

Can't Live With 'Em: r/AskFeminists VS. r/MensRights



Reddit NLP Analysis
Kate Dowdy

Problem Statement

After conducting NLP on content from the subreddits 'AskFeminists' and 'MensRights', how accurately can we predict which subreddit new content belongs to?

r/AskFeminists

“If you've wondered what most feminists think about certain things, what our response is to certain issues, how we think certain things should be handled, or why we have adopted the positions and stands that we have, this is your place to get your questions answered!”

Subscribers: 18.6K

r/MensRights

“The Men's Rights subreddit is a place for those who wish to discuss men's rights and the ways said rights are infringed upon.”

Subscribers: 198K

Top All-Time Scoring Posts:

r/AskFeminists

“What rights, historically denied to women, are you most grateful for today?”

(122 upvotes)

r/MensRights

“How to get banned from r/Feminism”

(30.5K upvotes)

Web Scraping / Processing

Web Scraping:

- Pushshift (submissions + comments)
- 25,000 comments + 5,000 submissions from each sub from past 12 months
 - There were only ~5K posts in AskFeminists over the last year; MensRights had significantly more

Data Cleaning:

- Tracked whether comment/post, whether 'removed' or 'deleted'
 - Comments/posts 'removed': MensRights: 223, AskFeminists: 1847
 - Comments/posts 'deleted': MensRights: 499, AskFeminists: 112
- Removed links, lemmatized words
- Cleaned up contraction remnants

Top 10 Phrases

r/AskFeminists

doesn['t] make sense	43.553172
false rape accusation	35.549279
difference men woman	32.353939
traditional gender role	29.493159
just don['t] think	26.809705
don['t] really know	26.759491
reflect feminist perspective	25.933302
innocent proven guilty	24.901007
rape sexual assault	24.534936
gt don think	24.006895

r/MensRights

false rape accusation	49.574165
men right issue	42.888562
men right movement	42.000943
international men day	35.123406
gender pay gap	34.188866
innocent proven guilty	33.367548
men right activist	26.564834
new york time	22.744520
doesn['t] make sense	21.240539
year old boy	20.441459

Top 10 Ngrams (1-2)

r/AskFeminists

wife ha	976.118614
men family	729.171984
fellow	677.549175
thing	612.694714
divorce	607.185712
partner violence	605.775949
lifestyle	550.282282
june	537.683578
voting	513.169382
greatest	450.885487

r/MensRights

woman	868.976790
men	826.650364
wa	579.965131
just	491.954841
like	476.984193
don	462.072782
people	423.281936
gt	407.147621
right	403.265087
feminist	391.284568

Top 15 Coefficients

(most influential words/phrases for predicting subreddit)

r/AskFeminists

pre determined	-5.681340
study seminar	-4.972181
talking federation autonomous	-4.799818
eff	-3.908666
anti sjw	-3.638601
desperate attempt	-3.596841
shitty yeah	-3.415897
pick apart	-3.373890
okay sexualize	-3.134356
exposed message	-3.017257
brainwashing hell drug	-2.897521
duo	-2.883798
distinguish	-2.873333
non binary particularly	-2.842388
appearance body	-2.806616

r/MensRights

member dominant	5.570271
bread scientist react violently	4.393380
argument moving	4.075883
message author	3.676436
sexually submissive	3.567055
brave serve	3.275661
mobile atm	3.235209
manifesting	3.220126
self righteousness simply	2.914714
atlassian accepted majority coder	2.782258
beg scrap	2.760610
slut slut	2.736139
expendable	2.723877
month son	2.716348
bias court	2.708516

Modeling Results

TFIDF Vectorizing:

- ngram_range = (1,4)
- max_features = 150_000
- stop_words = stop_words_2

Predicting MensRights (LogReg):

True Negatives: 11235

True Positives: 11067

False Negatives: 3774

False Positives: 3652

Logistic Regression

Train: 0.8656

Test: 0.7521

(penalty = Ridge)

Naive Bayes (Multinomial)

Train: 0.8834

Test: 0.7301

Support Vector Classifier

Train: 0.9840

Test: 0.7311

(kernel= 'rbf', C = 100,
gamma = 0.05)

Misclassified Examples

[Actually from MensRights]:

“Your a sick person for enjoying what he going through.”

“...Women in office make me feel emotionless for not showing more outward affection for a co-worker who lost a loved one. Wondering if there's a term for this.”

[Actually from AskFeminists]

“!!||* SWERF ALERT |*|*!!”*

“Actually, i think thats kind of what i was saying... its not some grand conspiracy! We are humans, humans want to be accepted and in that rush for acceptance we hurt eachother and ourselves. Its not a conspiracy, but it still exists... so lets fix it! We are idiots, but we can be better!”

Conclusion

More Questions Than Answers...

- What's the overlap in thread participants (are the MensRights subscribers posting on AskFeminists, or vice versa?)
- How have topics of discussion changed over time (compared to stories in the news?)
- How does sentiment analysis compare across subs?
- Have the subs become more radical/neutral since 2016? Does strictness of their mods make a difference?