CONTENT ANALYSIS OF MISOGYNISTIC TWEETS TARGETING INDIAN WOMEN CELEBRITIES

S. B. Patil

I/c Head, Department of Library & Information Science, Shivaji University, Maharashtra.

Corresponding Author - S. B. Patil

Email –sbp.lib@unishivaji.ac.in

ABSTRACT:

This paper aims to investigate the misogynistic comments made by the trollers on the twitter profiles of select Indian women celebrities. In the present context of the study, content analysis method is used to identify the misogynistic comments directed towards the twitter accounts of Indian women celebrities. 'Zipf's Law' is used for identifying the misogynistic words in user tweets. Academic Research Mozdeh software version 5.1.2.0.0 was used to collect and analyze the required data. The present study is restricted to the Indian women celebrities included in Forbes '2019 Celebrity 100' list. The results of a study revealed hot, pretty, fake, bad, sexy, forward, lost, hate, funny, and black as frequently occurred misogynistic words in the user tweets.

Keywords: Misogynistic, Tweets, Indian Women Celebrities, Content Analysis, Zipf's Law.

Introduction:

Nowadays the social media has been heavily used in almost all the fields of human endeavor. According to Statistica, in 2020 there are 518 million active social media users in India (Keelery, 2022). The twitter alone shares 24.4 million users and is continuously increasing (Statistica, 2022). It has become popular social media tool among the netizens. It exclusively facilitates various kinds of interaction, and that is used for extensive purposes, such as keeping in touch with friends, sharing multimedia, consuming news, engaging with public, celebrities and gathering real-time customer feedback. Twitter also provides a space for debate, humor, updates, products, gossip, and more besides. The celebrities often use twitter to engage with their fans, to reach out to the public, to promote their activities, agenda and get popular. Apart from these merits, the dark side of the twitter is alarming in the name of 'Online misogyny'. The female users, particularly the celebrities becoming the victims of online misogyny. They are subjected to more bullying, abuse, hateful language and threats than their male counterparts when online. An Amnesty

International IPSOS MORI poll in 2017 reported, "the women are much more likely to experience adverse psychological effects as a result of online harassment". As stated in the target 5.1 of United Nations fifth goal of Sustainable Development, it is important to end all forms of discrimination against all women and girls everywhere (United Nations, 2015).

With this background in mind, the researcher has attempted to investigate an increasingly well-known phenomenon: misogynistic comments, made by the trollers on the twitter profiles of select Indian women celebrities.

Definitions of Key Terms:

Misogynistic: According to Merriam Webster Dictionary (n.d.) "misogynistic means feeling, showing, or characterized by hatred of or prejudice against women."

Indian Women Celebrities: In the present context of the study, 'Indian women celebrities' refer to the famous female film actors, directors, singer/musicians, chefs, sports, and TV personalities mentioned in Forbes '2019 Celebrity 100' list.

Literature Review:

Firstly, Bartlett et al. (2014) analyzed 108,044 tweets that included the

word 'rape'. These tweets are collected using the public Twitter Application Programming Interface (API) from Twitter accounts based in the UK. The study found around 100 thousand instances of the word 'rape' used in UK-based Twitter accounts. Secondly, the researchers analyzed 131,711 tweets belonging to Twitter accounts based in the UK that used the words 'slut' and 'whore'. The study found a high proportion of colloquial /casual misogyny (35%) followed by abusive type of misogynistic words(20%) and generally misogynistic words(18%).

Konnelly (2015) examined a corpus of two hashtags – #Yes All Women and #He For She. The Systemic Functional Linguistics (SFL) approach is used to analyze the engagement, meanings, and functions manifested in the tweets. The quantitative results are interpreted by the researcher within the framework of Feminist Critical Discourse Analysis. The study noted that the engagement in both hashtags reveal that hashtags can be understood as semiotic linguistic resources and a facilitative device for asserting one's collective group identity and ideological affiliation.

