

CSE6240 Project Presentation

Identifying Misogynistic Rhetoric With AdviceSeeking Communities

Tanvi Bhagwat, Disha Das, Aiswarya Bhagavatula



Reddit in Popular Culture

- Reddit has been consistently among the top ten most frequented websites for the last few years
- Posts from the website often make way from the front page to mainstream media
- Its anonymity helps users engage in discourse freely

How Reddit's "Am I the Asshole?" became the definitive document of the Trump era

The popular subreddit might be the most philosophical place on the internet—and the perfect metaphor for how we live now.

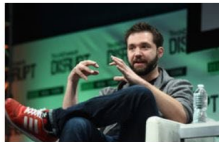
May 14, 2015

TECHNOLOGY

Reddit Introduces Anti-Harassment Policy

Reddit users who view or experience harassment on the service will now be able to e-mail the company and request to have content and offenders removed from the site.

By Mike Isaac



PRINT EDITION Reddit Enacts a Policy to Prohibit Harassment | May 14, 2015, Page B2

ANNALS OF TECHNOLOGY

THE WEB'S FAILED HUNT FOR THE BOSTON BOMBER

By Matt Buchanan April 19, 2013





Baseline Approach and Description

- The approach used by Farrell et al. in their paper, Exploring Misogyny across the Manosphere in Reddit, was used as a baseline for this project
- Communities selected identified with men's rights activism and incel idealism
- To characterize the misogyny they built 9 lexicons for the different levels of hate speech
- These constructed lexicons were used to calculate the amount of misogynistic posts per community, the amount of users posting such content per community, the top terms in each community, and the evolution over time of these communities for the different levels of misogyny.



Project Approach

- The communities analyzed by the baseline work were already infamous for their ideologies
- We aim to analyze advice- seeking communities and monitor spread of hate speech and toxic ideologies within these communities
- ‘Community Interaction and Conflict on the Web’ by Kumar et. al showed that users engagement drops when they start facing negativity in certain communities
- We seek to verify this within selected Reddit communities

Humans and technology

The “manosphere” is getting more toxic as angry men join the incels

Men from the less extreme end of the misogynistic spectrum are drifting toward groups that espouse violence against women, a new study suggests.

by **Tanya Basu**

February 7, 2020

MIT Technology Review



Novelty

- We seek to explore the effect of hate speech within the advice seeking community
- The baseline approach performed the following analysis:
 - Analyzing the most frequent misogynous words
 - Finding the percentage of misogynous posts
 - Finding the frequency of misogynous words used over time
- We introduce 3 new avenues of analysis:
 - Analysis of negativity for over 18 and under 18 users
 - Relation between negativity in title and comments
 - Whether unpopular posts receive more negativity
 - Where do the users with negative submissions post most frequently in



Data Acquisition

- Data from following Reddit communities were acquired: AskReddit, relationships, relationship_advice, AmITheAsshole and OffMyChest.
- Used a Reddit Search Application, Pushshift.io, to mine the posts from five subreddits and chose the top 1000
- Mined comments under posts using Reddit's API The Python Reddit Wrapper(PRAW)
- Data Preprocessing was performed to eliminate irrelevant columns and convert data in the csv file to suitable formats (DateTime Objects, sorting, etc.)



Data Properties

- Features of data used in this project:

Data Features	Utility in project
Post author name	Used this feature to analyze user involvement over time
Number of comments	Analyze popularity of post
Over 18 or not	Analyze appropriate content for young users
Score	Upvotes- Downvotes. Analyze Popularity of post
Title of post	Analyze misogynistic content
Comments	Analyze misogynistic content



Experimental Setup

- Lexicon of Misogynistic words had a selection from these categories: Belittling, Flipping the narrative, homophobia, hostility, patriarchy, physical violence, racism, sexual violence and stoicism
- The analysis was done on Google Colaboratory and Kaggle
 - Runtime Type: Python3
 - Hardware Accelerator: GPU
 - RAM: 12 GB

