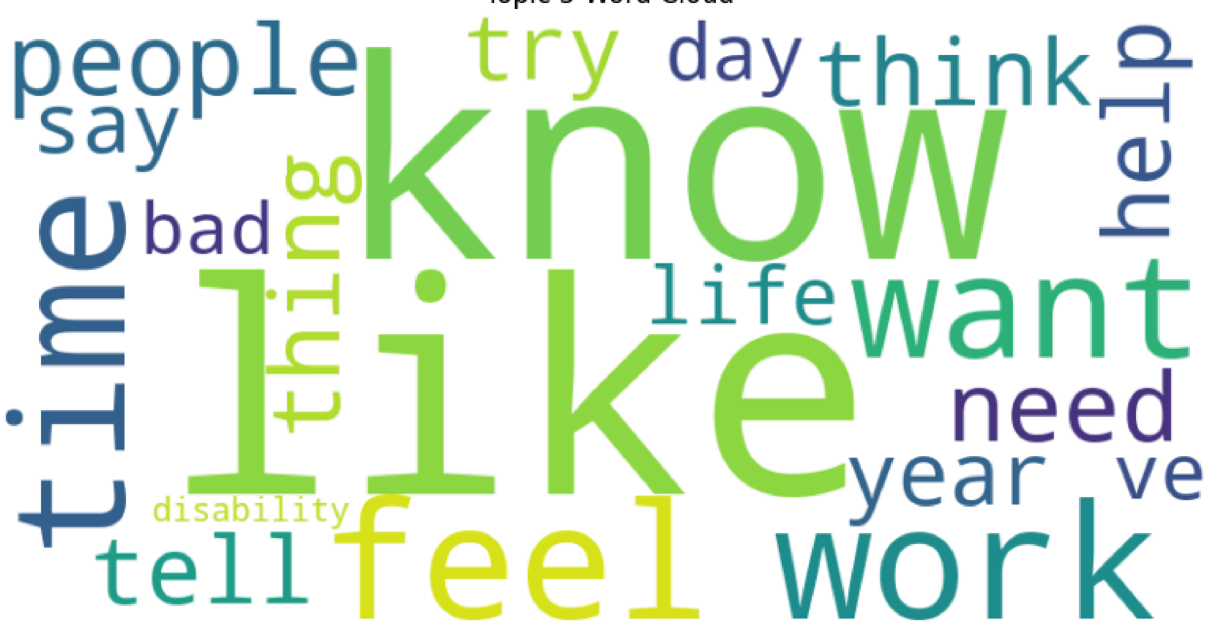
For each topic, these were the measurements in terms of ranking by prevalence that I then qualitatively analyzed:

=== Topics Ranked by Prevalence ===

1. Topic 5: Unkown (Weight: 1918.00) -> like | know | feel | work | time | want | people | need | thing | tell

From this list of words, it is hard to make a strong analysis since the most prevalent words are words such as ‘know’, ‘like’, feel, which don’t provide the strongest link to a theme when I think of qualitative associations. However, with words like ‘bad’, ‘work’, ‘need’, it could refer to the overall negative sentiment patients have when it comes to medical neglect. Omission of filler words might be the solution to this. However given my time constraints, that will not be performed in this study.

