

# Subreddit Classification: r/MensLib vs r/MensRights

Benjamin Dornel – DSI 18 – Singapore

# Problem Statement

- **Classify posts from r/MensLib and r/MensRights based on their title and body text**
- Help new users post in the right subreddit (subreddit names can be highly misleading)
- Allow moderators to easily identify posts that might have been posted in the wrong subreddit.



# r/MensLib vs r/MensRights

- **r/MensLib:**
    - Pro-feminist
    - Inclusive and broad definition of masculinity/manhood
    - Focus on positivity (aims to build a healthier, kinder, and more inclusive masculinity)
  - **r/MensRights:**
    - Anti-feminist
    - Supports traditional and outdated gender roles
    - Connected to the alt-right and other highly misogynistic subreddits
-

# Data Cleaning

- Retrieved 1,000 posts from each Subreddit using Python Reddit API Wrapper (PRAW)
- Removed URL links, HTML tags, text-based emojis and other irrelevant characters.
- Changed text to lower-case
- Combined title and text together

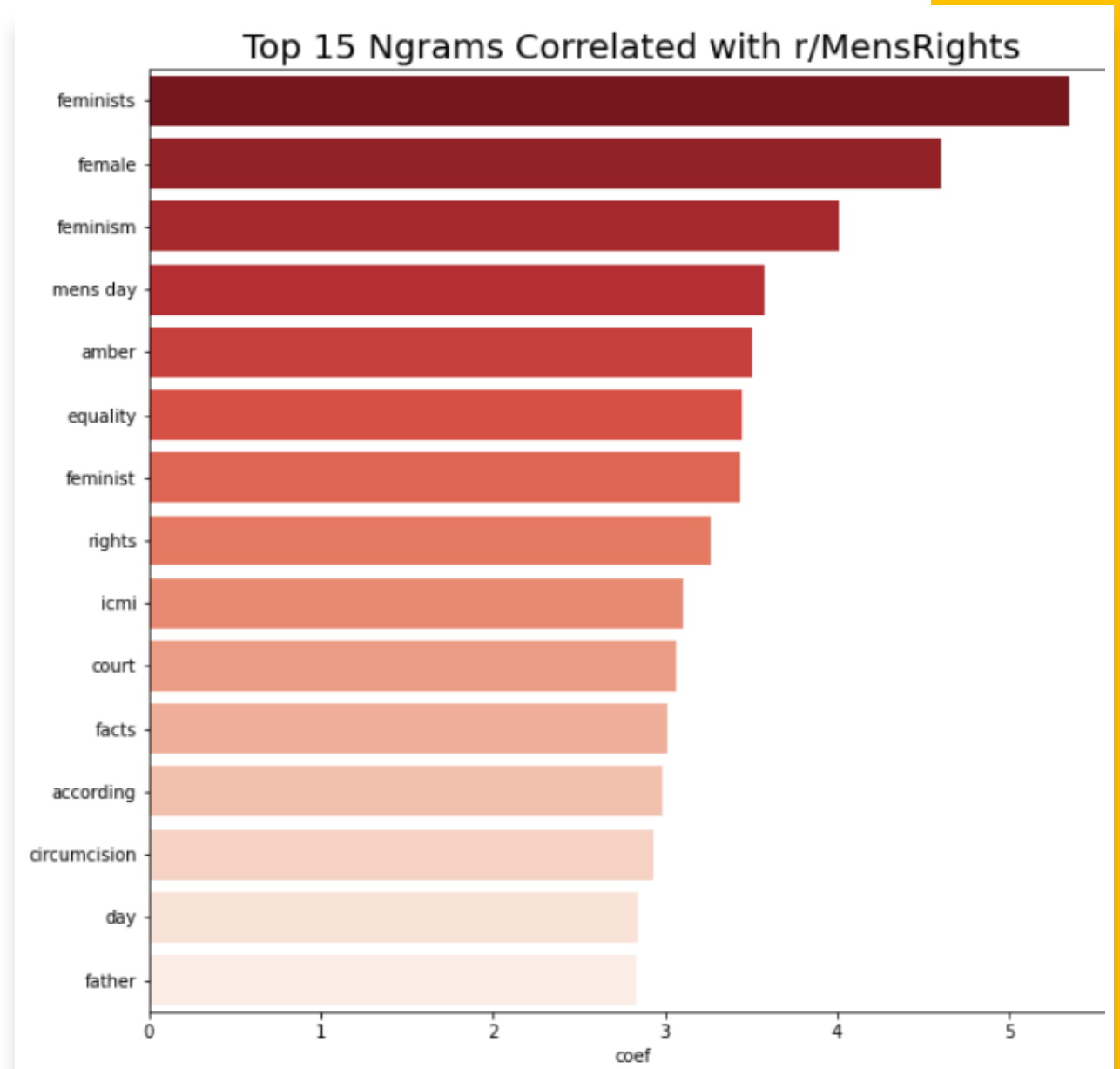
# Model Selection

- Logistic Regression with TfidfVectorizer:
  - Test accuracy: 0.834
  - F-score: 0.850
  - Includes only words or n-grams that appear in at least 4 posts
  - Excludes stop words and ignores terms that appear in more than 20% of posts
  - Uses Ridge regularization ( $\alpha = 0.1$  |  $C = 10$ )