Category of Misogyny	Num terms	Examples
Belittling	58	femoid, titties, stupid cow
Flipping the narrative	7	beta, normie, men's rights
Homophobia	126	dyke, fistfucker, faggot
Hostility	303	bitch, cunt, whore
Patriarchy	8	alpha male, subjugate, suppress
Physical Violence	73	hit, punch, choke
Racism	670	nigger, raghead, pikey
Sexual Violence	22	rape, sodomise, gangbang
Stoicism	33	blackpill, cuck, hypergamy

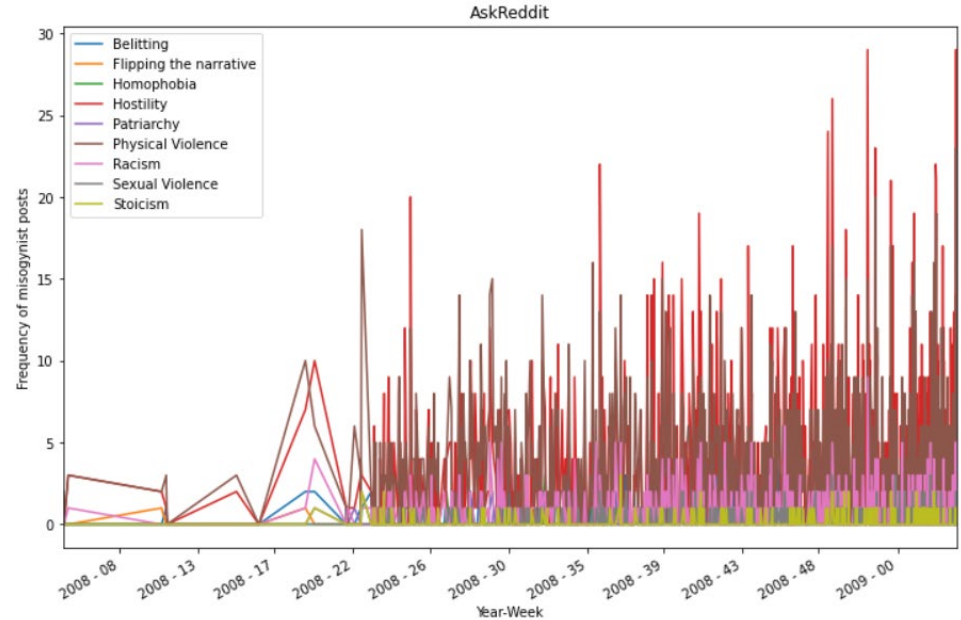


Data Analysis Approach

- For the sake of simplicity, data analysis of one Reddit community has been presented here.
- Reddit community information::
 - Community: AskReddit
 - Date range: 2008-02-08 to 2009-01-24
 - Number of posts: 1000
 - Avg number of comments per post: 137

Data Analysis and Results

- Detected number of misogynist words used in comments section of posts and plotted their frequency over time. This was done for each community, and for each category of misogyny from the lexicon.
- Observations:
 - Growing trend over the year
 - Could be due to more user engagement in the Christmas holiday season
 - Words associated with Physical Violence and Hostility see a steady growth





Data Analysis and Results

- Found the percentage of misogynous posts for each category of misogyny.
- A misogynous post has a comments section that contains any misogynistic words found in the lexicon.
- User most likely to come across references to Physical Violence (80%) , Hostility (78%) , Belittling (48%) and Racism (47%).

