

Analysis of Hate Speech and Cyberbullying in Social Media

Student Name: **Gaurav**

Roll Number: **2021147**

BTP report submitted in partial fulfillment of the requirements for the
Degree of B.Tech. in Electronics and Communication engineering on 13th
December 2024

BTP Track: Research

BTP Advisor: **Dr. Soibam Haripriya**

Abstract

Social media is one of the most innovative technologies in the world. The concept of people from around the world collaborating and communicating with each other in the comfort of their homes has transformed how human beings interact with each other. This enabled people to share their own views and opinions into cyberspace and reach numerous people. We can do so at the comfort of our homes and without necessarily revealing your identity to the outside world. However, this freedom is like the two sides of a coin as people abuse social media heavily for spreading hate and misinformation. Our previous work dealt with how fake news propagates through social media. In this study, we will see 4 case studies where people have suffered heavily due to the hate propagated against them in social media. The common link between all these cases is that all of them were targeted for their political views and for political causes. We will see the dynamics of how and why they were targeted and if the existing rules are robust enough to contain hate speech and online abuse.

Acknowledgments

I would like to express my sincere gratitude to my advisor, **Dr. Soibam Haripriya**, for providing me with this opportunity and all his guidance throughout this project. The regular meetings, constructive feedback, and constant support were of great help to me in accomplishing each task.

Contents

1	Introduction	iv
1.1	Introduction	iv
1.2	Motivation	v
2	Case Studies - Gender Based	vi
2.0.1	Rana Ayyub	vii
2.0.2	Meghan Markle	x
2.0.3	Hate speech against women dataset overview	xiii
3	Case Studies - Caste Based	xiv
3.0.1	Meena Kotwal (India)	xiv
3.0.2	Gauri Lankesh	xv
4	Analysis	xvii
4.1	Acts against online hate speech	xviii
4.2	Some supreme court jurisdictions	xviii
4.3	Some caste based hate speech tweets	xx
4.4	Some gender based hate speech tweets	xxii
5	Conclusion	xxvi
6	Future Work	xxvii

Chapter 1

Introduction

1.1 Introduction

Twitter, facebook, instagram, whatsapp, telegram - almost all of the world population has heard about these names. That is the kind of influence social media has on the general populace nowadays. Almost everyone uses at least one or more of these apps daily and are widely influenced by them. News channels also have a lot of influence on the general public. In the case of social media, it enables one to share their opinions and views on anything and everything and are protected by anonymity if they choose to be. This has allowed a lot of people to raise their voice for and against a lot of things and has made a great impact on many things. However, if a group of people doesn't approve of what someone has to say, or if they straight up reject or oppose their views and ideologies, it can be a stage for hate speech and abuse against individuals. These could be political, religious or any other reason. In this study, we are going to look at few examples who were heavily targeted by hate speech and abuses by people who didn't not agree with them on many things. All of these people have been targeted due to their political views. The violators targeted their gender, caste, sexual orientation etc but we will focus on just the first two here.

1.2 Motivation

Social media has become a major part of our daily lives, giving everyone a platform to share their thoughts and opinions. However, with this freedom comes the darker side of online hate and harassment, especially targeting people who speak up on political issues. Many public figures, activists, and journalists have faced severe abuse for expressing their views. This has led to a rise in online hate speech, which can affect people's safety, mental health, and reputation.

In India, where social media plays a huge role in political discourse, these issues are even more prevalent. We are motivated to study this topic because it's important to understand how hate speech spreads online and how it impacts individuals, especially when the hate is driven by political disagreements. By looking at real-life case studies, we aim to explore the extent of this problem and whether the current laws and policies are enough to address it. This study is also a step towards finding better solutions to ensure that social media remains a space for healthy and respectful discussions.

Chapter 2

Case Studies - Gender Based

Hate speech against women due to political disagreements

In the present world, it is pretty common to see women increasingly becoming involved in politics and expressing their political stances and opinions. Social media is a very important factor for this cause. But it has been increasingly been abused to propagate hate against women specifically. This could lead to women moving away from the social media platforms or limiting the amount of content they post. It is against free speech where you face abuse when you post stuff especially when its women. It thus means that women have limited capacity to engage with the issue of politics since they will face hate online. For our research, we have selected two highly recognized cases, namely Rana Ayyub and Meghan Markle, because they both went through severe hate online because of their political views and activism. Rana Ayyub is a journalist from India who has been targeted based on her views against the government, especially since she came out with a book on the Gujarat riots. Meghan Markle, Duchess of Sussex, has been put under criticism for speaking out on racial justice, walking away from the royal family, and sharing her personal issues, which created strong responses online. Both of them faced not only political attacks but also gendered abuse. Their personal lives and identities were distorted to bring them down. In their cases, we wish to study how being a woman in the public eye, especially with political disagreements, leads to more hate speech and online abuse.

2.0.1 Rana Ayyub

Rana Ayyub is an Indian journalist who gets hate online purely because of her political and religious ideologies. Her case reflects how social media has been used for targeted campaigns, misinformation, and organized harassment that is happening in really sensitive areas. This case can be an insight into the gendered abuse and religious biases as well, an example of how intersectional identities multiply the intensity of online attacks.