Demos (2016) collected tweets containing 'slut' and/or 'whore' words through Twitter's API using a text-analysis software platform (Method52). In all, 1.46 million tweets were collected between 23rd

April and 15th May 2016. Out of the total dataset, the study found significant proportion (56%) of pornographic tweets.

Raju (2018) has analyzed the misogynistic comments received for specific posts of two Indian film artists Facebook pages. He used the Excel add ins called IQ Macros 2018 for word frequency analysis. To identify the misogynistic words in the comments, the researcher consulted the list of misogynistic words compiled and maintained by Thalikar. The study found that many comments which are made on the Facebook page of the actresses were highly derogatory. The study shows that women irrespective of their social, professional, economic status often subjected to online trolls and bullying for having their own opinions. The study highlighted a need of creating awareness on cyber-crime and other existing laws on this issue. It also suggested for establishing centres for victims of online harassment.

Schwarz (2020) provides an insight into feminist activism by presenting reactions to a poster campaign conducted in 2017.This campaign was offline in nature. The analysis of the troll comments shows that troller use existing stereotypes to insult women in a misogynist manner while at the same time they avoid dealing with feminism in a serious way.

To identify the trends in online misogyny before and during the COVID-19 pandemic, Dehingia, Lundgren, Dey, and Raj (2021) analyzed 19.8 million tweets from five South-Asian countries: Bangladesh, India, Nepal, Pakistan, and Sri Lanka. The study period was November 2019 to October 2020. The study noted that India has more twitter users than other four countries. In the present context of the study, Support Vector Machine (SVM) model was used for text classification. The results of a study revealed around 0.05% of the tweets in selected dataset contains misogynistic content, with occurrence before and during the pandemic at the rate of 0.046% and 0.057%, respectively. The highest peak in misogynistic tweets was observed in May for India. The study found that the spikes in online abuse correspond with events related to feminist movements or gender rights across the nations under the investigation.

Objectives of the Study:

The objectives of the study are stated as below:

1. To conduct content analysis of user tweets related to the Indian women celebrities.
2. To apply 'Zipf's Law' for identifying the misogynistic words in user tweets.

S. B. Patil

Research Methodology:

According to Columbia University Mailman School of Public Health (n.d.) "Content analysis is a research tool used to determine the presence of certain words, themes, or concepts within some given qualitative data i.e. text." In the present context of the study, content analysis method is used to identify the misogynistic comments directed towards the twitter accounts of Indian women celebrities.

Firstly, the researcher has consulted Forbes **'2019 Celebrity 100'** list for identifying the women celebrities in India. The list includes 31 female celebrities. Out of which, account of Kangana Ranaut is suspended by twitter and official account of Sonakshi Sinha, Malaika Arora, and Zoya Akhtar were not found. Hence, the 27 twitter accounts are used as the final sample of study. The details are given in Table1.

Table 1: Sample of the Study

Rank in the List	Name of the Female Celebrity	Profession	Twitter Account Name
8	Alia Bhatt	Actress	@aliaa08
10	Deepika Padukone	Actress	@deepikapadukone
14	Priyanka Chopra Jonas	Actress	@priyankachopra
21	Anushka Sharma	Actress	@AnushkaSharma
23	Katrina Kaif	Actress	@katrinakaif_4
28	Shraddha Kapoor	Actress	@ShraddhaKapoor
29	Neha Kakkar	Singer	@iAmNehaKakkar
32	Jacqueline Fernandez	Actress	@Asli_Jacqueline
38	KritiSanon	Actress	@kritisanon
41	Parineeti Chopra	Actress	@ParineetiChopra
42	Sonam Kapoor Ahuja	Actress	@sonamakapoor
43	DishaPatani	Actress	@DishPatani
48	Sunny Leone	Actress	@SunnyLeone
57	Madhuri Dixit Nene	Actress	@MadhuriDixit
59	SonakshiSinha	Actress	Official account not found
63	PV Sindhu	Badminton player	@Pvsindhu1
66	Sara Ali Khan	Actress	@SaraAliKhan
68	TaapseePannu	Actress	@taapsee
70	KanganaRanaut	Actress	Account suspended
75	ShilpaShettyKundra	Actress	@TheShilpaShetty
76	MalaikaArora	Actress	Official account not found
79	DivyankaTripathiDahiya	TV Personality	@Divyanka_T

