

Polarization and its Life on Social Media: A Case Study on Sabarimala and Demonetisation

Ashutosh Ranjan

IIIT Hyderabad

ashutosh.ranjan@research.iiit.ac.in

Dipti Sharma

IIIT Hyderabad

dipti@iiit.ac.in

Radhika Krishnan

IIIT Hyderabad

radhika.krishnan@iiit.ac.in

Abstract

This paper is an attempt to study polarisation on social media data. We focus on two hugely controversial and talked about events in the Indian diaspora, namely 1) the Sabarimala Temple (located in Kerala, India) incident which became a nationwide controversy when two women under the age of 50 secretly entered the temple breaking a long standing temple rule that disallowed women of menstruating age (10-50) to enter the temple and 2) the Indian government's move to demonetise all existing 500 and 1000 denomination banknotes, comprising of 86% of the currency in circulation, in November 2016. We gather tweets around these two events in various time periods, preprocess and annotate them with their sentiment polarity and emotional category, and analyse trends to help us understand changing polarity over time around controversial events. The tweets collected are in English, Hindi and code-mixed Hindi-English. Apart from the analysis on the annotated data, we also present the twitter [data](#) comprising a total of around 1.5 million tweets.

1 Introduction

Social media, over the past few years, has become the fastest growing medium for expressing opinions for billions of people worldwide. People have always had opinions in history, some extreme and some moderate. What has changed today, is the way people can make their opinions reach masses (read millions) of people in seconds to minutes to hours, depending on the popularity of the person and the general controversiality (for lack of a better word) of the opinion. Earlier people could influence a bunch of people in their vicinity and sometimes, very rarely after going through multiple filters of sanity, they would be able to bring their idea to the world map. Today any random person is able to express their ideas to millions of

people, and thereby has the potential to influence masses much faster.

India is no different. Easy access to internet has brought millions of Indians to Twitter, where they express their opinions on topics that interest, affect them. Leaders use social media to make their opinions reach the masses as well. In a recent survey by Pew Research Center it was revealed that a significant percentage of people get their news from social media ([Pew Research Center, 2018](#)). This activity on twitter consequently helps people form opinions or change their already existing opinions on different topics. This paper attempts to study this change in opinion of the masses of people by observing and analysing their tweets on social media. We study the life of polarization on social media in the Indian context.

We decide to focus our efforts on two pressing, controversial and divisive issues - 1) Sabarimala 2) Demonetisation. We study how the polarity of opinions of people change with respect to these issues over time. The complete context of the two events is described next.

1.1 Sabarimala

The Sabarimala temple is a temple complex located at Sabarimala inside the Periyar Tiger Reserve in Pathanamthitta district, Kerala, India. It is the site of the one of the largest annual pilgrimages in the world with an estimate of between 17 million and 50 million devotees visiting every year.

According to the Memoir of the Survey of the Travancore and Cochin States, published in two volumes by the Madras government in the 19th century, women of menstruating age were denied entry into the Sabarimala temple two centuries ago, as all sexual intercourse in that vicinity is averse to the celibate deity Lord Ayyappa. In response to a Public Interest Litigation filed in 1991, the Kerala High Court had judged that the restriction of entry of

women ages 10-50 to the temple was in accordance with the usage prevalent from time immemorial, and it directed the concerned board to uphold the customary traditions of the temple. However, on 28 September 2018, the Supreme Court of India overturned the restriction on the entry of women, declaring it unconstitutional and discriminatory. On 2nd January 2019, two women under the age of 50 entered the shrine for the first time since the Supreme Court verdict, after attempts of many others failed due to protests by devotees.

The event of two women entering caused huge uproar on social media. One side of the debate argued for the conservation of tradition, and the other side argued for gender equality.