She is an independent journalist who has been the target of severe online hate speech because of her journalism and political views. She was heavily targeted because of a controversial book - Gujarat Files : The Anatomy of a coverup” which came out in 2016. From then, she has been heavily targeted by extremist groups in social media platforms , predominantly in Twitter. She has been heavily bad-mouthed on twitter and her personal details were made public in twitter, raising significant safety concerns. She was also targeted in the Bully bai app, which included several other muslim women.

She has been predominantly targeted for her political views, government criticism on topics like Covid-19, CAA, NRC, and Kashmir and the book she had published mostly.

Diving deep into why she is getting hate

Ayyub was an investigative journalist for a company named Tahleb when she covered the issue undercover. When her reports got out, she was able to send some of the top BJP leaders like Amit Shah to jail. This led to her facing significant hate against her back in 2010. Then she deleted her account and opened a new one. She continued to receive hate and threats over twitter and facebook and several rumours and videos were circulated. Then in 2014, after BJP came to power, she faced further threats. In 2016, she released her book Gujarat Files: Anatomy of a Cover-Up. This got her the labels of anti national and similar labels and faced a lot of death and rape threats. In 2017, one of her friend Gauri Lenkesh was killed for her views, which raised significant threats to her life too. This is a prime example of how online hate could translate into threats to one's life. Later on several targeted campaigns were launched against her and she was framed as an anti national person who supports various crimes. A deep fake porn video was also purported through several right wing twitter and facebook handles, even official accounts.

CAA law was passed by the government and she opposed it, which again caused a spike in hate against her. Almost all of her tweets received a heavy load of hate speech and smear campaigns against her. She also faced allegations of fraud for the fund raising she did during covid times. Well it is said that she receives hate speech every 14 seconds on social media - a scary statistic.

Works we referred

We have referred a study done by researchers from prominent news channels like Article 14 and ICFJ. They analysed around 8.5 million (collected between December 2019 and March 2022) tweets targeted against Rana Ayyub. Here is what we understood from the paper:

- Allegations of missing charity funds - She faced a smear campaign alleging her regarding the misuse of the money she raised during covid 19, leading to her assets being frozen by ED.
- She has been the target of coordinated online hate speech campaigns and trolling, largely on Twitter and Facebook. In their report, Hashtags like RanaAyyub, HateRanaAyyub, RanaAyyubExposed were used to amplify negative narratives.
- A simple sentiment analysis of tweets containing RanaAyyub (using Tweepy or Social Media Monitoring tools) can show around 70-80 percent negative sentiment, including abusive language, threats, and defamation and 20-30 percent neutral or positive (those in support of her work).
- Twitter and Facebook posts spread misinformation accusing her of being anti-national and involved with foreign powers. She has been accused of spreading fake news, anti-national rhetoric, and anti-India content.
- Over 50,000 tweets and comments have used negative hashtags targeting her. Around 60 percent of the hate tweets resorted to personal attacks, around 35 percent credibility and the rest for her belief.
- Gendered abuse is combined with religious slurs, such as “jihadi” and “terrorist”, reflecting an intersection of misogyny and Islamophobia. This overlap worsens the personal and professional toll on Ayyub.

- Hundreds of fake news articles were circulated across social media platforms, with the RanaAyyubExposed hashtag receiving over 10,000 mentions. The hashtag presstitutes (used to discredit female journalists) is one of the most common forms of abuse.
- The top 100 abusive terms include a mix of personal insults, credibility attacks, and religious bigotry. The term “hypocrite” emerged as one of the most frequent words used against Ayyub.
- Times Now (Right-wing Channel) ran stories accusing Ayyub of spreading misinformation and alleged foreign funding in her reporting. These reports were broadcasted to an audience of millions.
- Millions of viewers watched these negative portrayals. 50+ articles in major Indian media outlets such as Times of India, NDTV, and Hindustan Times have criticized her work.
- Doctored Video - Ayyub was portrayed as making derogatory comments about Muslims and India. These videos received wide circulation on WhatsApp and Twitter. A deep fake pornographic video was also circulated.
- In 2022, Ayyub was included in a “Bulli Bai” auction, where images of Muslim women were misappropriated and sold online. This event, along with other gendered Islamophobic attacks, exemplifies the coordinated nature of the abuse targeting her.
- Over 1 million views on platforms like Facebook and WhatsApp for doctored content. Hundreds of fake news websites spread false claims.
- The abuse and legal harassment not only damage Ayyub’s professional reputation but also pose significant threats to her physical safety.
- The psychological toll of the abuse is significant. Ayyub described how character assassination and threats to her personal safety often force her into a retreat from social media. She mentioned that when her critics target her as a woman, it has a powerful emotional impact, forcing her to go quiet in the face of online onslaughts.