Topic 5 Word Cloud

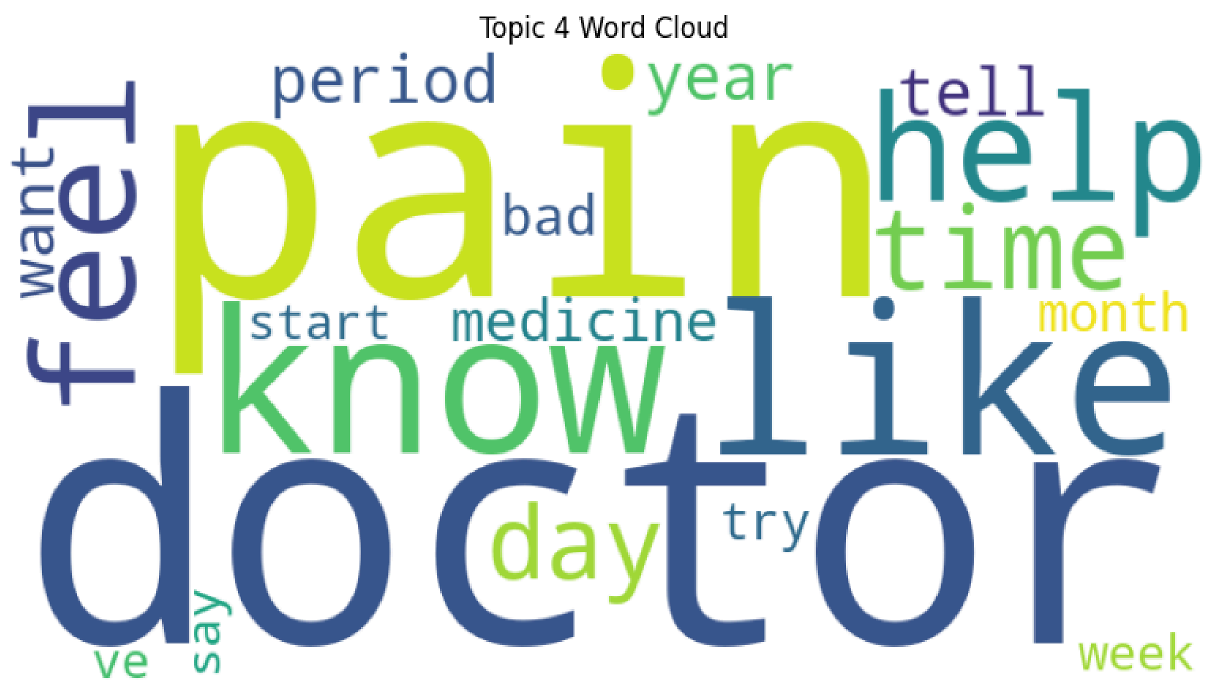


2. Topic 11: Seeking Treatment (Weight: 1001.43) -> pain | symptom | doctor | feel | year | test | like | time | issue | normal

This topic would be associated with people’s sentiments about symptoms/pain and seeking treatment from medical practitioners. With the top words being “pain”, “symptom”, and “doctor”, I would infer that this would be the main theme of this word group.

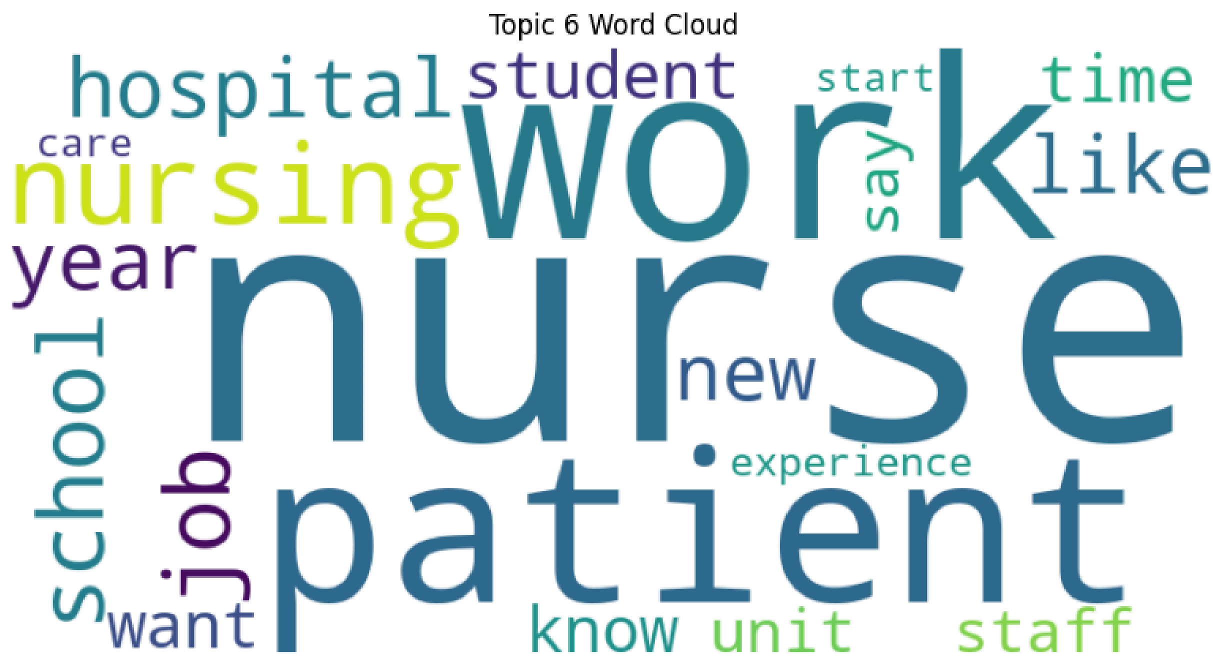
3. Topic 4: Pain (Weight: 959.76) -> doctor | pain | like | know | help | feel | time | day | period | tell