1.2 Demonetisation

In November 2016, the Government of India announced the demonetization of all 500 and 1,000 banknotes of the Mahatma Gandhi Series. It also announced the issuance of new 500 and 2,000 banknotes in exchange for the demonetised banknotes. The Prime minister of India Narendra Modi claimed that the action would curtail the shadow economy and reduce the use of illicit and counterfeit cash to fund illegal activity and terrorism.

The announcement of demonetisation was followed by prolonged cash shortages in the weeks that followed, which created significant disruption throughout the economy. People seeking to exchange their banknotes had to stand in lengthy queues, and several deaths were linked to the rush to exchange cash.

The issue came to relevance again in August 2018 when the Reserve Bank of India released a report, according to which approximately 99.3% of the demonetised banknotes, or 15.30 lakh crore (15.3 trillion) of the 15.41 lakh crore that had been demonetised, were deposited with the banking system, leading analysts to state that the effort had failed to remove black money from the economy. By many economists, the move is blamed for reducing the country's industrial production and slowing down its GDP growth rate.

Initially, the move received support from several bankers as well as from some international commentators. By others it was widely criticised as poorly planned and unfair, and was met with protests, litigation, and strikes against the government in several places across India. The stated failure of demonetisation and subsequent downfall

of India's GDP growth rate was also a major topic of interest during the 2019 general elections and cited by many opposition parties as a failure of the presiding government.

These two incidents were divisive and evoked different reactions in people. Both the events have periods where they are focus of attention, and with time the attention dies out. The change in polarity from before the event through the time of the event and then to the eventual end gives us some interesting insights. Here in this study we analyse the change in reactions at the mass level through different time periods, and try to answer the following questions:

- Is society inherently polarised?
- How does a divisive event affect the polarity?

2 Literature Review

Stephen Hawkins and Dixon (2018) published a study entitled "The Hidden tribes of USA". The study, through a survey, categorises the people of USA in 7 major political categories from extreme left to extreme right. They classify four of these seven non-extreme classes as the exhausted majority. The concept of exhausted majority implies that the majority of people lying in the middle of the polarity spectrum are tired from the extremist and tribal stands of the extreme left and extreme right people. They want them to strike a balance in their views and agree on a compromise instead of being engaged in political tribalism. In our work, we see a huge proportion of people keeping neutral opinions which aligns with the above mentioned work.

3 Data Collection

To analyse how people reacted to these two events, we scraped tweets from twitter using keyword search with the help of the api created by Jefferson Henrique (2018). The tweets scraped are collected according to the timeperiods of interest, as mentioned below.

3.1 Sabarimala

We scraped tweets with keyword "sabarimala" for five 3-month periods (Table 1), spread across two years, to study the changing sentiments of people with respect to the event. The 3-month periods are:

- Oct-Dec 2017, a supposedly non controversial period, long before Supreme Court's decision

Time Period	No. of Tweets
Oct - Dec, 2017	6,445
Mar - May, 2018	770
Oct - Dec, 2018	15,373
Jan - Mar, 2019	74,498
May - July, 2019	12,115

Table 1: Sabarimala Twitter Data

in September 2018. This period also coincides with the days of Mandalapooja festival (around 15 November to 26 December) which is when Sabarimala is the busiest ([Dainik Bhaskar, 2011](#)). Hence this period saw a lot of activity on Twitter.

- Mar-May 2018, another supposedly non controversial period. The activity on Twitter was a lot lower during this time with respect to the previous period.
- Oct-Dec 2018, right after Supreme Court Judgement (on 29th September, 2018), but before the date (2nd Jan 2019) the two women entered the temple for the very first time.
- Jan-Mar 2019, right after the two women entered the temple.
- May-July 2019, five to eight months after the controversy.

3.2 Demonetisation

We scraped tweets with keyword "demonetisation" for five 3-month periods (Table 2), spread across two years, to study the changing sentiments of people with respect to the event. The 3-month periods are:

- 8th Nov 2016 - 7th Feb 2017, right after the demonetisation was announced on 8th Nov 2016.
- May-July 2018, 5-8 months after the event.
- 9th Nov 2017 - 8th Feb 2018, 1 year after the announcement.
- 29th Aug 2018 - 29th Nov 2018, right after Reserve Bank of India's review that reported that 99.3% of the cash has been deposited back, effectively stating the failure of demonetisation.