S. B. Patil

81	SainaNehwal	Badminton player	@NSaina
82	Bharti Singh	Comedian	@bharti_lalli
83	ZoyaAkhtar	Filmmaker	Official account not found
87	Mary Kom	Boxer	@MangteC
88	Mithali Raj	Cricketer	@M_Raj03
89	RaveenaTandon	Actress	@TandonRaveena
90	SmritiMandhana	Cricketer	@mandhana_smriti
91	Harmanpreet Kaur	Cricketer	@ImHarmanpreet
93	KalkiKoechlin	Actress	@kalkikanmani

Secondly, the researcher has applied for academic research access through Twitter's developer platform (https://developer.twitter.com/en/products/twitter-api/academic-research). Personal and project related information was submitted in given prescribed application form. After the verification, twitter approved the researcher's application and provided access to utilize the academic level of the Twitter API. Twitter has provided the credentials like API Key and Bearer Token. Finally, these credentials are used to search and retrieve the required data from 'Academic Research Mozdeh' version 5.1.2.0.0. Mozdeh is a free Windows program for keyword, issue, time series, sentiment, gender and content analyses of (mainly) social media texts. It is developed by the Statistical Cyber metrics Research Group at the University of Wolver Hampton during the CREEN and Cyber Emotions EU projects (Statistical Cyber metrics Research Group, 2022).

Analysis and Interpretation of Data:
Distribution of Users Tweets:

The distribution of user tweets (N=20869) related to Indian women celebrities is shown in Table 2. Among the total output, highest user tweets (1487) are related to Sonam Kapoor Ahuja, followed by tweets related to Parineeti Chopra (1291) and Priyanka Chopra Jonas (1244). Only three tweets are concerned to Katrina Kaif. *This does not mean that she is not popular celebrity. This could be due to a number of factors, such as her focus on her film career over social media, or the fact that her fans are more likely to engage with her content on other platforms, such as Instagram.*

Table 2: Distribution of Users Tweets

Name of the Celebrity	Count of User Tweets	Rank
Sonam Kapoor Ahuja	1487	1
Parineeti Chopra	1291	2
Priyanka Chopra Jonas	1244	3
TaapseePannu	1139	4
Deepika Padukone	1073	5
Sunny Leone	1072	6
DivyankaTripathi Dahiya	1066	7
Madhuri Dixit Nene	1037	8
Smriti Mandhana	955	9
Mithali Raj	898	10
Shraddha Kapoor	866	11
Alia Bhatt	847	12
SainaNehwal	812	13
Anushka Sharma	773	14
DishaPatani	772	15
RaveenaTandon	764	16
Shilpa Shetty Kundra	759	17
PV Sindhu	713	18
Jacqueline Fernandez	630	19
Harmanpreet Kaur	621	20
Sara Ali Khan	610	21
Neha Kakkar	476	22
Bharti Singh	449	23
KritiSanon	212	24
Mary Kom	151	25
KalkiKoechlin	149	26
Katrina Kaif	3	27
Total =	**20869**	

Word Frequency Analysis of Misogynistic Words in User Tweets:

Word Frequency Analysis is conducted for identifying Misogynistic nature of user tweets related to select Indian women celebrities. To identify the misogynistic words, the researcher have referred the list of 122 misogynistic words compiled by Thalikar which is available in his personal blog at:http://sacraparental.com/2016/05/14/everyday-misogyny-122-subtly-sexist-words-women/ Researcher also noted the misogynistic words identified by Dr. Vasantha Raju N. (2018) in his conference paper titled as *"Is Indian*