This topic could be about doctors and pain. Potentially, how doctors view patients' pain, how doctors handle pain, and what diagnoses they give to their patients.



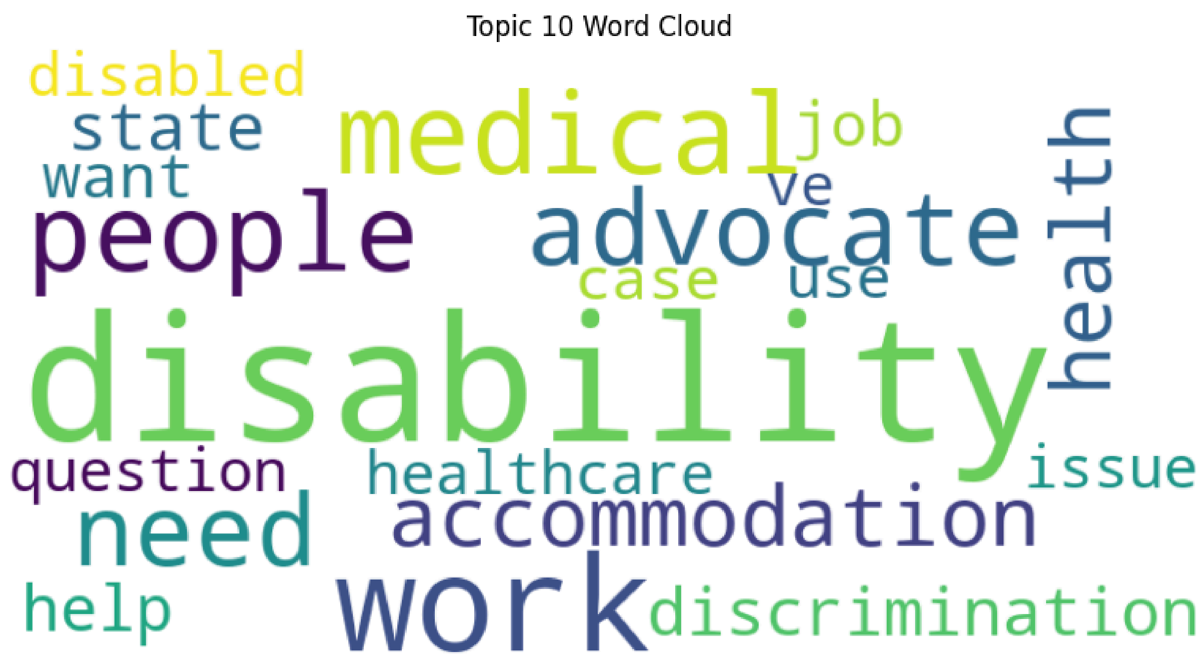
4. Topic 6: Nursing (Weight: 942.71) -> nurse | work | patient | nursing | job | school | hospital | year | like | new

It could be mainly about nursing specifically and their interactions with patients



5. Topic 10: Disabilities (Weight: 864.40) -> disability | work | medical | people | advocate | need | accommodation | health | discrimination | state

This topic would be about people with disabilities, how they would be discriminated against within the medical field, the accommodations they would need, and how they would want to be advocated for.