# r/MensRights:

## Anti-Feminist Rants

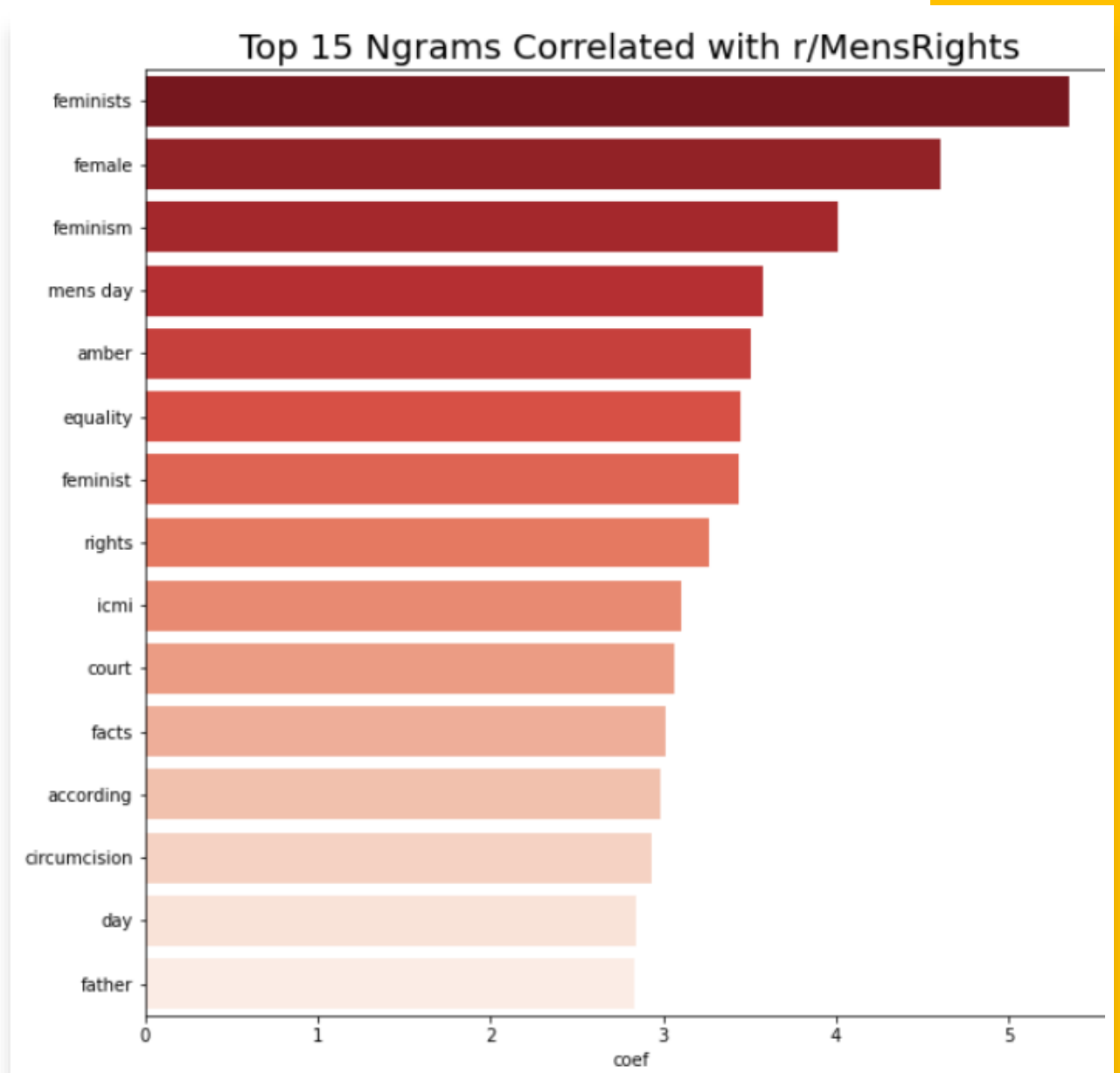
- So a common argument **feminists** LOVE to use is the “**feminism** cares about men’s issues too” bullshit line used, well the nice thing about today (international men’s day) is it really proves them wrong, go look at every single **feminist** sub on Reddit, every **feminist** page on Facebook, any feminists blog or journal.... guess what, radio silent on men’s issues on the one day of the year dedicated to men’s issues, yet every men’s rights page and sub is covered in IMD posts. **Feminism** straight up does not care about men’s issues and this proves it. **Feminism** is not about equality but supremacy. **Feminism** hates men.