Time Period	No. of Tweets
9th Nov'16- 8th Feb'17	711373
May - July, 2017	41461
9th Nov'17 - 8th Feb'18	58464
29th Aug - 29th Nov, 2018	45108
Feb - Apr, 2019	38534

Table 2: Demonetisation Twitter Data

- Feb 2019 - Apr 2019, during general election campaigning days just before the elections in Apr and May 2019.

4 Data Preprocessing

To make preprocessing easier, the tweets were separated into three groups based on the language of content, - 1) English 2) Hindi 3) Hindi-English code-mixed. Tweets containing words from any other language were removed. This separation was done with the help of the language identifier by [Bhat et al. \(2015\)](#).

All three groups of tweets were preprocessed separately as follows:

4.1 Preprocessing English Tweets

- Remove special characters, username mentions (beginning with @) and urls.

4.2 Preprocessing Hindi Tweets

- Remove all special characters, username mentions (beginning with @), and urls.
- Translate the whole tweet to English using python freeware translate.

4.3 Preprocessing Code-Mixed Hinglish Tweets

These tweets were in roman script, with Hindi words written in Roman script as well.

- Use language identifier ([Bhat et al., 2015](#)). to find all the Hindi words.
- Use transliterator ([Bhat et al., 2015](#)) to transliterate the Hindi words from Roman script to Devanagri script.
- Translate Hindi words in Devanagri script to English using python freeware translate.

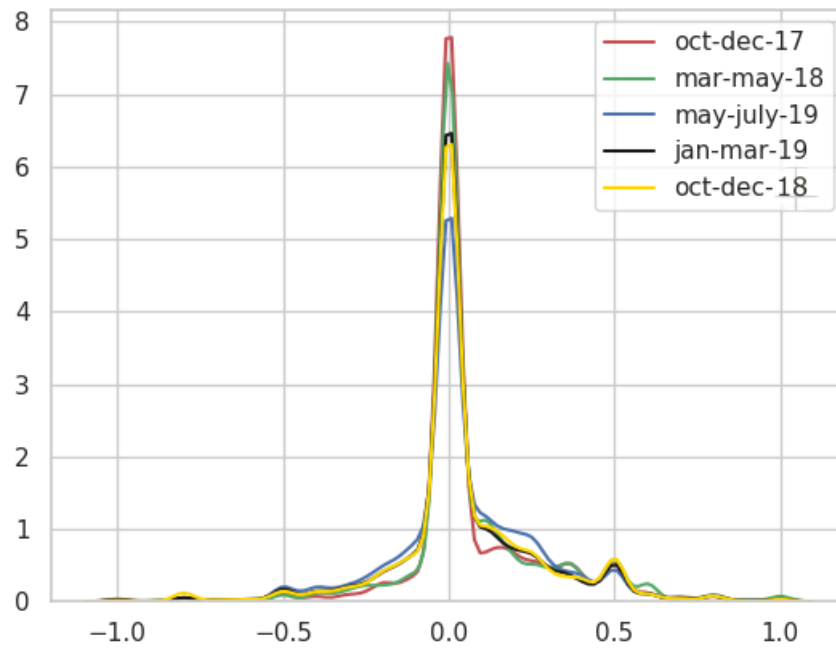


Figure 1: Sabarimala Gaussian Curves measuring polarity percentage

5 Methodology

5.1 Sentiment Analysis

We use the python library TextBlob (Loria, 2018) which uses naive bayes to find the sentiment of each of the tweets on a scale of -1 to 1. We say that the polarity of a tweet increases as we move away from 0 to either side.