6. Topic 2: Hospital Care (Weight: 673.45) -> patient | hospital | say | tell | day | doctor | nurse | pain | care | come

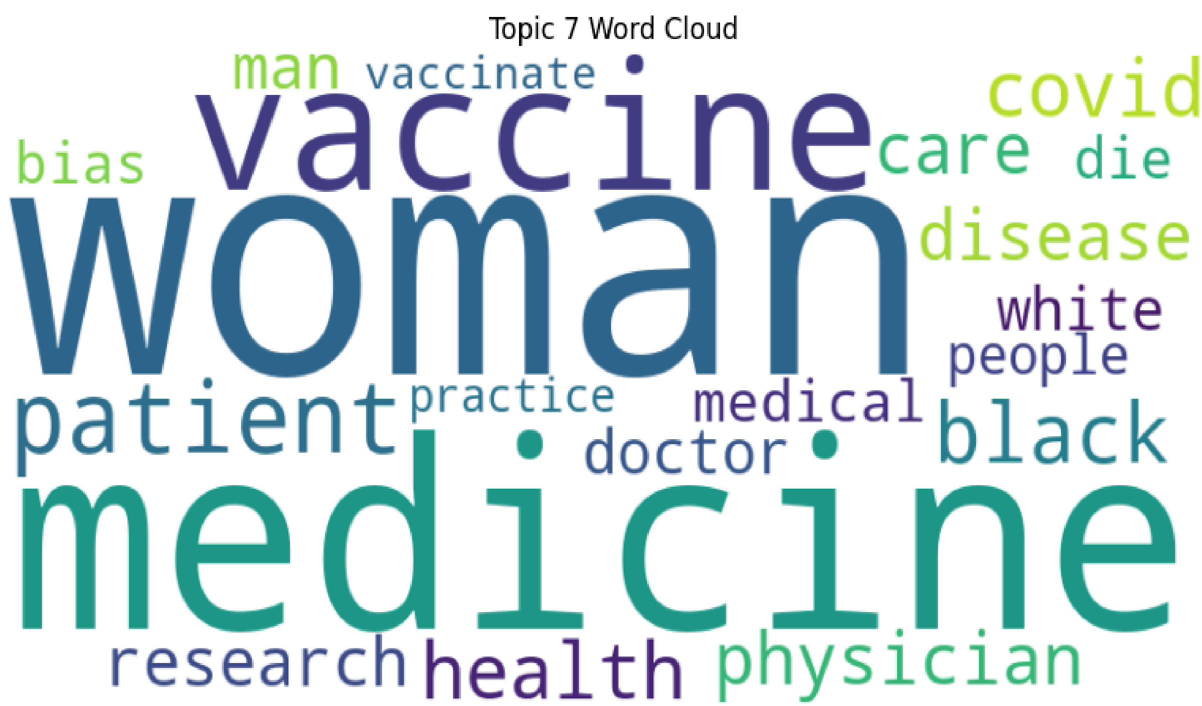
This topic would be mainly about hospital care. With the words 'patient', 'hospital', 'nurse', 'doctor', the major theme could be about patient experiences with hospital care and how that relates to medical negligence.

7. Topic 3: Disability (2) (Weight: 532.58) -> people | disabled | post | disability | like | comment | community | think | person | know

Based on the keywords, this could be another topic about the disabled community; however, I am unsure what specifically. Potentially, it could be general discussions of disabled patients' experiences with medical neglect, but what in particular is what I am unsure of.