Social Media Space Misogynistic? A Study of Facebook Comments on Two Indian Female Film Artists Observations on Portraying Women in Cinema." The researcher has also manually checked the vocabulary downloaded from Mozdeh software to identify the misogynistic words used in the user tweets. The identified misogynistic words are shown in Table 3 and Figure 1. The misogynistic words in green colour are identified through Thalikar's list while the words in red colour are available in conference paper of Dr. Vasantha Raju N. (2018). The researcher had manually identified remaining words. The original data source i.e. vocabulary downloaded from Mozdeh software contain stop words, twitter account names, hashtags, abbreviations etc. After manual screening the researcher excluded the same by applying necessary filters in excel. Top forty misogynistic words (N ≥ 17)with their frequencies and ranks are presented in Table 3. It is significant to mention that the misogynistic word 'hot' (130) is ranked top in the list followed by 'pretty' (106), fake (94), bad (92) and sexy (80).

Table 3: Word Frequency Analysis of Misogynistic Words

Misogynistic Words	Frequency	Rank
Hot	130	1
Pretty	106	2
Fake	94	3
Bad	92	4
Sexy	80	5
Forward	64	6
Lost	53	7
Hate	50	8
Funny	47	9
Black	45	10
Paid	41	11
Private	41	11
Killed	38	13
Worst	36	14
Dumb	33	15
Emotional	33	15
Princess	32	17
Damn	31	18
Attack	30	19
Boycott	27	20
Crush	27	20
Crying	27	20

S. B. Patil

Easy	27	20
Low	27	20
Warm	27	20
Hell	26	26
Common	25	27
Stupid	24	28
Feminist	23	29
Hottie	23	29
Racist	23	29
Shameless	23	29
Cry	22	33
Loose	22	33
Tear	22	33
Diva	19	36
Ban	18	37
Drug	18	37
Cheap	17	39
Dump	17	39
❖ N ≥ 17 (N = Frequency of Misogynic Words)		

Figure 1: Word Cloud of Misogynistic Words

Application of Zipf's Law:

Zipf's Law is concerned about the words in a text. According to the Zipf's law, the frequency of occurrence of a word in a text is inversely proportional to the rank of that word in that text (Savanur & Sangam, 2019).

It is mathematically represented as:

F ∝ 1/R

Where, F = Frequency of occurrence of a word in a text

R = Rank

In Table 3, the misogynistic words with its frequencies and ranks are given. Following scatter graph show that **as frequency decreases, the numerical value of rank increases.**
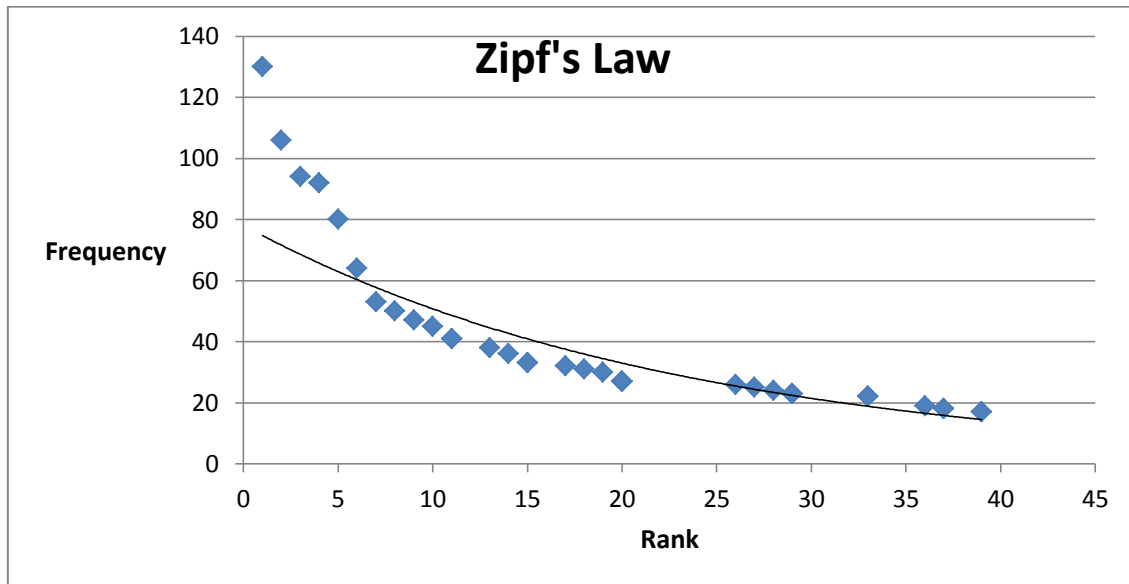
Graph 1: Zipf's Plot