5.2 Emotional change through the different periods

We used the NRC Emotion Lexicon (Mohammad and Turney, 2013) to find the emotions expressed and their change through the different periods (Figures 2 and 5). The NRC Emotion Lexicon is a labelled dataset that associates words with emotions, which can be of eight types - anger, sadness, disgust, joy, surprise, trust, anticipation and fear. As a bonus, it also associates each word with either positive or negative connotation.

6 Experiments and Results

6.1 Sabarimala

- Overall negative sentiment increases after the Jan 1 incident. (Figure 1)
- Opinions polarized due to the event (on both

sides), in comparison to a non-event (non-controversial) normal state. The graph shows a significant decrease in the neutral population in the subsequent periods. (Figure 1)

- 5-8 months after the polarizing event, the opinions are still polarised and even more so than they were at the time of the event. (Figure 1)
- The number of tweets increase drastically during the period in which the event took place wrt. previous non controversial periods. It also decreases drastically in the period 5-8 months after the event. (Figure 3)
- Grouped emotion bars show increased anger, sadness, negativity. (Figure 2)

6.2 Demonetisation

- The tweets in the first period, i.e. right after when the event happens, are huge in number and decrease drastically in the subsequent periods. (Figure 6)
- The graph shows a significant decrease in the neutral population in the subsequent periods. (Figure 4)