2.0.2 Meghan Markle

Meghan Markle is a former actress, who is currently the Duchess of Sussex. She was married to Prince William of the Royal family in 2018, when she retired from her acting career. She is originally from the US and was active in social media throughout her life. She has organised many non profit organisations and has been vocal about a variety of issues. Meghan has been vocal about racial justice, gender equality, and mental health awareness, issues that often carry political weight. She has consistently championed women's rights and has been very vocal about it. Her work with organizations advocating for gender equality has intersected with political conversations about women's empowerment and representation. In the 2020 US election, she indirectly supported Joe Biden. Her public statements were viewed by many as political, particularly since members of the British royal family are traditionally expected to remain neutral. In 2021, there were speculations that she would be running for president in the next election which faced significant backlash all over the media. Meghan's political involvement has been polarizing. Critics argue she is overstepping the traditional boundaries of the royal family and has attracted a lot of hate because of that online.

Diving deep into why she is getting hate

Meghan was born to biracial parents. Before marrying Prince Harry, she has been married before and was divorced a few years before. Also she was very loud on issues like racial justice, gender equality, mental health which led to conservatives finding her views and opinions uncomfortable. She has been misrepresented and racially abused before her marriage. A notable example of the slur she faced was when she was said to have been descended from slaves while ignoring her career. Also the ultra conservative people in the UK saw this marriage as a scandal, as the royal family has been associated with a symbol of whiteness and blood purity and watching a person of mixed origin becoming a member of the royal family irked many. Some critics also found it unacceptable to see a divorced, American, biracial actress marrying into the royal family.

Then later on , Meghan and Prince Harry decided to step down from the royal family and they moved to the USA, away from the royal family. This also led to a lot of negative commentary over her influence. Most of the hate comes from the conservatives who see her as someone who

defies the royal family and appears more modern and progressive. A lot of the online hate she gets is in the form of racism, sexism, traditionalism and mostly media manipulation.

Works we referred

The most controversial event happened on twitter when there was news that she might be running for president in the next election. The study we referred focused on this event and they used a qualitative research and document analysis focussing on twitter comments. This study havent mentioned mentioned the exact number of tweets they had analysed but they do mention that it was a significant number.

Their study have identified six different categories of hate speech.

- Behavior (42.8%)
- Class (18%)
- Disability (10.7%)
- Ethnicity (10.7%)
- Religion (14.3%)
- Gender (3.5%)

The most frequent category was hate related to behavior, while categories like race and sexual orientation were not prominent in this specific study, which contrasts with previous research.

Moreover four types of impoliteness strategies were identified in the tweets:

- Bald on Record Impoliteness (32.8%)
- Positive Impoliteness (32.8%)
- Negative Impoliteness (22%)
- Sarcasm or Mock Politeness (12%)

Bald on Record and Positive Impoliteness were the most common strategies used by Twitter users in expressing hate.

Functions of Hate Expression:

- Expressing Unpleasant Feelings (33.3%)
- Entertaining the Audience (6.2%)
- Mocking the Figure and Expressing Disagreement (60.4%)

The dominant function was mocking Meghan Markle and expressing disagreement with her potential political ambitions.

All of the results show that the hate that is being directed towards Meghan is multi faceted having various categories. All of the background details we have mentioned in the beginning have contributed to the kind of hate she is receiving. The research highlights the complexity of online hate, which combines personal biases and impoliteness strategies.