# r/MensRights:

False Accusations Against Men  
e.g. Amber Heard

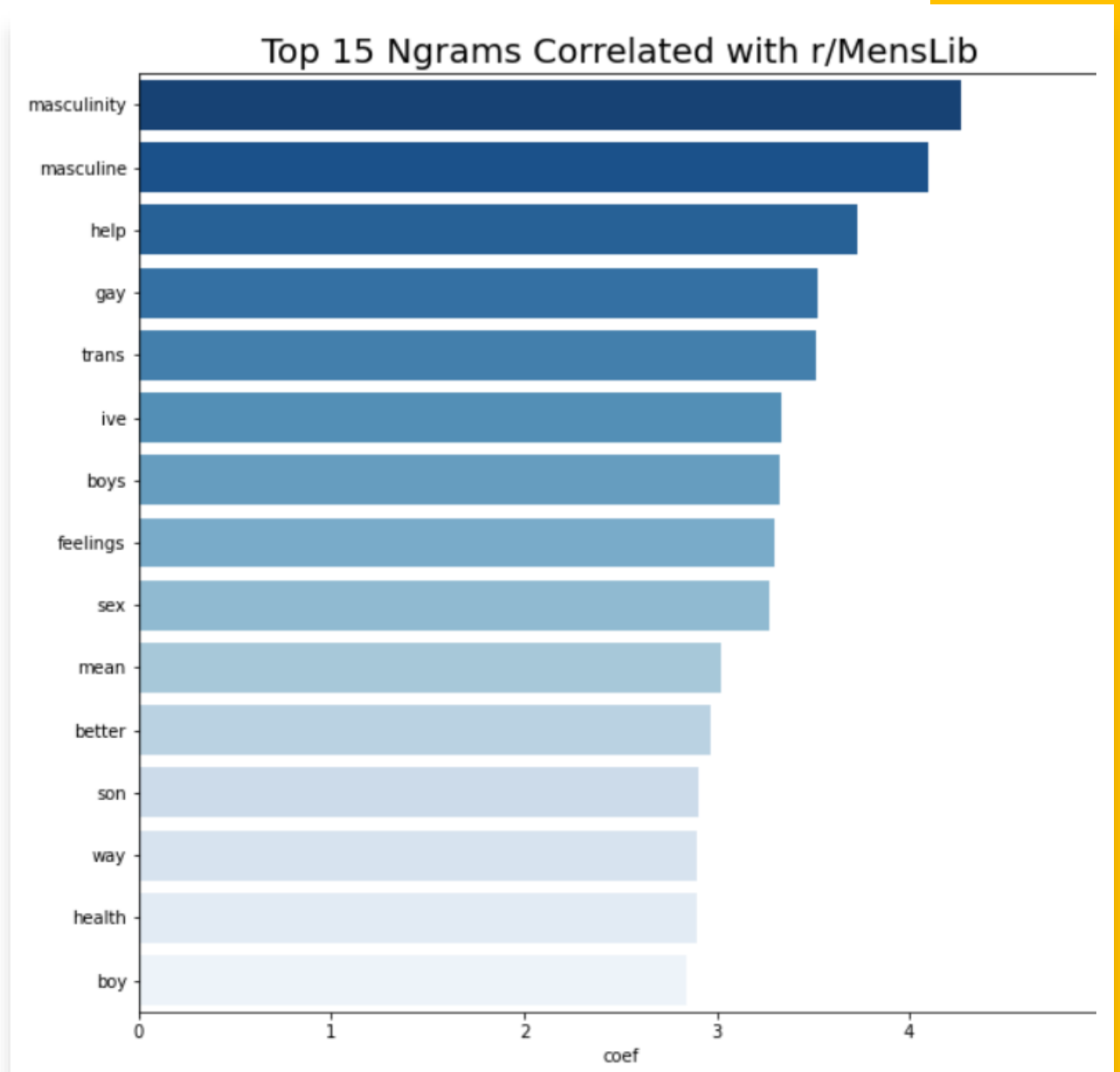
- “Sign to push for **Amber Heard** to be fired. Why? Because men are targeted and women get let off”
- “**Johnny Depp** loses to amber heard. Where is the justice?”
- “**Johnny Depp** Is Being Blackmailed by **Amber Heard**”
- “The **amber heard** situation is the best way to expose misandrist, and how people believe women more. People who defend **amber heard** are misandrist who knows nothing about false allegations. And the people defending her are so caught up in the “believe all women” bull shit that they are being completely blind of the truth.”



# r/MensLib:

## Questioning 'Traditional' Masculinity

- “Whether I’m on a basketball court, in a boardroom, at the pub, or in the bedroom, my **masculinity** will be different in each context [...] It’s a performance, and it changes based not only on the setting, but also the other players taking part in the scene.”
- “It's obvious that \*toxic\* **masculinity** is bad, but what about concepts of **masculinity** and femininity as a whole? Are they necessary to keep around, or do they only exist to put people in boxes and perpetuate stereotypes? “