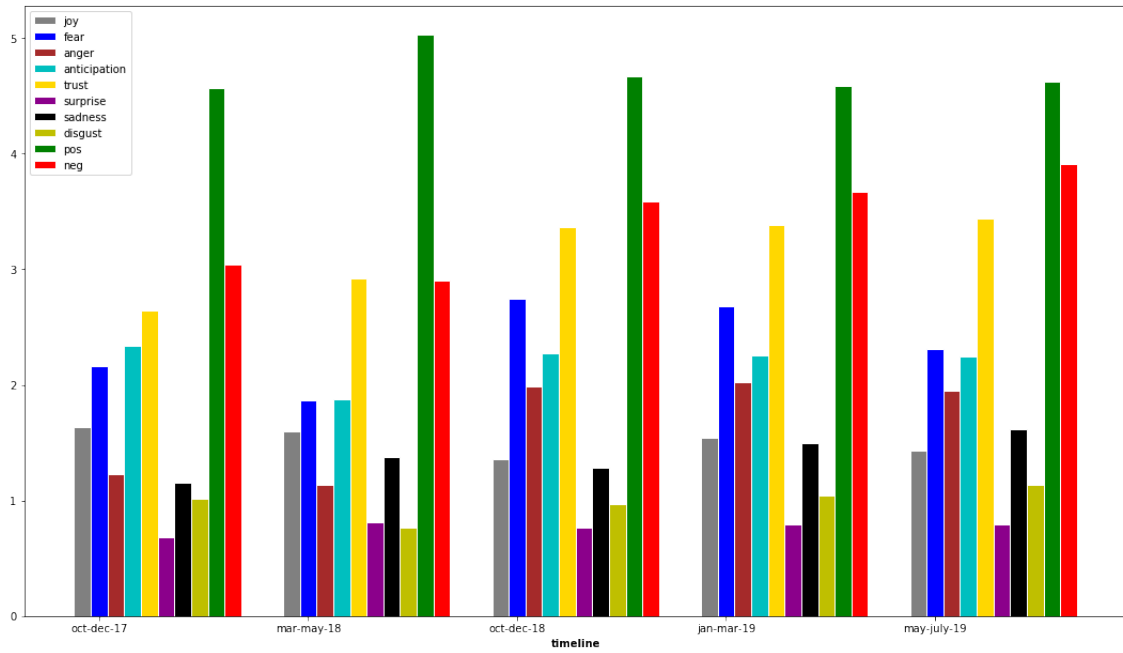


Figure 2: Emotion category Sabarimala

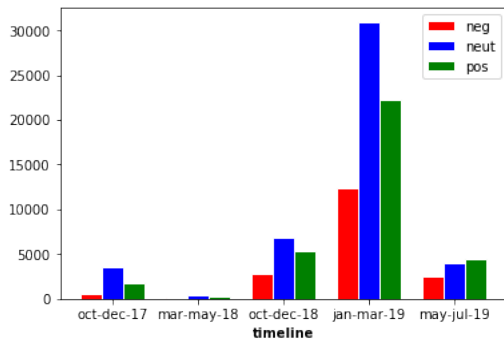


Figure 3: Sabarimala tweet frequency

- The graph shows an increase in percentage of tweets towards negative and positive sides, signalling increase in percentage of polarized reactions. (Figure 4)
- Negativity, sadness emotions have increased through the timelines with the maximum negativity and sadness being in the election period. (Figure 5)

7 Key Observations

- Our study confirms the "exhausted majority" observation of the Hidden Tribes study. Majority of people have neutral opinion. Can be observed both in Sabarimala and Demonetisation.

- The activity at the time of the event is a lot more than at other times.
- Polarity before the event is negligible.
- Once an event has happened the polarity remains even after a significant amount of time.
- Since 1) the number of people have decreased significantly, 2) percentage of neutral people have decreased significantly 3) People with higher polarity have increased. 1, 2 and 3 make us believe that people with strong opinions remain long after the event while people with not so strong opinions disappear as shown in the analysis of the above two events.

8 Limitations

- Limitations of translation and transliteration.
 - Transliterating Hindi to English transliterate only the words and leaves the sentence structure same.
 - Translation changes the words and the negativity of the words may not be captured in the translated words.
- Limitations of sentiment analyser.
 - The sentiment analyser uses naive bayes and ends up focussing on the toxicity of specific words, rather than getting the sentiment of the entire sentence.