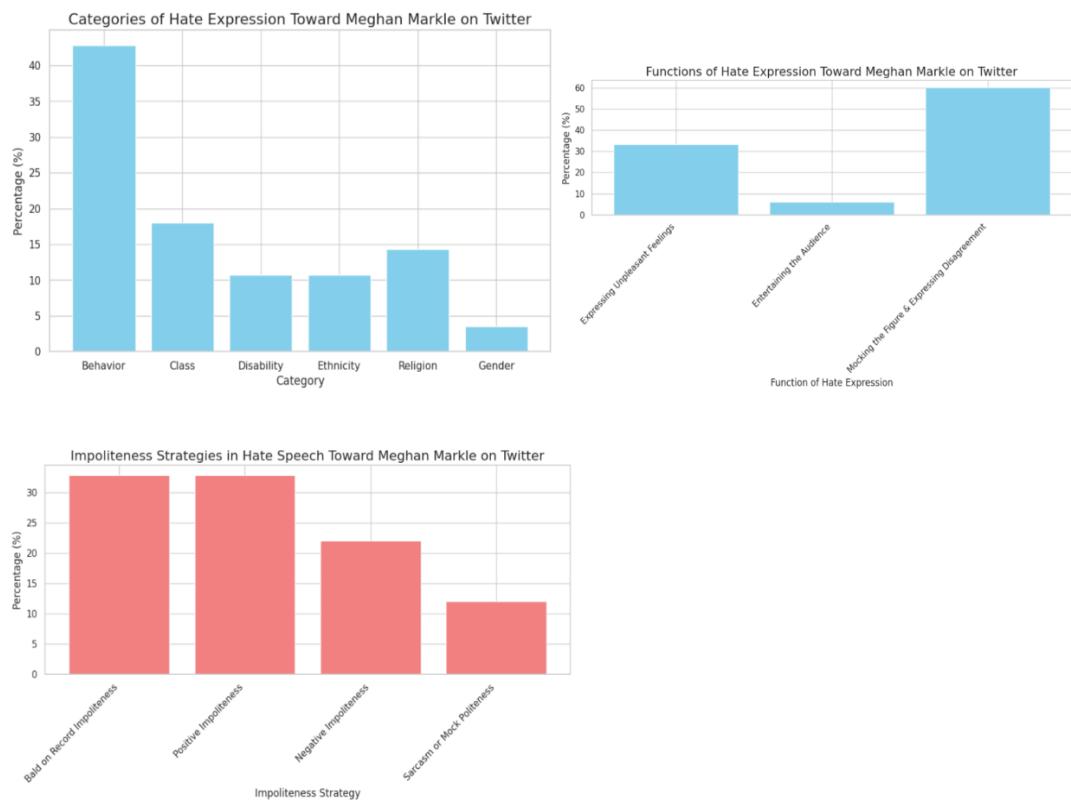


Figure 2.1: Hate analysis against Meghan

2.0.3 Hate speech against women dataset overview

We have also analysed a dataset containing around 20,000 tweets which mostly contained gender based hate speech tweets. We have read the paper titled "Automated Hate Speech Detection and the Problem of Offensive Language" from which we got the dataset. We also got the hate speech lexicons from Hatebase which is a repository for hate words their researchers found out from their analysis. So based on the most common words they had provided, we compared those words with out tweets dataset. The results for the most common words we got can we visualised here

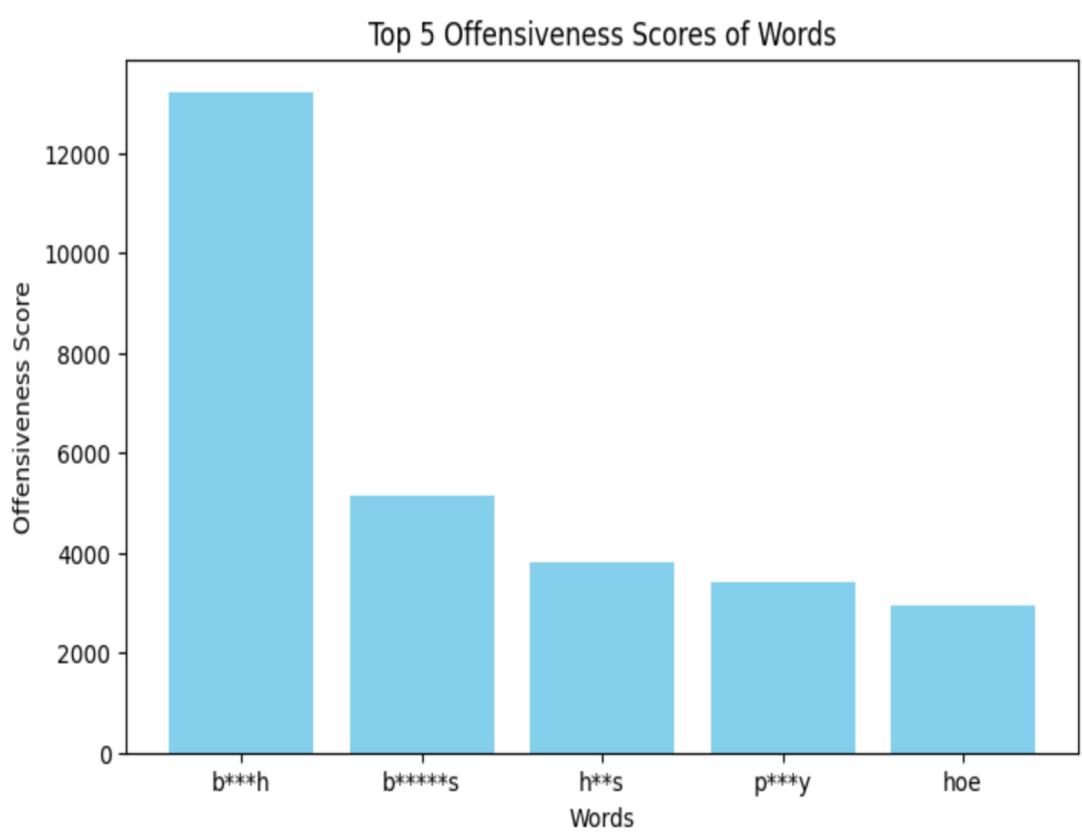


Figure 2.2: Hate analysis - Targeting women

Chapter 3

Case Studies - Caste Based

Caste-based hate speech is a form of discrimination that is very common in the social structure of societies, especially in regions like South Asia. Social media has become a hotspot of this kind of hate speech especially in the name of caste. India is a country which has been plagued with this caste ideology and people are still suffering because of this. Even if the society has come a long way in countries like India, remnants of these feelings of discrimination still runs deep within a lot of people's minds. The cases of Meena Kotwal and Gauri Lankesh were chosen because they exemplify the distinct difficulties of individuals who have been targeted with caste-based hate speech

3.0.1 Meena Kotwal (India)

Ms. Meena Kotwal is a human rights defender, journalist and founder of the online news channel and website 'The Mooknayak', which raises the issues of the Dalit minority.