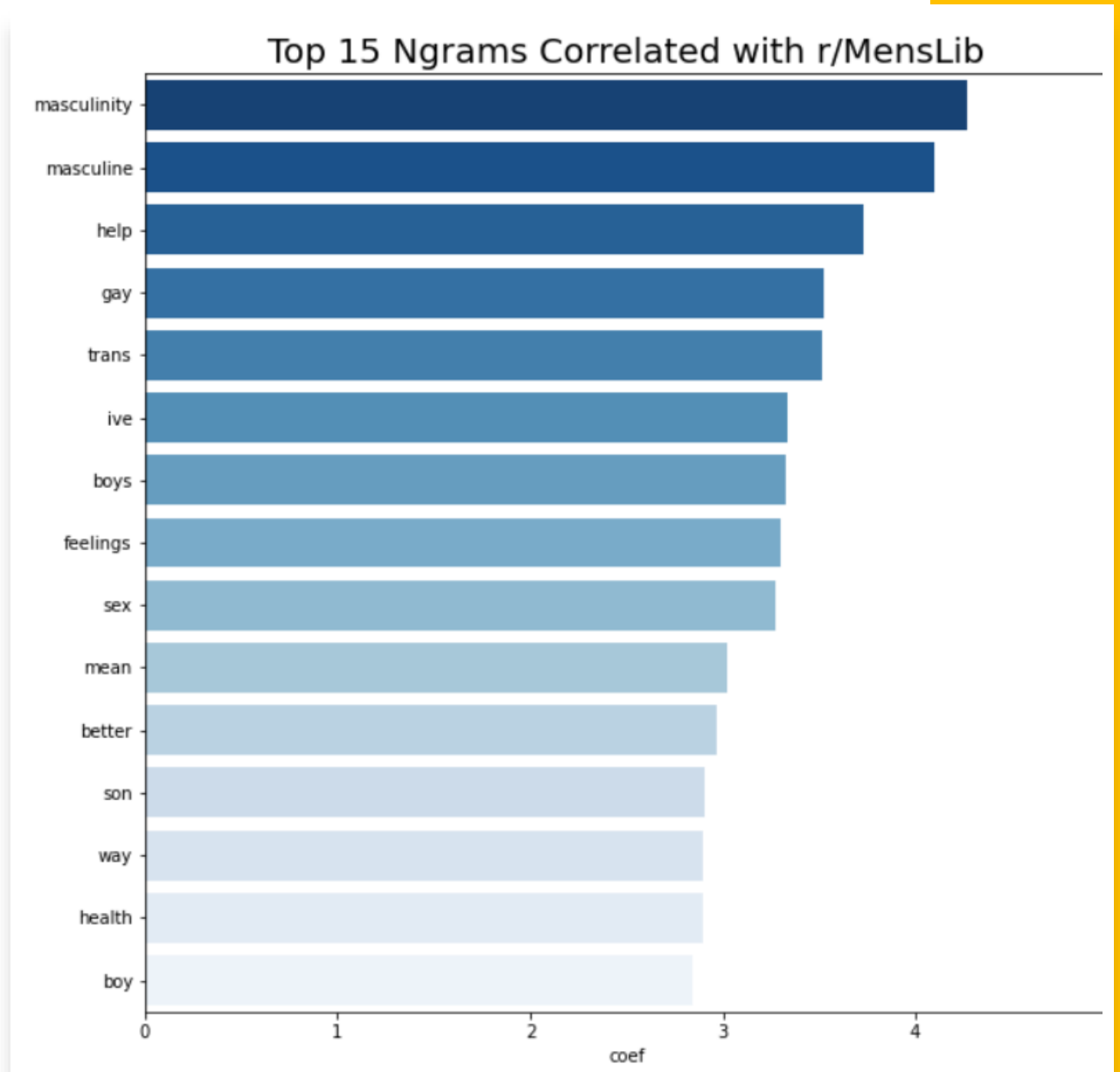




# r/MensLib:

## Discussion of LGBTQ/POC Issues

- “It didn't take me very long even as a kid to realize that I didn't really fit the societal image of what a **black man** should be... I never did grow into common cultural image of **black men** as tall, muscular, sexy, confident”
- “I am also a **disabled** man. I will never be a "protector" or any number of modern dominant culture stereotypes [...] Not meeting those stereotypes AND finding a powerful masculinity in myself is an active part of my liberation”
- “I am a **trans man** and "man" is an active part of who I am, day and day out”



# Where the Model Fails

## Shared Issues & Topics:

- “Could 6 hour shifts improve men's [mental health](#)? Finland and Sweden are slowly moving towards a 6 hour shift workday while maintaining the 8 hour pay. Do you think this could help men's [mental health](#)? Keep in mind companies see an increase in productivity and they could hire 30% more workers for the extra shift and employees are less stressed!” – misclassified as r/MensLib
- “The Trial of [Johnny Depp](#) and [Amber Heard](#). Hi all, first off, I wanted to say I love this group. Everyone here seems so level headed, engaging and takes time to read posts carefully. What a fabulous lot. I have been thinking about the current domestic abuse trial between [Amber Heard](#) and [Johnny Depp](#)” – misclassified as r/MensRights

## Short Posts:

- “Supreme [Court](#) says federal law protects [LGBTQ](#) workers from discrimination. There is a bias against men in family [court](#) says Karen Straughan.” – misclassified as r/MensRights
- “[Creative](#) Men support group. I have a weekly men's [support](#) group for [Creative](#) Men or men that aspire to be more [creative](#). We meet on Saturdays from 1-3pm CST. The topic today will be [collaboration](#) and [networking](#). “– misclassified as r/MensLib