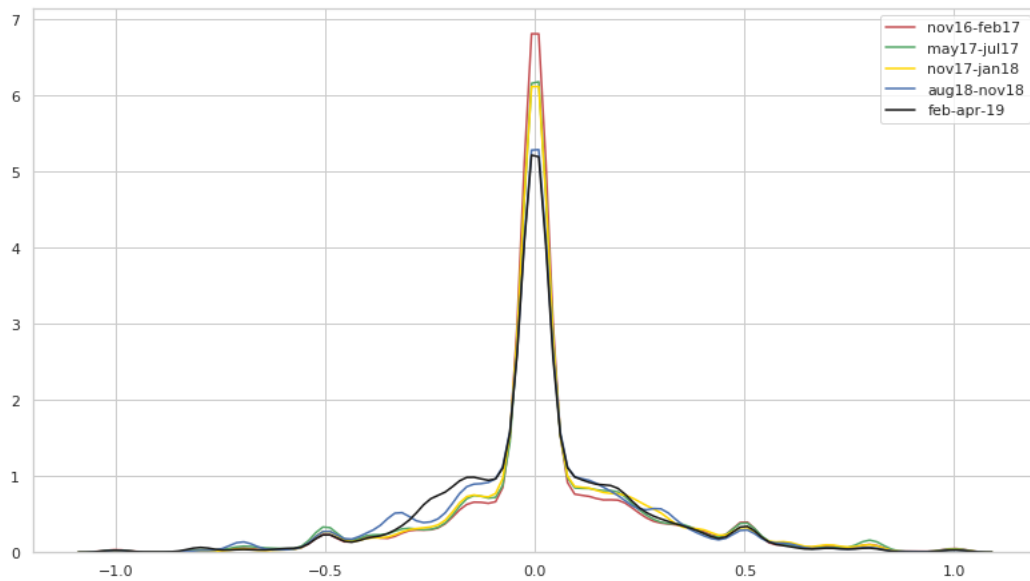


Figure 4: Demonetisation Gaussian curves

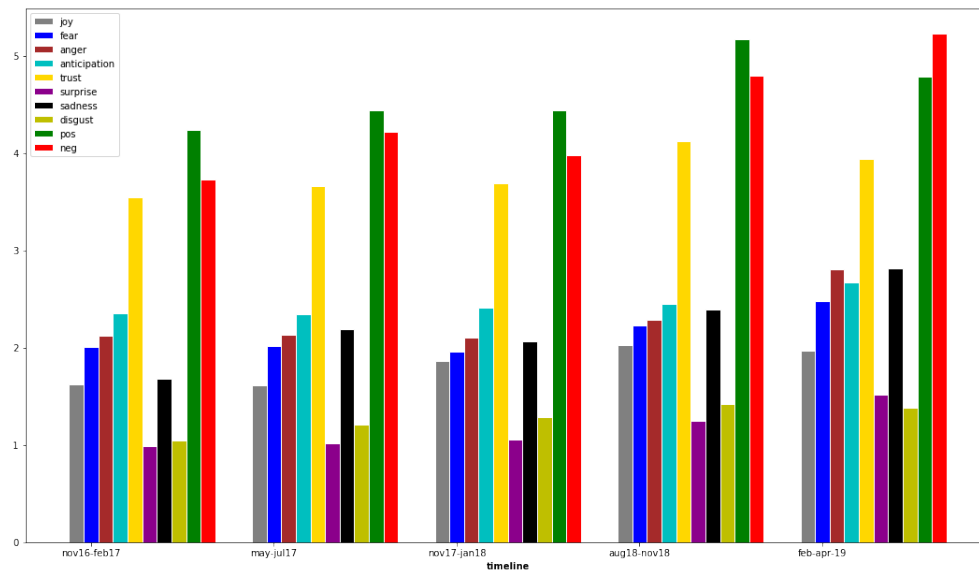


Figure 5: Emotion category Demonetisation

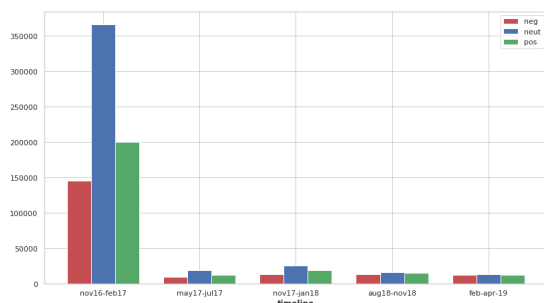


Figure 6: Demonetisation Tweet frequency

• Limitations of emotional category finding

tool.

- Each word is associated with one or more of the eight emotions. The analysis counts the total number of words with each emotion over the whole corpus.
- Ideally we would want an emotion to be associated with each tweet.
- Hence even though the overall emotion of the whole corpus is captured by word-by-word emotion categorisation, tweet-by-tweet emotion categorisation might help us gauge the emotions expressed

better.

9 Conclusion

We realize that:

- Majority of people hold neutral opinions.
- The polar opinions that seemingly remain after months are because of the people who were polar even before.

References

- Irshad Ahmad Bhat, Vandan Mujadia, Aniruddha Tam-mewar, Riyaz Ahmad Bhat, and Manish Shrivastava. 2015. [Iiit-h system submission for fire2014 shared task on transliterated search](#). In *Proceedings of the Forum for Information Retrieval Evaluation, FIRE '14*, pages 48–53, New York, NY, USA. ACM.
- Dainik Bhaskar. 2011. [Why millions throng sabari-mala shrine](#).
- Jefferson Henrique. 2018. [Getoldtweets api](#).
- Steven Loria. 2018. Textblob.
- Saif Mohammad and Peter Turney. 2013. [Crowdsourcing a word-emotion association lexicon](#). *Computational Intelligence*, 29.
- Pew Research Center. 2018. [Americans are wary of the role social media sites play in delivering the news](#).
- Miriam Juan-Torres Stephen Hawkins, Daniel Yudkins and Tim Dixon. 2018. The hidden tribes of usa. In *More in Common*.