Diving deep into why she is getting hate

She challenged Manusmriti by burning it because according to her it support gender discrimination and caste inequalities .Such acts considered offensive by conservative members of society. Being a Dalit woman journalist she try to raise the issues related to the caste and gender, so she face more hate and unfair treatment because of her work and her profession. She brings lighten the systematic inequalities by raising issues and the concern for Dalit and deprived communities

with the help of her channel “The Mooknayak”. Those who are in favor of maintaining the status quo may turn hostile against this movement. When Kotwal tweeted about the caste slur used against her, a group of dominant or the upper caste people comes in action to bully her. Many started hashtags which demands her arrest, even when she was a victim of caste-based harassment ,under the claim that she was “trying to criticise Hinduism

- A 2021 report by Equality Labs found that 25% of Dalits experienced caste-based slurs on social media, and 19% faced caste-based threats. Troll farms often coordinate attacks against Dalit activists, using WhatsApp, Twitter, and Facebook to amplify casteist abuse. For instance, Ms. Kotwal’s phone number was shared on WhatsApp groups after her act of burning the Manusmriti.
- A study by Amnesty International India found that 37% of online abuse targeting Indian activists was caste-based, and many platforms were slow to act against such hate speech.

3.0.2 Gauri Lankesh

Gauri Lankesh was a journalist who worked as an editor in Lankesh Patrike, a weekly built by her family. She was extremely vocal against caste based oppression and also was extremely against the right-wing hindutva ideology. She was facing a lot of online hate for her views and opinions. She was tragically murdered in front of her home by three men in September 2017. This event illustrates how hate in cyberspace turns into violent attacks in real life and highlights the pressing need for prevention.

This case provide a specific perspective into focus on the functioning of caste hate speech in cyberspace, its interface with other discrimination, and broader implications for social equality and justice.

Diving deep into why she is getting hate

Lankesh was against to right-wing Indian politics and Hindutva philosophy. She used to write regularly against laws and practices that supported religious hate and caste system and Support for the Underprivilege

Her writing highlighted issues that the Dalits and other marginalized segments face. She offended agencies that advocated for the older caste system.

Apart from casteist and ideological intolerance, Lankesh had faced sexist harassment as a woman journalist. Such threats were magnified in cyberspace, where internet trolls used abusive language to insult her and make violent threats. Campaigns against her were planned by online hate groups and troll farms. Memes that attacked her secular beliefs, videos that misrepresented her remarks, and many attempts to name her as "anti-national" all this increases the harassment in her life which she experienced.

Works We Referred

Equality Labs - Caste in the Digital Age

This report discusses how hate speech based on caste develops on social media and what the murder of Gauri Lankesh represents among the outcomes of unfiltered digital hate. Social Media Matters - Caste-Based Hate Speech in the Digital Age

An extensive analysis of online casteist abuse, through examples such as Lankesh's, to provide evidence of real-life.

UNESCO - Online Violence Against Women Journalists This international report clarifies the intersectionality of gender-based and ideological abuse experienced by women journalists such as Gauri Lankesh.

Amnesty International - Troll Patrol Report It focuses on the scale and nature of online abuse targeting activists, using examples from India, including the case of Gauri Lankesh.

Digital Witness Lab - How Misinformation Fueled the Gauri Lankesh Murder Examines how false narratives and cyber campaigns foster an atmosphere of hostility that leads to her murder.

Citizens for Justice and Peace (CJP) - Online Threats and Offline Consequences Explores how threats and campaigns against individuals like Lankesh turn into a reality in the offline space.

Chapter 4

Analysis

We have used a pretrained BERT model from huggingface that effectively calculates the probability of a sentence being classified as hate and a sentence being classified as a non hate text. BERT models are known for their classification capacity and it has performed pretty well for our task. The huggingface link for the model is Hate-speech-CNERG/dehatebert-mono-english and it is based in the paper "Deep Learning Models for Multilingual Hate Speech Detection" written by Sai Saketh Aluru et al. published in arxiv.