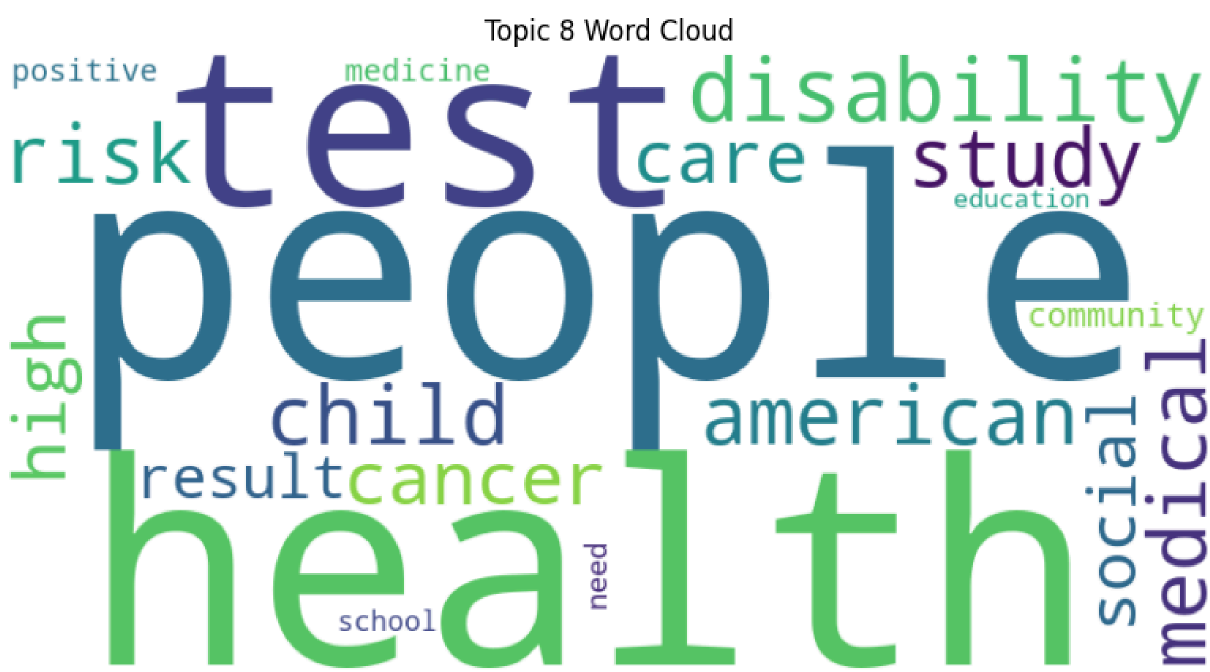
8. Topic 7: Black Women and Covid (Weight: 492.32) -> woman | medicine | vaccine | patient | black | covid | health | physician | care | disease

Based on the keywords, this topic could be about black women and COVID-19. Potentially noting inadequacies in care during COVID and how physicians treated black women during this time.



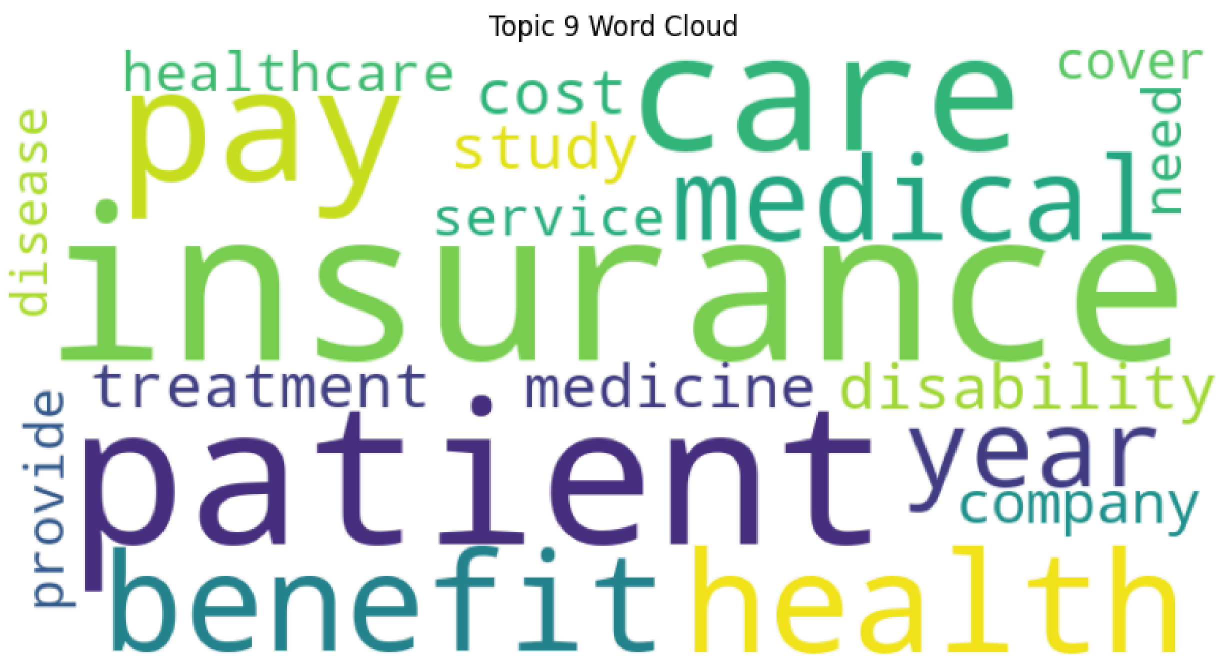
9. Topic 8: Childcare and Health in the US (Weight: 483.64) -> people | health | test | disability | child | american | study | medical | risk | cancer

Based on these keywords, this could be about childcare and health. A bit off topic for my research, but could be an important factor when studying family medicine and inequalities within it.