Major Findings of the Study

- *In the present study, celebrities from the Indian film industry dominate the sample, followed by celebrities from the sports field.*

- It was observed that the sample comprises single comedian (Bharti Singh), a film director (Zoya Akhtar), a singer (Neha Kakkar) and a television actress (Divyanka Tripathi Dahiya).

- In the Forbes 2019 Celebrity 100 list, Alia Bhatt ranked 8th, followed by Deepika Padukone at 10th and Priyanka Chopra Jonas at 14th position.

- Smriti Mandhana, Harmanpreet Kaur, and Kalki Koechlin had received the lowest rank in the Forbes '2019 Celebrity 100' list.

- The highest user tweets 1487(7.12%) were related to Sonam Kapoor Ahuja.

- It was noteworthy to find out that the famous actress Katrina Kaif had received lowest user tweets (3).

- The five celebrities named Sonam Kapoor Ahuja, Parineeti Chopra, Priyanka Chopra Jonas, Taapsee Pannu, and Deepika Padukone share 6236 (29.89%) user tweets.

- By considering the quantity of users' tweets, except Mithali Raj no other sports celebrity is included in top ten.

- For the time span of 1/2/2022 to 28/2/2022, the present study has identified 40 misogynistic words used in the user tweets concerned to Indian women celebrities.

- It was noticed that the top five misogynistic words (hot, pretty, fake, bad, and sexy) together shared 502 (32.17%) frequencies among the total output.

- The misogynistic words viz. ban, drug, cheap, and dump ranked low in the list.

- The results of a study revealed hot, pretty, fake, bad, sexy, forward, lost, hate, funny, and black as frequently occurred misogynistic words in the user tweets.

Conclusions:

The issue of misogyny identification in social media is complicated and has many aspects. Twitter has long been criticized for the seemingly hands-off approach to online harassment and abuse when defending postings of harmful content (Wagner, 2017). After facing the much more criticism Twitter started working with its Trust and Safety Council, taking a more active stance in

S. B. Patil

harmful content conversations (Flynn, 2017). The council's areas of focus include Online Safety and Harassment, Human and Digital Rights, Suicide Prevention and Mental Health, Child Sexual Exploitation, and Dehumanization (Twitter, n.d.).

There have been several instances of high-profile cases of women celebrities, sportswomen, journalists, and politicians who have experienced misogynistic violence and endured bullying insults on Twitter. In the present context of the study, the researcher attempted to investigate an increasingly well-known phenomenon: misogynistic comments, made by the trollers on the twitter profiles of select Indian women celebrities. Forbes '2019 Celebrity 100' list was used as the source of the data collection. Majority of the women celebrities included in this list belongs film industry. It does not include celebrities from fields like media, journalism, social work, politics etc. The study has identified 40 prominent misogynistic words used in the user tweets. The findings of a study revealed hot, pretty, fake, bad, sexy, forward, lost, hate, funny, and black as frequently occurred misogynistic words in the user tweets concerned to Indian women celebrities. In order to identify the misogynistic words used in the communication, a standard lexicon or a

database is required. The study highlighted a need to develop the machine – learning model to detect the misogynistic tweets. To overcome the problem of online misogyny, the government/ non-government organizations, higher educational institutions, and media firms etc. should organize gender sensitization programmes for the masses. Finally, it is suggested that the user should avoid conventional policing of the celebrities.

