

Understanding Caste System in Nepal: Surfacing changes in public sentiment on Twitter over time

Samyak Shrestha, '25

IPHS 200 Programming Humanity (Fall 2022) Prof Elkins and Chun, Kenyon College

Abstract

This project analyzes sentiment values taken from Twitter regarding the Nepali Caste System, which broadly borrows the classical Hindu *Chaturvarnashram* model, consisting of four broad social classes or varna: **Brahmin**, **Kshatriya**, **Vaishya**, **Sudra**. The caste system defines social classes by a number of hierarchical endogamous groups often termed *jaat*. Through this analysis, I aim to show major breakthroughs and disasters related to the caste system and how its dominance is becoming more transient.

Introduction/ Context

What is now commonly referred to as Hinduism was, in fact, a literary philosophy known as "Brahmanism," which enunciated the goals of a small, Sanskrit-educated social minority. The Brahman writings from texts were used to create the so-called quaternary hierarchy. This classification system was also textual or theoretical as it only existed in scrolls and had no link to reality on the ground. Since the direct rule of the British Raj in 1858 and the first censuses in the late 1860s, it has been clear that they tried to categorize the entire "Hindu" population into these four groups. However, the population's wide range of responses to caste identity made it impossible to neatly fit colonial or Brahman doctrine. Imposed by colonial powers and perpetuated by local elites, the caste system separated the population into the functional classes in the Hindu system called varnas where Brahmins functioned as priests, scholars, and men of letters; Kshatriyas were warriors responsible for maintaining order and providing security for the community; Vaishyas were farmers, herders, craftsmen, and merchants; and Shudras were the lowest level of workers, whose only mission was to serve the three other classes. Additionally, anyone outside of the varna was regarded as an outcast, untouchable, or Dalit.

Future Work

I would like to use more than one source to collect data from, for example, newspaper articles, or even Reddit. Furthermore, I would like to use topic modeling to analyze exactly how topics shift over time. For example, I will be able to disentangle and specify various topics under each cluster of tweets around peaks and valleys.

References/Acknowledgements

Elkins, K. (2022). *The Shapes of Stories: Sentiment Analysis for Narrative* (Elements in Digital Literary Studies). Cambridge: Cambridge University Press. doi:10.1017/9781009270403

J. C. (n.d.). GitHub - jon-chun/sentimentarcs_notebooks: SentimentArcs: a large ensemble of dozens of sentiment analysis models to analyze emotion in text over time. GitHub. Retrieved December 16, 2022, from https://github.com/jon-chun/sentimentarcs_notebooks

Chun, J. (2021). SentimentArcs: A Novel Method for Self-Supervised Sentiment Analysis of Time Series Shows SOTA Transformers Can Struggle Finding Narrative Arcs. arXiv. <https://doi.org/10.48550/arXiv.2110.09454>

Adhikari, D. (2015, August 5). Nepal caste system leaves dalits with secret suffering. जगारण मिडिया सेन्टर. Retrieved December 16, 2022, from <https://jagaranmediacenter.wordpress.com/2015/08/05/nepal-caste-system-leaves-dalits-with-secret-suffering/>

I would like to thank Professor Elkins and Professor Chun for their continued support and guidance in making this project possible. They have been accommodating and have helped me build various aspects of this project. One of my favorite classes!

Methodology

To retrieve data from twitter, 'snsrape', a python library for social networking services was used. The library ended up scraping 1,505 raw tweets, where it was given the parameters of keyword = 'nepal caste system', max_results = 50,000. However, since snsrape collects data according to volume, we were limited in searching for tweets under a certain time period.

After saving the scraped tweets in a csv format, I was ready to parse and clean its contents through three steps: basic Twitter text preprocessing, translating unicode icon emoji and text/punctuation emoticons to english words, and finally advanced Twitter text preprocessing. By uploading my raw csv file, I cleaned it and moved on to my sentiment analysis.

For computing the sentiment values, I used an ensemble of four models which are Lexicon-based sentiment analyzers (VADER, TextBlob, DistilBERT, Twitter RoBERTa). Using these models, I could generate various plot analyses and their sentiment values.