10. Topic 9: Insurance (Weight: 429.57) -> insurance | patient | care | pay | health | benefit | medical | year | disability | study

This would add another realm to this by including the financial aspects of care. Mainly focusing on insurance, medical fees, debt, and so on.



11. Topic 1: Time and Events (Weight: 410.14) -> like | feel | day | year | ve | time | eye | mg | start | eat

This could just be about different times that certain events may happen. Which would make sense as to why this would be the weakest topic out of the 11.

Topic 1 Word Cloud

Analysis

When comparing RoBERTa's and VADER's sentiment analysis, we can see that there is an overall negative sentiment in VADER's analysis; however, the emotional category "approval" is the highest in RoBERTa's analysis. This could be since VADER typically captures generally negative and positive sentiment and places them into non-specific categories. RoBERTa subdivides these negative emotions like anger, confusion, and annoyance into separate negative sentiments, which would make it seem like approval equals an overall positive sentiment. However, when added together, we would be able to see the overall negative sentiment that VADER shows.

Additionally, within the approval category, when examining the posts qualitatively, we get a range of different tones and contexts:

Title: Response to D.O. surgeon discrimination

Top Emotion: approval (0.89)

Title: Need advice for dealing with discrimination at hospital

Top Emotion: approval (0.90)

Title: How doctors can respond to discrimination from their patients

Top Emotion: approval (0.91)

Title: Doctors, how's age discrimination in your industry? Whether it be too old or too young.

Top Emotion: approval (0.97)

From this list, we can see that there isn't a strong positive sentiment regarding posts within the approval category, and the context of these posts could've been missed in the software's analysis. But most of them appear to be declarative statements and inquiries about discrimination rather than explicitly showing approval. This would mean while there may be posts that do have positive sentiments associated with the approval category, there are also a significant number of posts that are associated with negative or neutral sentiments that *acknowledge* medical discrimination but don't necessarily raise a strong positive sentiment, thus complicating and confusing the approval category and raising the need to conduct a more qualitative analysis to understand the context of each post.

For the topic groups, Topic 5 seems to have the most prevalence, with "like" being the most dominant word within this topic. However, I am unsure how this could correlate with my previous graphs mentioned above when it comes to minority sentiments about medical bias. However, with the following top 2 topics: Topic 11 and Topic 4, I've determined that these would be major topics about pain and seeking treatment. More than likely sentiments about patients who complain about their pain and negative sentiments about going to the doctor.

Limitations and Discussion

Within this project, one major limitation is my lack of analysis in examining sentiment over time. While I have collected the top 1000 posts from each Reddit group to have the data be more representative, the timeline of posts collected would be a significant limitation, especially when trying to analyze sentiment changes with things like new healthcare policies, health leaders, and organizations, medical research, and news events involving the medical industry as well. And with the models themselves, it could prove difficult for VADER and RoBERTa to analyze data over time since languages change over time; thus, it could be difficult to analyze nuances in emotion and different word structures, such as slang.

Despite cleaning and fine-tuning, the data could still contain variables left unaccounted for that could skew my dataset, thus affecting the results. To fix this, more data exploration/experimentation should be done to explore what possible variables should/shouldn't be accounted for to further optimize the script. Additionally, there could be other ways to measure sentiment that I may not be accounting for, which could potentially limit my analysis of the data.

Another limitation is the lack of qualitative data within my project. While prevalence and quantitative analysis help us capture the gravity of the issue of medical bias/discrimination, there is a lack of highlighting people's personal experiences when seeking treatment, especially with the word clouds. To fix this, omitting filler words such as 'like', 'as', 'and', etc, could yield more specified word groups that could create a clearer linkage to qualitative themes related to people's experiences with medical neglect.