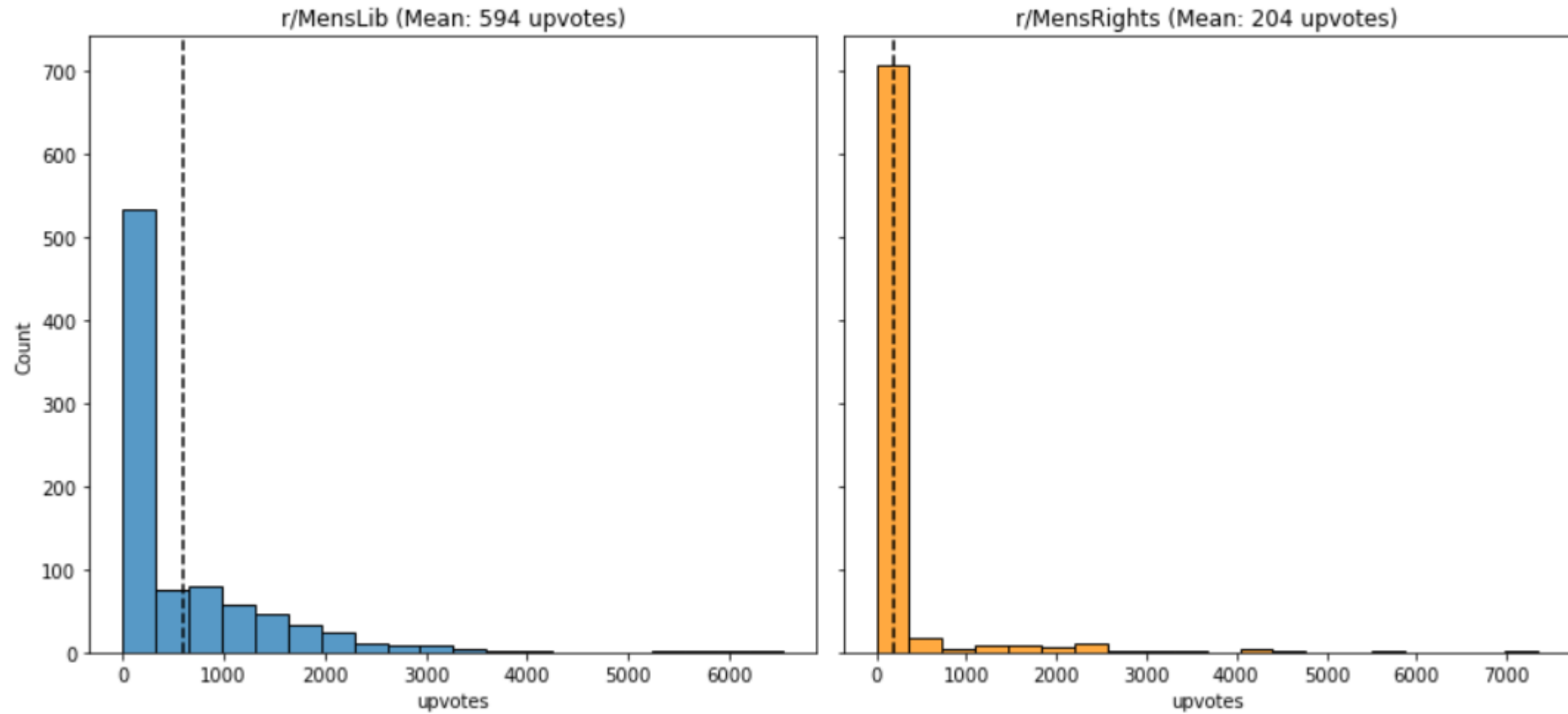
# Conclusion

- When it comes to the wire, both subreddits care deeply about similar issues facing men (e.g. male suicide, male parenting), even if their approach and beliefs are fundamentally different.
- To further improve model accuracy, we'd ideally need to train our model to recognize more abstract concepts such as the level of introspection or sentiment within the post.
- This would allow our model to deal with issues that are topical to r/MensRights (e.g. Johnny Depp and Amber Heard) and classify posts not just by mention of a name, but also by syntactic patterns that suggest introspection and openness such as like 'I have been thinking' or 'I'd love to hear people's views'.