References:

1. Bartlett, J., Norrie, R., Patel, S., Rumpel, R., Wibberley, S. (2014). Misogyny on Twitter. Demos. Retrieved from http://www.demos.co.uk/files /MISOGYNY _ON_ TWITTER.pdf

2. Dehingia, N., Lundgren, R., Dey, A. K., & Raj, A. Trends in online misogyny before and during the COVID-19 pandemic: Analysis of Twitter data from five South-Asian countries. Available at:https://data2x.org/wp-content/uploads/2021/03/UCSD-Brief-3_BigDataGend erCOVID19SouthAsianMisogyny. pdf

3. Demos (2016). *The use of misogynistic terms on twitter*. Retrieved from https://www.demos.co.uk/w pcontent/uploads/2016/05/Misogy nyonline.pdf.

4. Flynn, K. (2017, December 18). The "Twitter purge" Nazi reckoning has begun: Here are the rules. *Mashable*. Retrieved from https://mashable.com/2017/ 12/18/twitter-purge-neo-nazi-reckoning-newrules-hate-speech/

5. ForbesIndia.com (2022).2019 Celebrity 100. Retrieved 2 March 2022, from https:// www.forbesindia.com/lists/2019-celebrity-100/1819/1

6. http://sacraparental.com/2016/05/1 4/everyday-misogyny-122-subtly-sexist-words-women/ Accessed on 2/3/2022

7. https://worditout.com/word-cloud/create Accessed on 2/3/2022

8. https://www.amnesty.org/en/latest/ news/2017/11/amnesty-reveals-alarming-impact-of-onlin e-abuse-against-women/ Accessed on 8 March 2022.

9. Keelery, S. (2022). Number of social network users India 2015-2040. Retrieved 8 March 2022, from https://www.statista.com/statistics/ 278407/number-of-social-network-users-in-india/#:~: text=With%20the%20ease%20of% 20internet,media%20platforms%20 as%20of%2 02020

10. Konnelly, A. (2015). Activism: Identity, affiliation, and political discourse - making on Twitter. *The Arbutus Review*, *6*(1), 1-16. doi:10.18357/ar.konnellya. 612015

S. B. Patil

11. Merriam-Webster.(n.d.).Misogynistic.In *Merriam-Webster.com dictionary*. Retrieved March 7, 2022, from https://www.merriam-webster.com/dictionary/misogynistic

12. Raju,V. (2018). Is Indian social media space misogynistic? A study of Facebook comments on two Indian female film artists' observations on portraying women in cinema. In: *National Conference on Social Media for Digital and Social Inclusion: Cultural, Technological and Economic Implications*. Mysore: CSSEIP, University of Mysore, Mysore.

13. Savanur, S., & Sangam, S. (2019). Zipf's Law in Governance. In *8th National Conference on Scholarly Communication and Scientometrics* (pp.1-16). Mumbai: ACTREC. Retrieved from https://www.researchgate.net/publication/337621758_Zipf's_Law_in_Governance

14. Schwarz, S. (2020). I wasn't chastised properly: On trolls and misogyny. In S. Polak& D. Trottier (Eds.), *Violence and Trolling on Social Media: History, Affect, and Effects of Online Vitriol* (pp. 217–232). Amsterdam University Press.

https://doi.org/10.2307/j.ctv1b0fvrn.14

15. Statistica (2022). Leading countries based on number of Twitter users as of October 2021*(in millions).* **Retrieved**8 March 2022, fromhttps://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/

16. Statistical Cyber metrics Research Group (2022). About Mozdeh. Retrieved 2 March 2022, from http://mozdeh.wlv.ac.uk/

17. Twitter. (2022). About Twitter: Trust and Safety Council. Retrieved 10 March 2022, from https://about.twitter.com/en/our-priorities/healthy-conversations/trust-and-safety-council

18. United Nations. (2015). Sustainable Development Goal 5 | Achieve gender equality and empower all women and girls. Retrieved 8 March 2022, from https://sdgs.un. org/goals/goal5

19. Wagner, K. (2017, April 10). Twitter fights to protect anonymous users more often than you'd think. *Recode*. Retrieved fromhttps://www.recode.net/2017/4/10/ 15244754/twitter-lawsuit-government anonymous-users