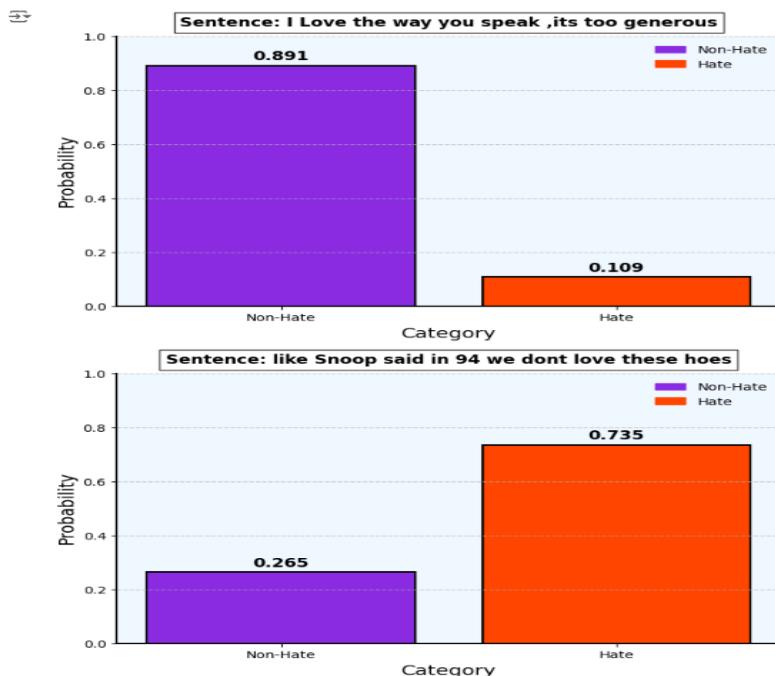


Figure 4.1: Examples of our model output

4.1 Acts against online hate speech

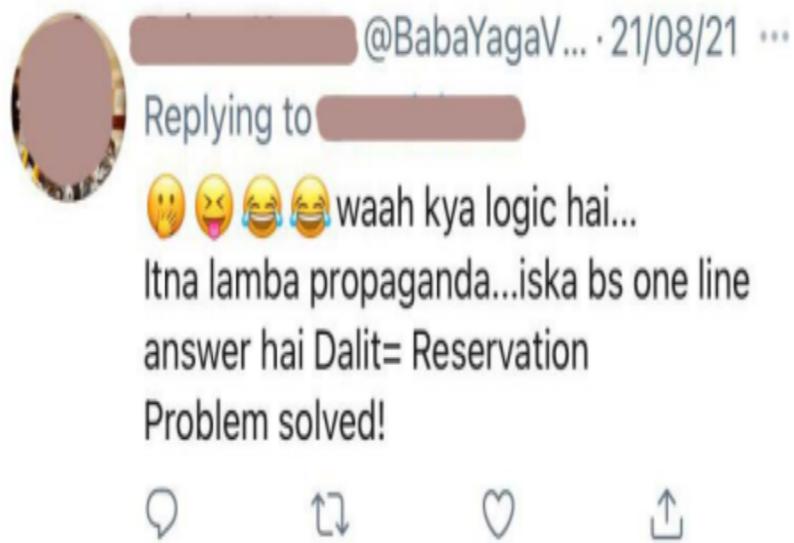
- Section 66A of the IT Act: This act target the offensive messages which used to be sent online over platform messages but was stopped down in 2015 for having uncertainty.
- Section 67 of the IT Act: Penalty over sending the obscene or dirty material online and which may lead to hate
- Section 69A of the IT Act: This act can block the public access to online content for sovereignty and security concern which can include hate speech.
- Indian Penal Code: Sections 153A, 295A, and 509 are applicable to online hate speech if it defamation, or gender-based insults.

4.2 Some supreme court jurisdictions

- Shreya Singhal v. Union of India (2015): The Court struck down Section 66A of the Information Technology Act, 2000, This act target the offensive messages which used to be sent online over platform messages. The provision was considered uncertain, leading to a possible misuse against free speech. Court want a clear legal definition to avoid restrictions on online expression.
- Pravasi Bhalai Sangathan v. Union of India (2014): In digital era court considered the challenges posed by hate speech, but restricting itself from issuing new guidelines, with supporting statement that existing laws like Sections 153A and 295A of the Indian Penal Code are sufficient.
- Amish Devgan v. Union of India (2020): during a television program a journalist makes a derogatory remarks against a religious figure. The Court looked at the nature of hate speech by the harm caused to persons and society. The court stated a fact that free speech does not involve expressions that establish hate or disturb the peace in the society.
- Patricia Mukhim v. State of Meghalaya (2021): Journalist Patricia Mukhim got some charges under Sections 153A and 505(1)(c) of the IPC for a Facebook post condemn an

attack on non-tribal youths. The Supreme Court quash the FIR, by stating that her post was a call for justice not for spreading the hate amongs the communities.

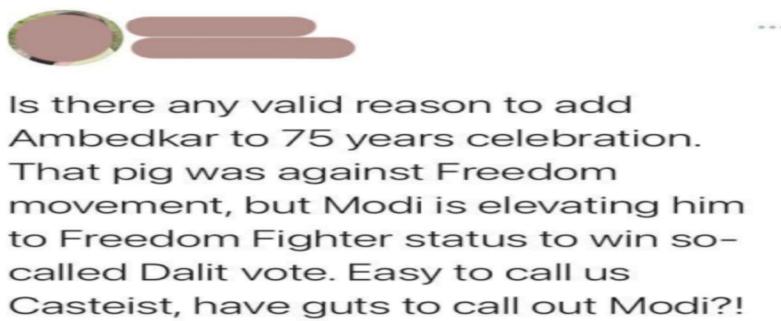
4.3 Some caste based hate speech tweets



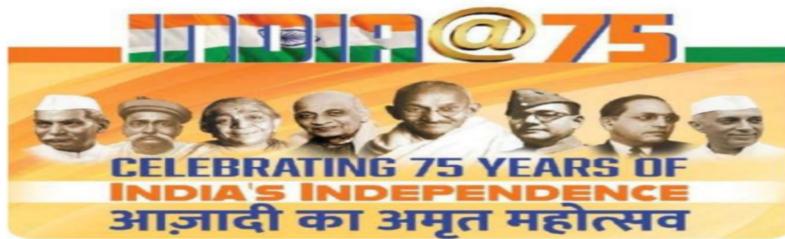
@BabaYagaV... · 21/08/21 ...