# Appendix

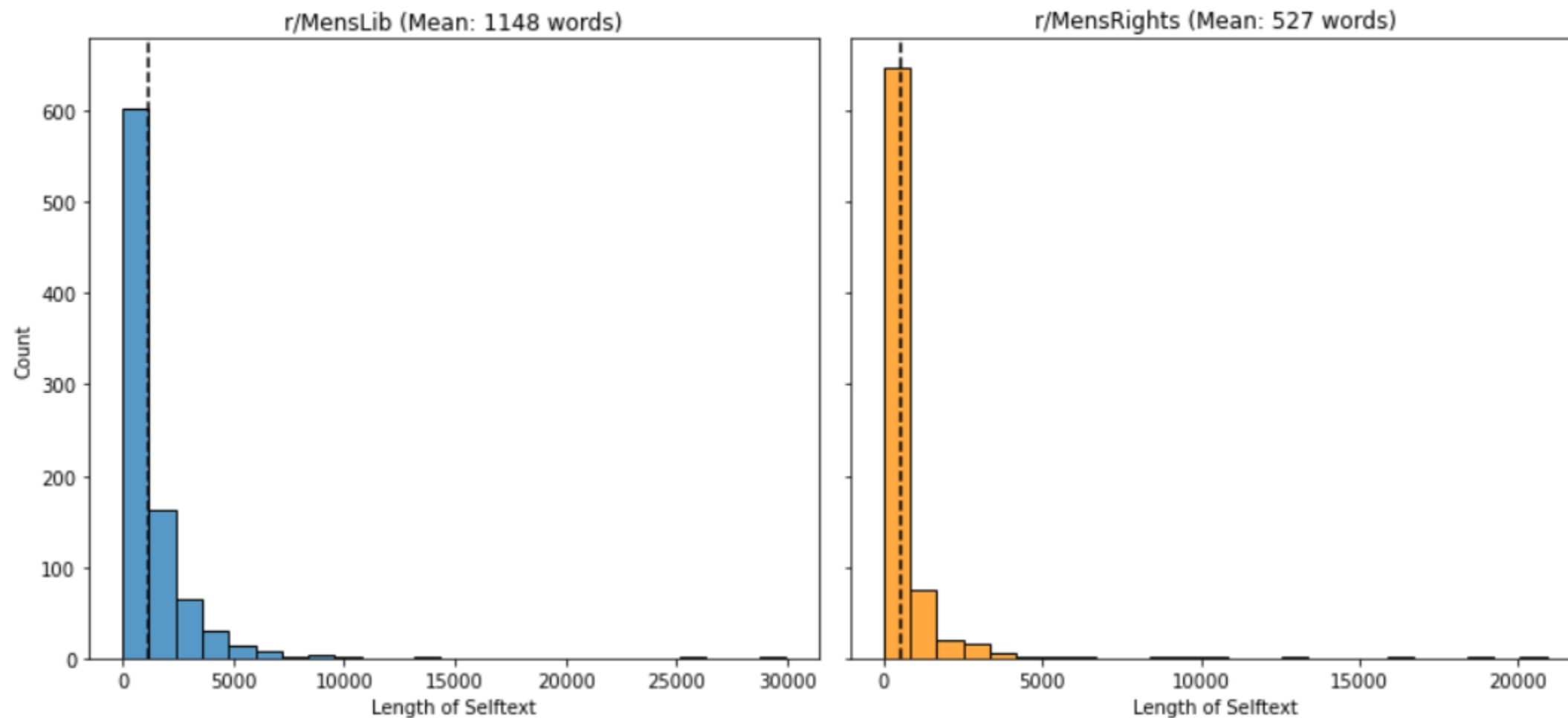
# EDA

## Upvotes



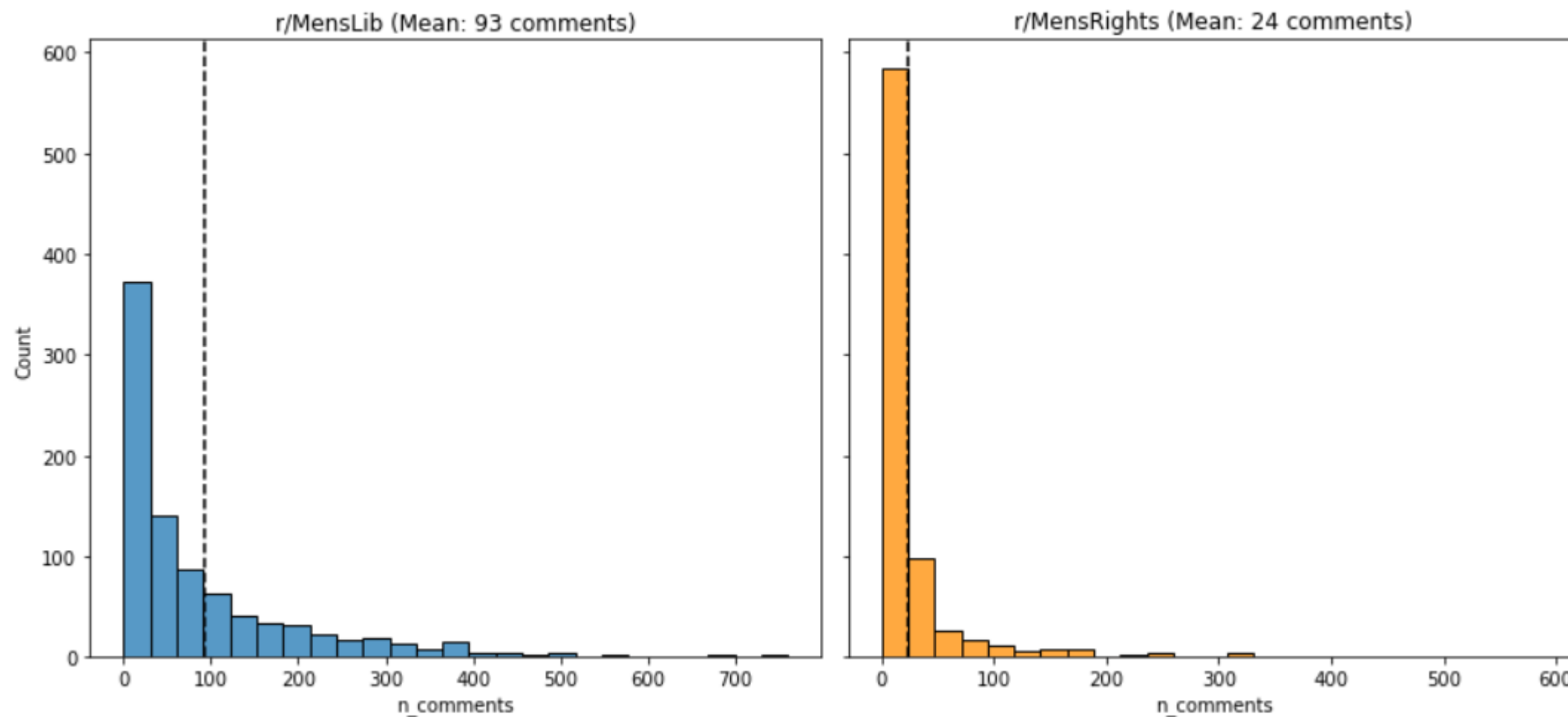