Community	% of Mis. Posts
Belittling	48
FlippingNarr	4
Homophobia	11
Hostility	78
Patriarchy	6
P. Violence	80
Racism	47
S. Violence	18
Stoicism	11



Data Analysis and Results

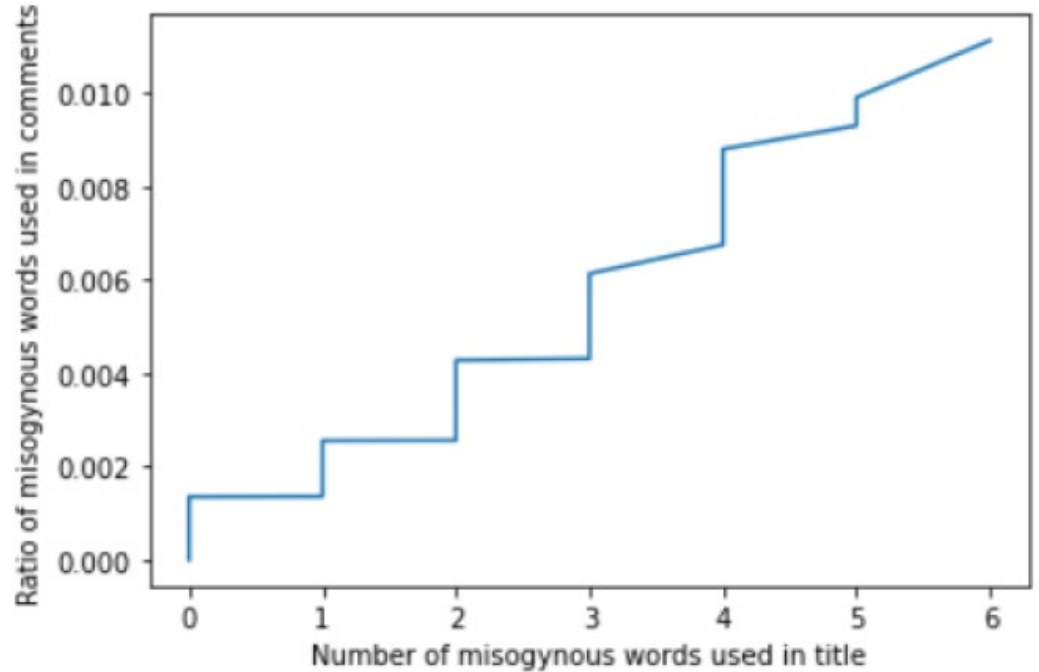
- Found the top misogynous words and their frequencies used across all categories.
- Most frequently used words include Hate, Hit, Cut, Kill etc.
- These words point to some violent attitudes in the comments section of a community that elicits further hateful discussions.
- Top 5 words from some popular categories are given below

Homophobia		Hostility		Patriarchy		P. Violence		Racism		S. Violence	
Word	Freq	Word	Freq	Word	Freq	Word	Freq	Word	Freq	Word	Freq
Fag	35	Hate	479	Oblige	13	Hit	406	Black	328	Rape	101
Faggot	22	Hurt	264	Suppress	12	Cut	350	Bigger	189	Pound	47
Homo	17	Bitch	238	Overwhelm	11	Kill	322	Nigger	19	Conquer	15
Queer	16	Beat	229	Omega	11	Force	242	Arab	17	Spank	10
Fags	15	Asshole	201	Compel	7	Kick	178	Crackers	17	Lolita	9