TextBlob was selected from the ensemble of models because it aligned most closely with with real world events and allowed us to view each peak and valley in detail.

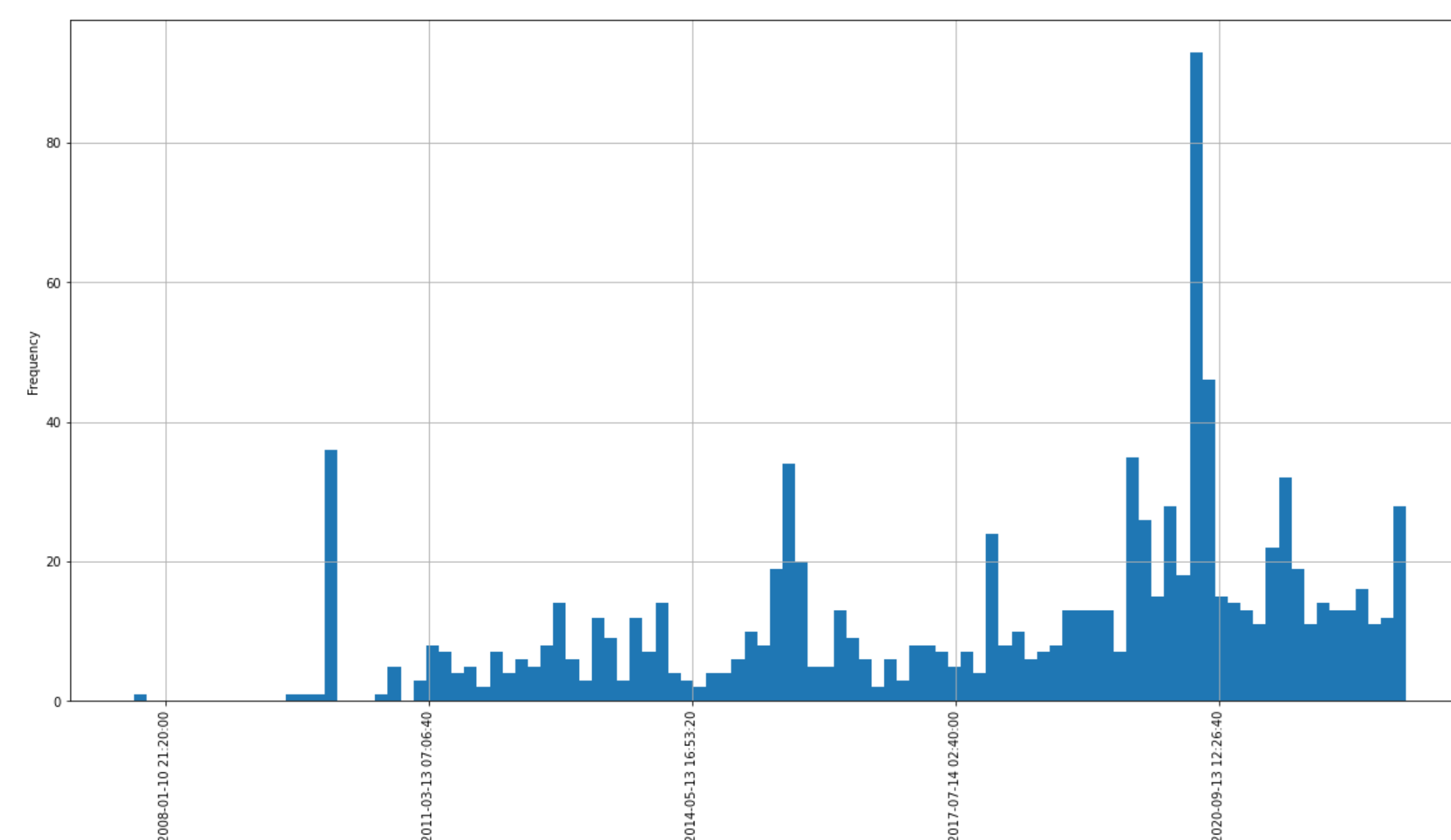
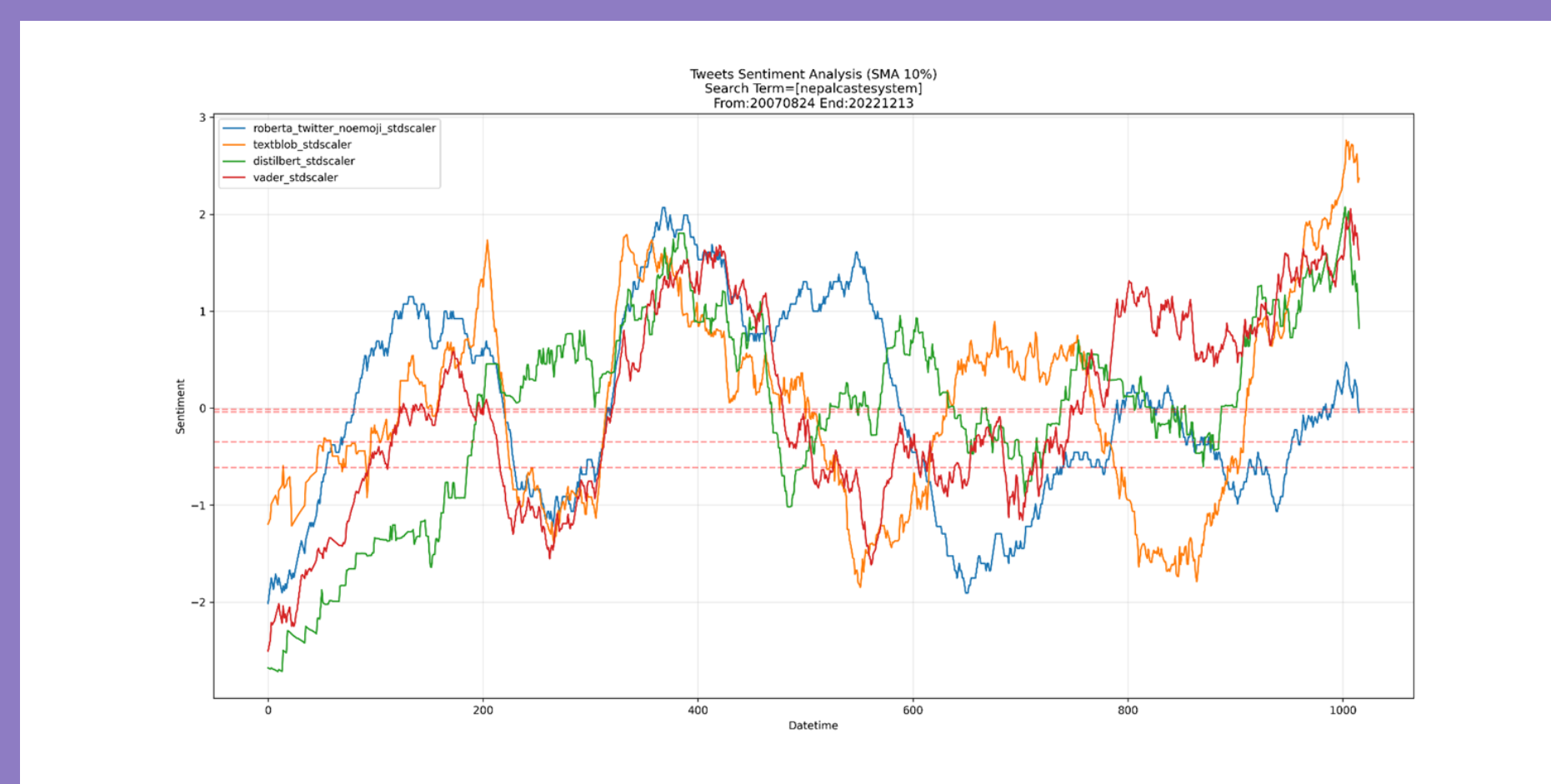


