

Exploring Hindu Nationalism Discourse on Reddit through Text Analysis

This project uses sentiment analysis and topic modeling methods to visualize similarities and differences in sentiments and topics between two subreddit communities differing by size and level of extremism.



Author:
Ananya Pujary

Affiliation:
Data Analytics and Computational
Social Science Program, University of
Massachusetts Amherst

Motivation

- Increasing Hindu nationalistic sentiments in India around the 70s (Siyech, 2021)
- Growing extremism on social media translating into the real world as hate towards religious minorities (Saraswat, 2022)
- Lack of literature on the Reddit, the third most used SNS platform in India (Saraswat, 2022)

Research Questions

- What are the dominant themes discussed in prominent Indian subreddits, and as related to the Hindutva ideology?
- What are the dominant sentiments of users on Indian subreddits, and as related to the Hindutva ideology?
- Do some Indian subreddits engage in more extreme discussions of the Hindutva ideology than others?

Data and Methodology

Data

- r/India (2M members; 747 rows) & r/HindutvaRises (6.4K members; 924 rows)
- Collected on December 8th, 2023 using RedditExtractoR
- Time period of 1 month, sorted by newest posts
- 8 columns (6 character, 2 numeric) and each row is a post
- Columns: date published, title, text, title_text, timestamp, number of comments, subreddit name, url
- Data cleaning: Tokenization (removed punctuation, number, symbols, urls, English and Hindi stopwords), corpus creation

Methodology

- Sentiment Analysis:** with the NRC dictionary including emotions: anger, anticipation, disgust, fear, joy, negative, positive, sadness, surprise, trust. Used to understand user emotions related to topics; polarity score from -1 to 1 calculated by subtracting negative from positive sentiment scores
- Structural Topic Modeling:** to extract four overarching topics from the subreddit datasets split by polarity scores (<0: negative polarity, = 0: neutral polarity, >0: positive polarity), generating three models for each subreddit