Data Analysis and Results

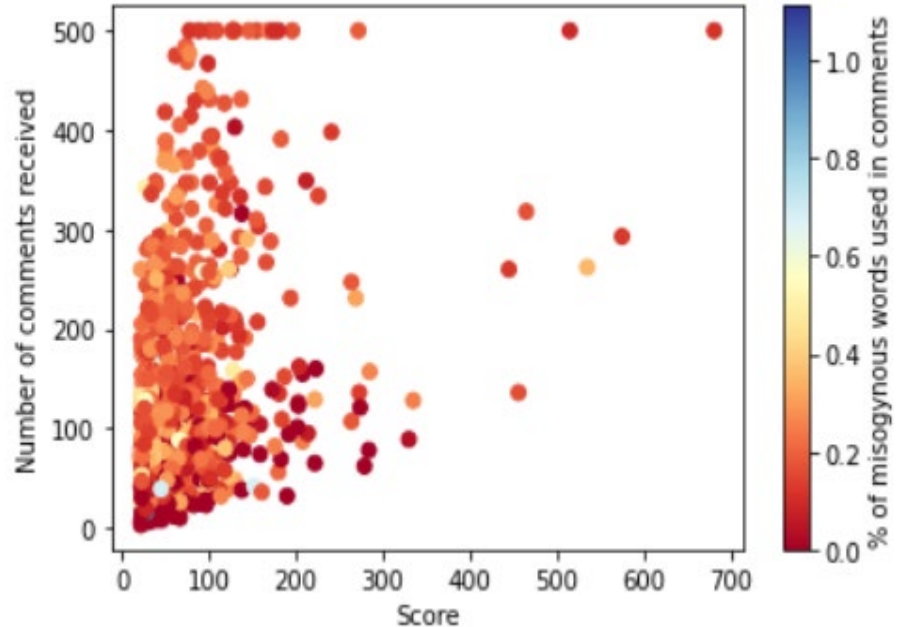
- Found correlation between frequency of misogynistic words in the title of post and comments section of post
- Found a linear relation
- Questions asked in relation to misogynous content invited more misogynous discussions in the comments section





Data Analysis and Results

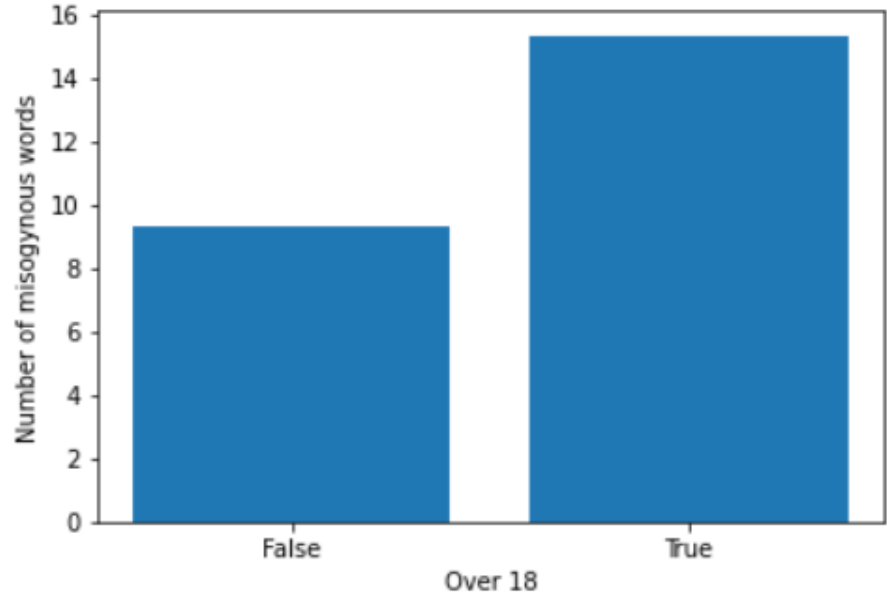
- High number of comments and low score shows that the post received a lot of downvotes.
- Posts having lesser score than the number of comments generally have a high percentage of misogynous words (0.2% - 0.5%) in the comments section