The way VADER and RoBERTa process language could be a limitation as well. When examining posts, posts that show things like sarcasm could be misrepresented as positive/negative/neutral or be given an emotional category that doesn't truly represent what the posts are conveying. For example, within my dataset, there is this post:

Title: Another shining success of private equity taking over a county hospital

Top Emotion: admiration (0.92)

While this post may seem admirable when not read sarcastically. When adding a sarcastic tone, this would likely be a negative post about private equity firms hurting rural communities. Additionally, this adds another layer, because it isn't just an experience of an individual and medical neglect, but a response to medical neglect from an individual about institutions causing inequity. This evolves into a larger issue of not just how specific cases involving people linked directly to medical discrimination are, but also how their loved ones, friends, and communities are affected by this sort of residual trauma.

This could also be the case when trying to examine cultural differences and nuances that could be misrepresented as well, especially when it comes to slang terms and medical terms that RoBERTa and VADER would have trouble analyzing, along with posts that may seem positive, but could reflect an overall negative situation. While I have input keywords that I would think directly reflect minorities and their medical experiences, there could be an issue of underrepresentation, which could skew the data towards highlighting posts that are not relevant to my research question. Additionally, conversations around medicine could take place outside of the US, which would shift the focus of this study. For example:

Title: 2 More Japanese Medical Schools Admit Gender Discrimination

Top Emotion: approval (0.91)

Even though this post is talking about sexism within Japanese Medical Schools, my focus would not be on Japan, which modifies the data away from a US-based study. However, it can raise inquiries about sentiments felt globally.

Future Implications

While these limitations with the script itself raise the need for a secondary qualitative analysis of the data. By making my inferences and interpretations for categories and topic groups, I've highlighted more nuanced language around this issue that is too simplified when run through a script. However, while obtaining the quantitative data, which makes measuring prevalence easier, it highlighted the need for human intervention. And by utilizing my own thoughts, experiences, and knowledge as a social scientist to convey the narrative presented to me through these quantitative methods and put the pieces of the "puzzle" together, so to speak. I have raised important inquiries about how layered these responses can be, and for the future, it would be worthwhile to study how the trauma of medical discrimination affects communities who haven't directly dealt with it, but feel its effects through their loved ones and the institutions they interact with.

This script I have created additionally, is a good basis to examine different social scientific issues that could be used to examine other social issues/trends that could be efficient for policy analysis when looking for current sentiments about a certain topic. And with even more refinement, it could be used as a tool to collect sentiment over time. Additionally, other scripts could be created to scrape from other websites like Facebook, YouTube, Twitter, and more, which could expand the data and make it more robust by collecting from different sources to achieve a more nuanced and varied result, which could be useful for advancing social scientific research, and implementing problem-fixing policies to mend the issue of medical neglect/bias. Additionally, this script can easily be made more accessible to other computational social scientists through GitHub so that they can modify the script based on mine, which can act as a blueprint. This could encourage more collaboration on this project to flesh it out more.

Conclusion

Disparities in healthcare are a persistent and complex issue that demands urgent attention. Structural inequalities, socioeconomic factors, and healthcare system barriers contribute to inequities in access, treatment, and outcomes for racial and ethnic minorities. By doing this project, I have shed light on potential issues related to the medical field utilzing a inovative mode of analysis, and by examining the sentiments of underrepresented groups by examining the sentiment of minority groups on Reddit through web-scraping, language data analysis, and measuring prevelance, we can gain insight into how we should evolve healthcare to work towards building a healthier future for everyone.

3

Appendix

For my code, I used VS Code to code with Python. Within the .ipynb file (Jupyter file), it should contain all of the necessary code and comments to replicate and modify my data, along with the presentation I have created for my project.

Github Link: https://github.com/Awest51/Personal-Projects

Bibliography

Covan, Eleanor Krassen. "Chronic illness: Misunderstood, misdiagnosed, and mistreated among women." Health Care for Women International, vol. 43, no. 1–3, 9 Feb. 2022, pp. 1–4, https://doi.org/10.1080/07399332.2022.2028470.-

Plaza, Cecilia. "Miss Diagnosis: Gendered Injustice in Medical Malpractice Law." Health Care for Women International, vol. 39, no. 2, 2018, pp. 91-139. https://journals.library.columbia.edu/index.php/cjgl/article/view/5701/2780.- .

Aizura, Aren Z. "Feminine transformations: Gender reassignment surgical tourism in Thailand." Medical Anthropology, vol. 29, no. 4, 24 Nov. 2010, pp. 424–443, https://doi.org/10.1080/01459740.2010.501314..