Analysis

Sentiment Analysis

Polarity Score Distribution

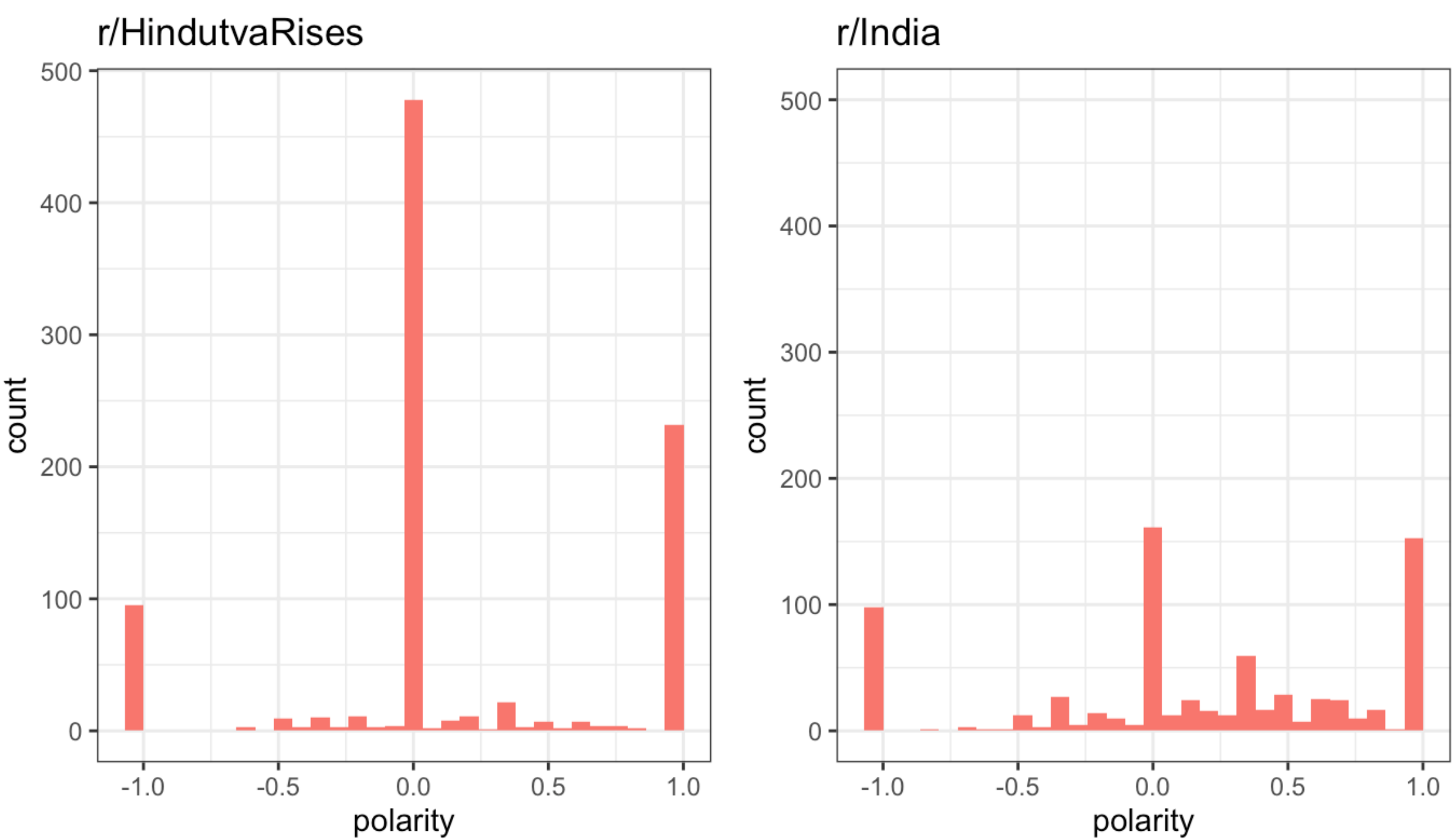


Figure 1. Distribution of polarity scores in each subreddit

Structural Topic Modeling



Figure 2. Topic models split by polarity score for both subreddits

Findings and Conclusion

- Partial support to the hypotheses.
- Smaller subreddit not as extreme as expected since most posts classified as neutral. Positively coded posts talk about Hindu rituals and extremist ideas, while negative ones discuss other religions.
- More general topics and positive polarity in r/India.
- Limitations**
 - Cross-sectional data may not be very representative, and was not collected at a particularly relevant time period (e.g. elections season).
 - Not very generalizable since inferences are drawn from two subreddits only.
 - Tokenization not very effective.
- Future Directions**
 - Compare more subreddits over a relevant time period (e.g. during election season).
 - Experiment with more K-values to see how topics in STM change.

References

Chakraborty, S., Pal, J., Chandra, P., & Romero, D. M. (2018a). Political Tweets and Mainstream News Impact in India: A Mixed Methods Investigation into Political Outreach. Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies. <https://doi.org/10.1145/3209811.3209825>.

Manuvie, R., & Chatterjee, S. (2023). Automated Sentiment and Hate Speech Analysis of Facebook Data by Employing Multilingual Transformer Models. *Computation and Language*, 1–11. <https://doi.org/https://doi.org/10.48550/arXiv.2301.13668>.

M. N., P. (2023, February 28). India's YouTube Vigilante is Wanted for Murder. *Wired*. <https://www.wired.com/story/indias-youtube-vigilante-monu-manesar-murder/>.

Saraswat, R. (2022, January 25). Out of Sight: How Reddit became a "safe space" for hate in India. *TheQuint*. <https://www.thequint.com/leon/web-culture/out-of-sight-how-reddit-became-a-safe-space-for-hate>.

Siyech, M. S. (2021). An Introduction to Right-Wing Extremism in India. *New England Journal of Public Policy*, 33(2). <https://scholarworks.umb.edu/cgi/viewcontent.cgi?article=1809&context=nejpp>.