Replies to [redacted]

waah kya logic hai...
Itna lamba propaganda...iska bs one line
answer hai Dalit= Reservation
Problem solved!



Is there any valid reason to add Ambedkar to 75 years celebration. That pig was against Freedom movement, but Modi is elevating him to Freedom Fighter status to win so-called Dalit vote. Easy to call us Casteist, have guts to call out Modi?!



 · 21/08/21 ...

Replying to 

This thread is bullshit

Q 1 ↴ ❤ 6 ⬆

 ... · 21/08/21 ...

It's an isolated case with a dull-it.

Q ↴ ❤ 1 ⬆

 **T M Krishna** ✅ @tmkrishna · 11m ...
brahmins getting upset about the term brahmanism is like whites taking offense to the expression white supremacy!

Q 8 ↴ 62 ❤ 228 ⬆

 **Rupa Murthy** @rupamurthy1 ...

Replies to [@tmkrishna](#) and [@waglenikhil](#)

The term Brahmanism is not offensive. The adjectives preceding the term Brahmanism commonly used by Brahmin haters is often offensive.

If White = White supremacy,
then
TM Krishna = Hypocrite Racist ?



Radical Dalit

@believe_0369



They say ब्राह्मण, I hear नंगे भिखमंगे।

[Translate Tweet](#)

5:18 pm · 30 Apr 20 · Twitter for Android

4.4 Some gender based hate speech tweets



Suchita Tiwari @Suchita_1998 · 51m

2014 के पहले हमने दलितों का उत्पात, और बाबा साहेब, बाबा साहेब, का सफूगा नहीं सुना था, जब से यह भाजपा आई है चमटो ने अत्याचार की हदें पार कर दी और इस बाबा साहेब शब्द से मेरे कान बोध गए?

आज अगर बाबा साहेब जिंदा होते तो मैं उनके मु़ह में 50 ग्राम पेशाब जरूर डालती।



just harry

@henryscoisin

[Follow](#)

#MeghanMarkle is not one of your little ugly a** british aristocratic vapid do-nothing b*****es and y'all gonna learn.



just harry

@henryscoisin

[Follow](#)

The way that Harry and Meghan aren't part of that s***, messy, toxic institution anymore is most pleasing to me.



just harry

@henryscoisin

[Follow](#)

Meghan after having to deal the wedding dress, tiara, venue, flowers, food, music, her father's imbecility and this b**** throws a tantrum in the middle of her bridesmaids' dress fitting over

[←](#) Tweet



Rana Ayyub

@RanaAyyub

I hate India and Indians!

2:48 PM · 20 Apr 18

24 Retweets 569 Likes



Air India @airindiain · 2d

Replying to @RanaAyyub

Your one way flight to Pakistan is ready.



659

2369



Rana Ayyub

@RanaAyyub · [Follow](#)



26.4 thousand tweets, most are abusive, rape and death threats, calling me a terror sympathiser. Most tweets are by the Indian right wing & Saudi nationalists who are attacking me after i posted a tweet in solidarity with Yemen and calling out the Saudis. Hello [@TwitterIndia](#)



Jodie Ginsberg @jodieginsberg

So **@Twitter** is censoring journalist **@RanaAyyub**'s account in India at the behest of the authorities whom her work seeks to hold to account...Meanwhile the relentless **#onlineviolence** she faces largely goes unpunished.

Nikhil Dadhich @nikhildadhich 9h
एक कुतिया कुते की मौत क्या मरी सारे विल्से
एक सुर मे विलविला रहे हैं। 😢😢



Nikhil Dadhich
@nikhildadhich
Businessman|Garments Manufacturer|
R.K.Trading Co.|Hindu Nationalist| Honored
To Be Followed By PM Sh. @narendramodi Ji
Archive @nikhilsuntwal
♀ Surat, India ⚡ Born on 9 October
7,637 FOLLOWING 28.1K FOLLOWERS



Jagrati Shukla
@JagratiShukla29

Follow

So, Commy Gauri Lankesh has been
murdered mercilessly. Your deeds always
come back to haunt you, they say. Amen.
#Bengaluru
#GauriLankesh

9:43 PM - 5 Sep 2017

194 Retweets 355 Likes



← Tweet
Suchita Tiwari
@Suchita_1998

यह चमरिया @KotwalMeena की
का माल खा खा के बौरा गई है? शायद
यह भूल गई है कि इनकी दादी और मां
ने कितनी लेट्रीन कमर पर रखकर ढोई
है?