Figure I. Tweets retrieved by date

Sentiment Analysis

All three peaks and valleys are included in the 2010-2022 period. Some peaks are very sharp and suggest a single tweet can change the direction of the sentiment (1,5). While others show more sustained discussion over time with a similar sentiment (3,6).

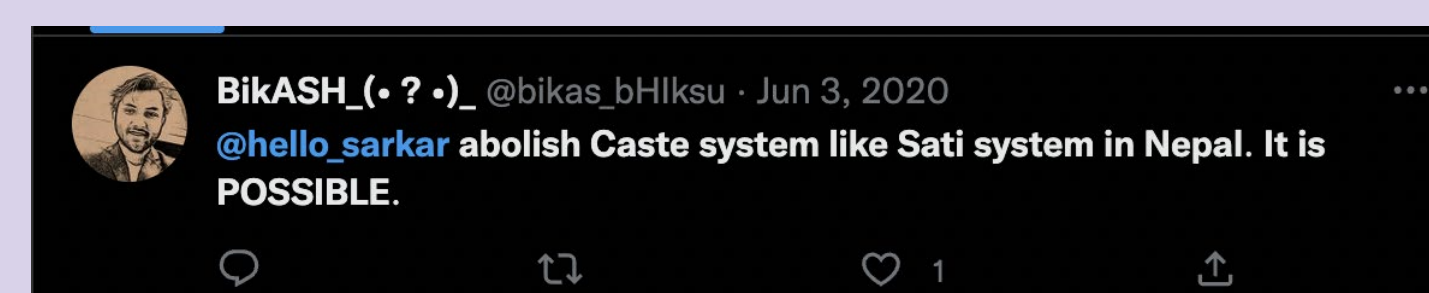
Observation 1: This first peak corresponds to the claim that Nepal in fact did not have the Hindu caste system during the Kirati rule (800 BC–300 AD).



Observation 2: This second peak corresponds to the “unique” characteristics of the caste system in Nepal, which necessarily isn't a positive correlation.



Observation 3: This final peak corresponds to the time period in June of 2020 which was a year filled with activism against cruelties of the caste system and the empowerment of Dalit voices. This was also the time that the hashtag #DalitLivesMatter was trending all over Nepali social media as well as many charities, rallies, and protests were held in solidarity. Similarly, there is belief in the citizens that the caste system can be completely “abolished” like how the Sati (a practice in which a widow sacrifices herself by sitting atop her deceased husband's funeral pyre) was in 1920. This point was plateaued and hovers around the 0.000 sentiment value due to the positive and negative tweets, which neutralized the group of tweets.



Observation 4: This first valley corresponds to the popular news article about discrimination against Dalits in 2015. It was labeled “Nepal Caste System Leaves Dalits with Secret Suffering” published by the Jagaran Media Center, and was a hot topic for many news sources.