Gravlee, Clarence C., and Elizabeth Sweet. "Race, ethnicity, and racism in Medical Anthropology, 1977–2002." Medical Anthropology Quarterly, vol. 22, no. 1, Mar. 2008, pp. 27–51, https://doi.org/10.1111/j.1548-1387.2008.00002.x..-

Carter, Chelsey R. "Health Disparities and health omissions: Pushing Medical Anthropology Forward in the United States." Journal for the Anthropology of North America, vol. 24, no. 2, Oct. 2021, pp. 65–73, https://doi.org/10.1002/nad.12153..-

Osborn, Chandra Y. "Health literacy." Home Healthcare Nurse: The Journal for the Home Care and Hospice Professional, vol. 26, no. 4, Nov. 2007, pp. 255–257, https://doi.org/10.1097/01.nhh.0000316707.45504.0c.-

Futterman, June, et al. "Medical educators' perceptions of race in clinical practice." BMC Medical Education, vol. 24, no. 1, 4 Mar. 2024, https://doi.org/10.1186/s12909-024-05232-5..-

Hoffman, Kelly M, et al. "Racial Bias in Pain Assessment and Treatment Recommendations, and False Beliefs about Biological Differences between Blacks and Whites." Proceedings of the National Academy of Sciences of the United States of America, U.S. National Library of Medicine, 19 Apr. 2016, pmc.ncbi.nlm.nih.gov/articles/PMC4843483/#abstract1.

Thomson, Jonny. "'terminalism' - Discrimination against the Dying - Is the Unseen Prejudice of Our Times." Big Think, 11 Nov. 2023, bigthink.com/thinking/terminalism-discrimination-against-dying/.

Holt, Brianna. "RFK Jr's Claim about Black People's Immune Systems Is 'Unscientific and Terrifying.'" The Guardian, Guardian News and Media, 5 Feb. 2025, www.theguardian.com/us-news/2025/feb/05/rfk-jr-hearing-black-people-immune-systems.

Caruana, Edward Joseph, et al. "Longitudinal Studies." Journal of Thoracic Disease, U.S. National Library of Medicine, Nov. 2015, pmc.ncbi.nlm.nih.gov/articles/PMC4669300/.

Mao, Yanying, et al. "Sentiment Analysis Methods, Applications, and Challenges: A Systematic Literature Review." Journal of King Saud University - Computer and Information Sciences, Elsevier, 30 Apr. 2024, www.sciencedirect.com/science/article/pii/S131915782400137X?via%3Dihub.

Puschmann, Cornelius, and Alison Powell. "Turning Words into Consumer Preferences: How Sentiment Analysis Is Framed in Research and the News Media - Cornelius Puschmann, Alison Powell, 2018." Sage Journals, 7 Sept. 2018, journals.sagepub.com/doi/10.1177/2056305118797724.

Younis, Eman M.G.. "Entiment Analysis and Text Mining for Social Media Microblogs Using Open Source Tools: An Empirical Study." Cloudfront, International Journal of Computer Applications , Feb. 2015, d1wqtxts1xzle7.cloudfront.net/34273592/Mkemal_one_2-libre.pdf?1406116235=&response-content-disposition=attachment%3B+filename%3DMohammedKemal.pdf&Expires=1702730913&Signature=AxRSMuR5GDAedKK-f0v9ckVgENd8o7aznfsFGW7Ezj8xbQ5dvFdEe-AwT3oCXJ1SpCgeha4CYwOSxZBt~BQoVg3AjInTy0NROmU3EDRufZqB684ImaxEfXn2PuGTB8hD8CpM5nEXPF0l6rwYIU2dZq8a6ZAewAsQnt3fTNqsDYf6hi2Aoy2S-jLVJmldQYr5hY1tq4hIO4gResYlpfS44Gb3ugJPFqiZbl8k~WE69tT6CHhGvdMljsWocXHqY2vWUr7OTZOBdZcNPG2rKqN8KD5ExOmaQtHfdR8ijawkK4YGYykS4kwrHUHMfIrHoyenIs4jJQsItbSnk5oi5SWE2Q__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA.

Bobbitt, Zach. "How to Use the Elbow Method in Python to Find Optimal Clusters." Statology, 3 Jan. 2023, www.statology.org/elbow-method-in-python/.