# EDA

## Average Post Length



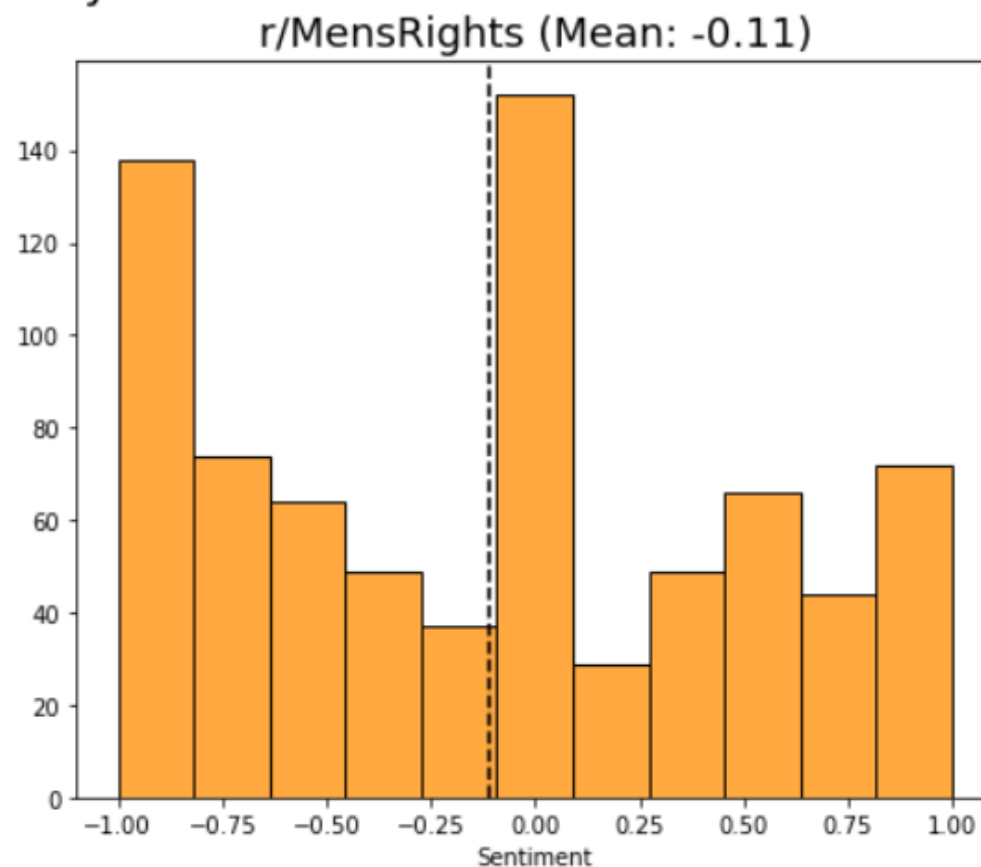
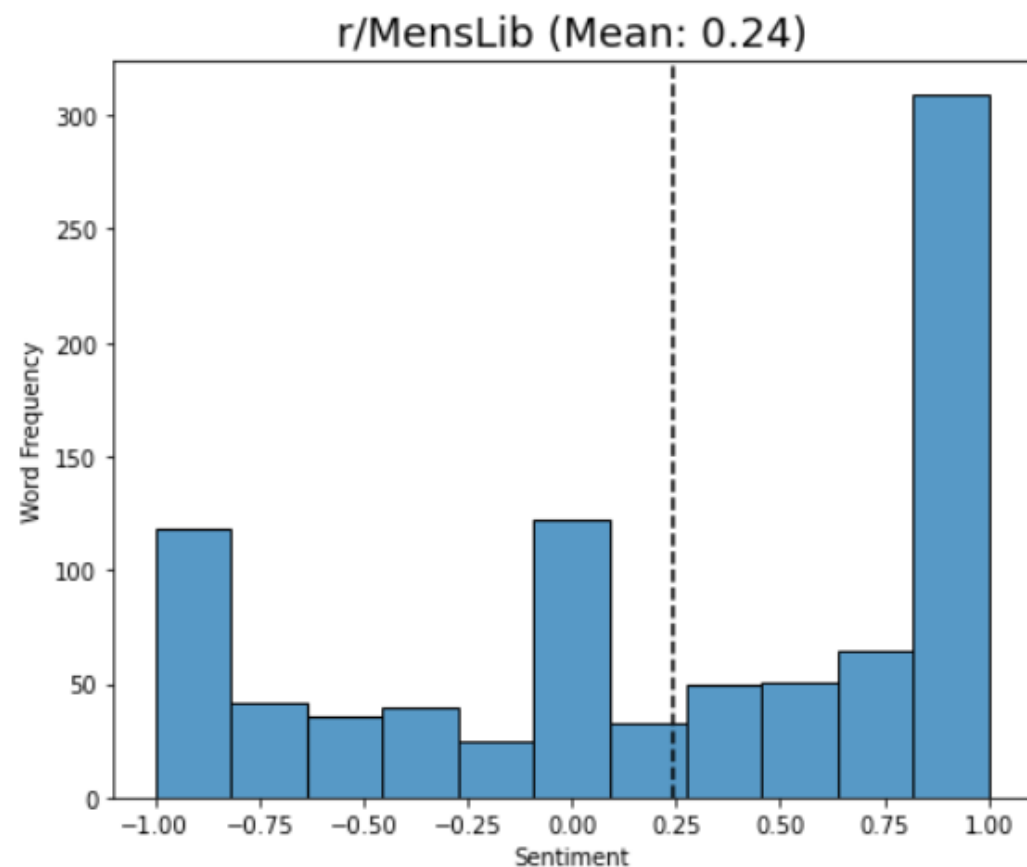
# EDA