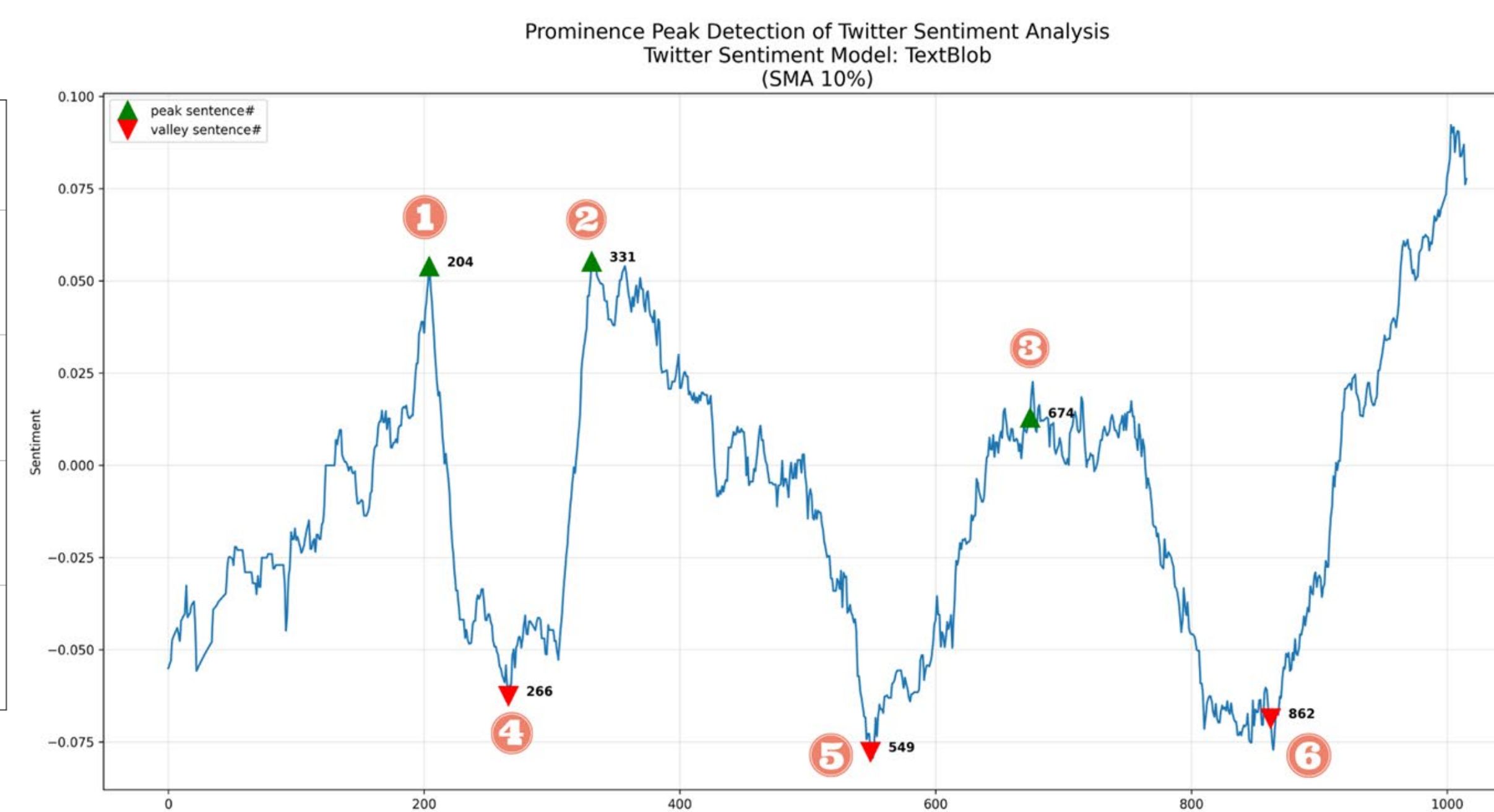


Figure II. Sentiment Analysis: Caste System in Nepal

Sentiment Analysis (continued)

They bring up problems regarding inter-caste marriages, bans from temples and government offices, denial of service, and discrimination in employment opportunities.



Observation 5: This valley corresponds to the discourse supporting Dalit Movements during 2020. While the Dalit movement was at a high, people were also discussing the origins of caste-based discrimination during British colonization and recalling freedom movements such as that of Bhimrao Ramji Ambedkar in India.



Observation 6: This final valley corresponds to the recent (2021-2022) discourse and hashtag #NepaCasteSystem, a caste system within a specific caste known as the “Newas” or “Newars”. It is a system by which Newārs, the historical inhabitants of Kathmandu Valley, are divided into groups on the basis of the Vedic varna model and divided according to their hereditary occupations. Conversations on the Newa caste system are still very unspoken and invisible in Nepali society.



Conclusion

Through the sentiment analysis, we found that earlier tweets (between 2010-2016) had fewer valleys and more peaks, possibly due to the lack of discourse and awareness about the marginalization of Dalits, untouchables, and other outcasts. They focused more on taking pride in having a “unique” caste system from India or that it was a product of British Colonization. As we progress further in time, people have become more accountable for the impacts of the caste system, and how modern society as a whole has been perpetuating it for centuries. Moreover, various human rights abuses and injustices have come to light in Nepal regarding the caste system and how it treats the “lower” castes.

Not only are people creating conversation, but they are also bringing change in Nepal. Especially the youth have started non-profit organizations, made appeals to legislation, and hosted multiple rallies to help improve the situation of the Dalits and others disadvantaged by the caste system.

Overall, I enjoyed brainstorming and analyzing ideas that were so close to home and observed the change in norms in Nepal. With the new generation being so active, I have faith in Nepal's future going forward.