Data Analysis and Results

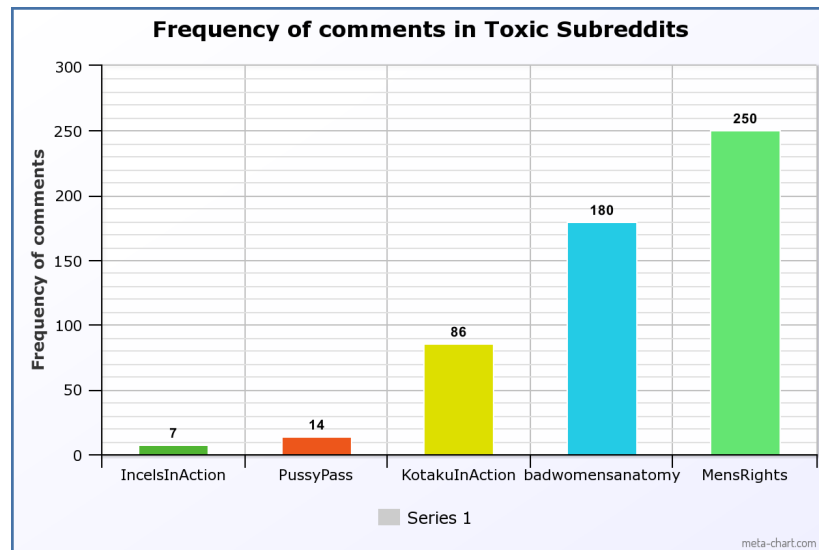
- Number of misogynous words found in posts for users above 18 and under 18 was plotted
- For AskReddit community, underage users were exposed to less hostile comments as compared to users above 18





Data Analysis and Results

- So what communities do the users engaging in negative discourse contribute the most to?
- These communities have been on Southern Poverty Law Center's hatewatch



Baseline Result Comparison

Belittling		FlippingNarr		Homophobia		Hostility		Patriarchy		P. Violence		Racism		S. Violence		Stoicism	
word	freq	word	freq	word	freq	word	freq	word	freq	word	freq	word	freq	word	freq	word	freq
female	121549	mgtow	38862	faggot	7453	hate	89389	suppress	1188	hit	37142	black	49552	rape	58396	incel	111373
dumb	31312	beta	33450	fag	1856	bitch	54609	betabuxx	1023	kill	29006	bigger	15545	cock carousel	3750	chad	36081
roastie	9143	normie	16975	fags	1407	pussy	47869	omega	786	cut	21888	nigga	3575	pound	2703	cuck	31482
boobs	8354	mra	424	homo	1278	hurt	26440	subjugate	300	force	17740	nigger	1030	conquer	1543	cope	26169
failure	7952	overthrow	368	dyke	879	cunt	24150	oblige	255	attack	12061	niggas	1006	incest	1378	blackpill	13621

Table 1 : Baseline Results

Belittling		FlippingNarr		Homophobia		Hostility		Patriarchy		P. Violence		Racism		S. Violence		Stoicism	
Word	Freq	Word	Freq	Word	Freq	Word	Freq	Word	Freq	Word	Freq	Word	Freq	Word	Freq	Word	Freq
Female	203	Beta	32	Fag	35	Hate	479	Oblige	13	Hit	406	Black	328	Rape	101	Fuel	47
Dumb	196	Overthrow	5	Faggot	22	Hurt	264	Suppress	12	Cut	350	Bigger	189	Pound	47	Cope	37
Failure	108	Prevail	5	Homo	17	Bitch	238	Overwhelm	11	Kill	322	Nigger	19	Conquer	15	Rope	33
Bird	80	Vanquish	1	Queer	16	Beat	229	Omega	11	Force	242	Arab	17	Spank	10	Chad	7
Birds	56			Fags	15	Asshole	201	Compel	7	Kick	178	Crackers	17	Lolita	9	Friendless	1

Table 2 : Our Results



Future Work

- Due to time restrictions, we could only perform the analysis on 5 subreddits. The work can be extended to other subreddits to gain a larger dataset.
- Additionally, more in-depth analysis can be made to identify a trend in the content projected by a given user to spot possibilities of severe problems.
- Algorithms using the concepts of Machine Learning and Deep Learning can be incorporated to train a model on the dataset to alert moderators quicker, or indicate to the user how seriously to take advice espoused by fellow Redditors.