## Number of Comments



# EDA

## Sentiment Analysis

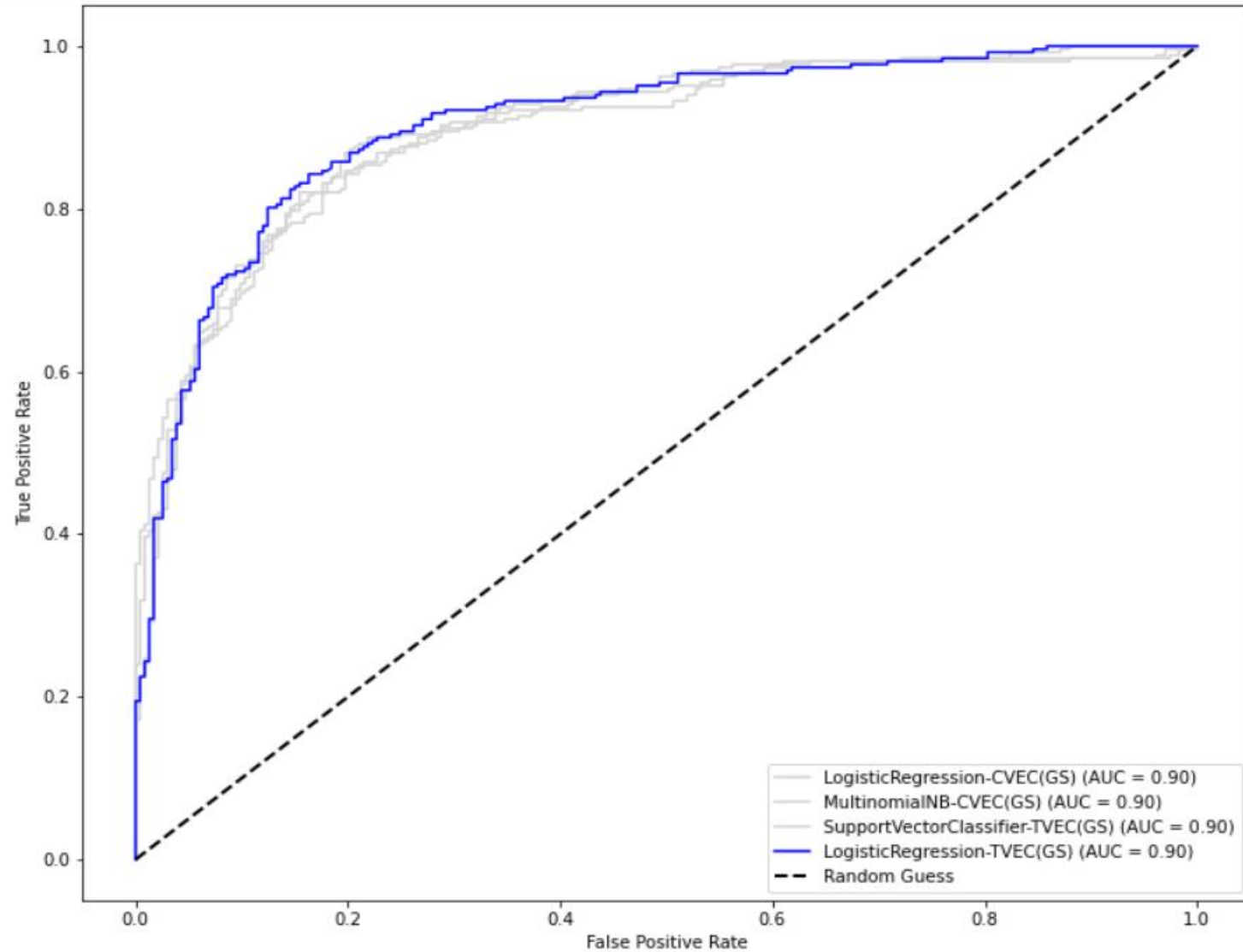


```
# Initialise VADER
sid = SentimentIntensityAnalyzer()

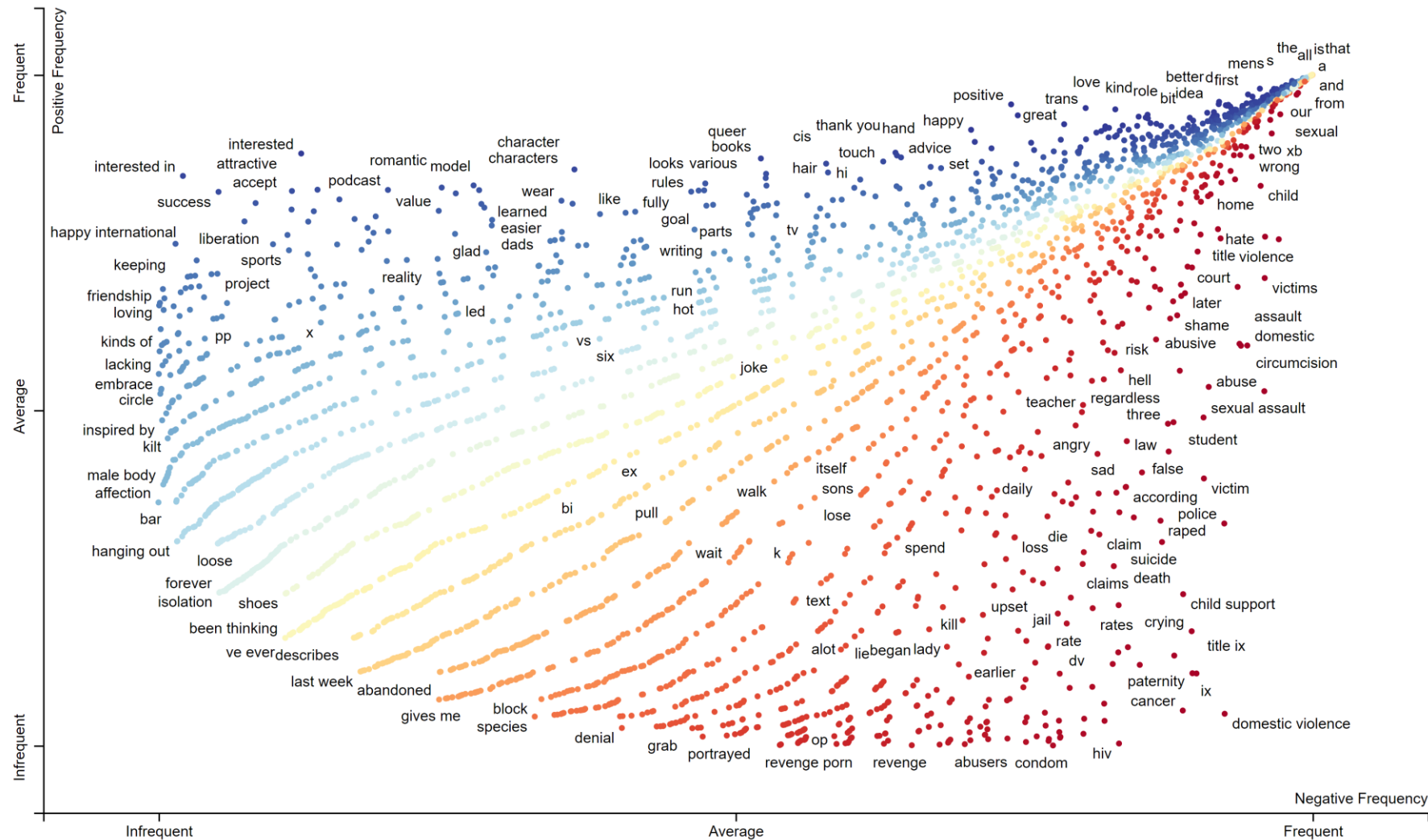
# Analyzing sentiment with VADER
sentiment_df['scores'] = sentiment_df['combi_text'].apply(lambda x: sid.polarity_scores(x))
```



# AUC-ROC Curve



# Sentiment - Alternative Classification Measures



| Top Positive      | Characteristic |
|-------------------|----------------|
| positive          | masculinity    |
| trans             | covid          |
| masculinity       | subreddit      |
| love              | patriarchy     |
| good              | feminists      |
| healthy           | tldr           |
| friends           | rmenslib       |
| better            | reddit         |
| us                | menslib        |
| kind              | misandry       |
| feel like         | youtube        |
| guys              | sexism         |
| myself            | twitter        |
| d                 | feminism       |
|                   | masculine      |
| Top Negative      | Characteristic |
| violence          | symere         |
| abuse             | phimosis       |
| victims           | incel          |
| rape              | instagram      |
| child             | circumcision   |
| domestic          | andor          |
| circumcision      | circumcised    |
| assault           | shaming        |
| domestic violence | femininity     |
| court             | metofo         |
| sexual            | nt             |
| police            | sexist         |
| wrong             | antifeminist   |
| ix                | mras           |
|                   | nonbinary      |

# Model Performance

|   | model | vectorizer | train    | test  | roc      | specificity | sensitivity | f_score  |
|---|-------|------------|----------|-------|----------|-------------|-------------|----------|
| 0 | lr    | tvec       | 0.993133 | 0.834 | 0.830086 | 0.817241    | 0.887640    | 0.850987 |
| 1 | svc   | tvec       | 0.972532 | 0.826 | 0.821503 | 0.806122    | 0.887640    | 0.844920 |
| 2 | nb    | cvec       | 0.893562 | 0.824 | 0.824275 | 0.845560    | 0.820225    | 0.832700 |
| 3 | lr    | cvec       | 0.987983 | 0.824 | 0.823456 | 0.837736    | 0.831461    | 0.834586 |