अरे दो वर्क का अच्छा खाना क्या खाने
लगी अपने आपको सर ऊंचा करके
बात करने लगी है?

Translate Tweet

Meena Kotwal (मीना कोटव...) 23h
मैं दोबे के साथ कह सकती हूँ कि दंदना कटारिया के
वक्त ही इन्हीं जोस्ट-शोर से आवाज उठते तो आज
मो शाम के साथ सामा ना होता लैकिन दुड़द कि
प्राप्ति जरूरी है इन्हीं क्षमा के लिए मैं जानती हूँ कि वे

Chapter 5

Conclusion

Hate speech and cyberbullying are two of the most persistent threats to digital and society harmony. Using the support of gender-based and caste-based hate speech, this research demonstrates how systemic discriminations is maintain which is increasing through social media platforms. The selected cases reflects the hate speech from organized campaigns against certain voices to threats that translate into offline violence.

There are different groups and each face unique problems, so we need working solutions that consider their specific needs. Governments, tech companies, and communities should come together to making better and stronger laws, hold online platforms responsible, and teach people how to stay safe online. We also need to improve how current cyber laws are enforced to keep up with new challenges like AI-created hate content which is nowday a new source of spreading the hates.

This statement focus on the importance of tackling hate speech early to prevent its harmful effects on the society . It tells the need to put balance between protecting free speech and shielding people from harm. And also, it calls for creating online spaces that encourages the marginalized voices and also reduce the effects of hate speech, and ensure a safer, more inclusive digital environment for everyone.

Chapter 6

Future Work

We can go through multiple cases, Currently we just focused on gender based and the caste based hate speech whereas there are many forms of hate speech on which we can work on.

We can create a Real-Time Monitoring Systems: By using some model we can create a system which develops the advanced real-time algorithms to detect and mitigate hate speech which may lead to faster and more effective solutions.

Collaborate with experts from sociology, law, and psychology to understand the socio-political factors contributing to hate speech and cyberbullying.

We can develop an artificial intelligence systems to find abuses such as sarcasm and coded language with less bias and more transparent system design.

Campaigns to update cyber laws against modern digital threats such as hate speech through AI.

Engage the users by making campaigns and spreading information regarding what impact hate speech might give, so we can appeal the correct practice on the web.

Chapter 7

Bibliography

- Equality Labs. (2021). "Caste in the Digital Age." Available at: <https://www.equalitylabs.org/>
- Amnesty International. (2021). "Troll Patrol Report." Available at: <https://www.amnesty.org/en/latest/research/2021/troll-patrol-report/>
- Article 14. "What 8.5 Million Tweets Targeting Rana Ayyub Tell Us About Online Violence." Available at: <https://article-14.com/post/what-8-5-million-tweets-targeting-rana-ayyub-tell-us-about-online-violence>
- New Yorker. "The Politics of Murder in Britain." Available at: <https://www.newyorker.com/news/news-desk/the-politics-of-murder-in-britain>
- IT for Change. (2022). "Profitable Provocations: How Social Media Platforms Enable Hate Speech." Available at: <https://itforchange.net/profitable-provocations-how-social-media-platforms-enable-hate-speech>
- ResearchGate. "The Representation of Meghan Markle in Social Media Posts." Available at: https://www.researchgate.net/publication/353121118/The_Representation_of_Meghan_Markle_in_Social_Media_Posts
- UNESCO. "Online Violence Against Women Journalists." Available at: <https://en.unesco.org/themes/safety-and-security/online-violence-against-women-journalists>
- Citizens for Justice and Peace (CJP). "Online Threats and Offline Consequences." Available at: <https://cjp.org.in/post/online-threats-and-offline-consequences>
- What 8.5 Million Tweets Targeting Rana Ayyub Tell Us About Online Violence The Failure To Stop It[<https://article-14.com/post/what-8-5-million-tweets-targeting-rana-ayyub-tell-us-about-online-violence-the-failure-to-stop-it>]

tell-us-about-online-violence-the-failure-to-stop-it-62d104dd20f4b]

- She experiences are largely facilitated by Twitter, where she has over 1.5 million followers.
The abuse routinely feature threats of death and rape.[<https://www.icfj.org/sites/default/files/2023-02/Rana%20AyyubCase%20StudyICFJ.pdf>]
- <https://www.trtworld.com/opinion/twitter-hate-86-percent-of-anti-muslim-content-comes-from-us-uk-india-12784562>
- The Representation of Meghan Markle in Facebook Posts: A Discourse Historical Approach (DHA), Iman M. Mahfouz et al., [https://www.researchgate.net/profile/Iman-Mahfouz/publication/328479774_The_Representation_of_Meghan_Markle_in_Facebook_Posts_A_Discourse_Historical_Approach_DHA.pdf]
- Hate Expression Found on Twitter as a Response to Meghan Markle, Auranissa Putri Riyadisty et al., [<https://e-journal.usd.ac.id/index.php/IJELS/article/view/4421/2579>]