

Proposal: Analyzing Ideology and Emotion: A Textual Examination of Indian Supreme Court Justices

Ayush Lahiri
PPOL 6801-01

March 23, 2024

1 Problem Statement

Since 2014, the BJP, led by Prime Minister Narendra Modi, has held significant power, with policies favoring Hindu majoritarianism. There are concerns that the Supreme Court's independence is being compromised, with claims that nominations of Supreme Court Judges are increasingly aligned with the BJP's ideology, reducing the court's ideological diversity (Dieterich, 2017). Despite the Supreme Court's longstanding reputation of being highly independent of the polity and objective (Jha, 2024), similarity of judge ideology can have long term effects on judicial quality such as setting singular precedents, further affecting future judgements via citations, lack of diversity in both legal argumentation and amongst bench members, wherein varying perspectives are key in enhancing deliberation (Orentlicher, 2018).

This study firstly aims to investigate ideological diversity among Supreme Court Justices across all cases, as well as by the core subject matter of the case such as business arbitration, constitutional amendments, religious disputes etc. This will be done by ascertaining semantic latent attributes from textual representations of oral judgements given by a judge across all cases they have presided over in a year, wherein the latent attribute aims to closely represent judicial ideology. Arguments presented by judges in a case can provide sub-conscious ideological leanings of a judge, such as sternness or sympathy for one type of case participant such as the state, which can further vary on the case context. These leanings can appear regardless of who the judge's decision favors, which is more often driven by objectivity.

Secondly the study looks to ascertain the use of emotion vs reasoning by Supreme Court Justices across all cases and by case subject. Evidence exists that as cases become more salient, justices also cater to increasingly relevant non-legal audiences (Jessee et al., 2022). Given the correlation between emotion and political ideology (Gennaro & Ash, 2021), I look to investigate whether a scaled measure of emotion and cognition is significantly correlated with ideological leanings.

2 Data

The study utilizes the Indian Legal Documents corpus (Malik et al., 2021) which contains judgement summaries for 35000 cases heard by the supreme court between 1958 and 2020. The judgement summaries contain information on the names of the presiding judge and names of the plaintiff and defendant. Further the dataset contains 56 documents annotated by 5 different legal experts. Annotated information contains the legal experts' prediction of the judgment, and highlighted the sentences they think are explanations for the judgment. The dataset is publicly available upon

request to the authors and has been acquired. In its raw format the corpus exists in a comma separated value format, and is unstructured, with unpredictable breaks of a single document over varying number of rows in the csv file, while other documents exist within a single line.

3 Methodology

I aim to utilize pre-trained text embedding models such as word2vec and doc2vec and fine tune the model on my corpus. I plan to experiment with two methodologies to find the average word embeddings for each justice in a given year (where document refers to the judgement recited for each case), i.e. (i) Finding the average of all document embeddings of a justice for a given year (ii) Appending metadata for the Justice name and year into respective documents and find the embeddings associated with name year pair. I then intend to follow the methodology outlined by Rheault and Cochrane (2019). The average n-dimensional vector yielded, representing the latent attributes of a judge, will then be compressed in dimensions through principal components analysis and the components explaining the largest variance in the data can be interpreted as ideological scores. I repeat the same exercise, but for judgements for specific case topics. These case topics will be discerned through an unsupervised topic modelling approach, specifically an implementation that leverages BERT based topic modelling.

Secondly, to understand use of emotionality versus reason among judges I follow the methodology as outlined by Gennaro & Ash (2021). The process involves isolating the embedding vectors related to words of emotion and reason based on thematically categorized words available in widely validated dictionaries such as Vader, and finding the average embeddings across all words for emotion and cognition respectively. For each judgement we then find the cosine similarity between the average embedding vector for all words in the judgement and the average vector for emotion and reason respectively. Scaling the two similarities with respect to each other yields a score of how emotional/reasonable judgements are.

Finally we can then assess, whether emotion is significantly correlated with our ideological scores through a multinomial logistic regression model by looking at the judge ideology as a categorical dependent variable and our scaled emotionality score as the independent variable.

4 Evaluating Success

Due to their being little research done on judicial text in the Indian context, interpretation of principal components of a judge's embeddings presents as an opaque task and would require manual validation of these scores against perception of judges. These perceptions while available across media publications, do not exist as a central repository or in a collatable format, and hence I would be relying on manual research of media publications and personality profiles of judges. Ideally, I would expect that the composition of supreme court is continually diverse in ideology and more emotional on topics such as constitutional amendment and religion/caste/gender disputes. I would find a more liberal leaning less diverse supreme court after 2014, most surprising, as it suggests that the BJP's expanding power may have not reached the Supreme court as popular narrative suggests.

Bibliography

Dieterich, C. (2023, December 27). Narendra Modi's grip on the Indian Supreme Court. *Le Monde.fr*. https://www.lemonde.fr/en/international/article/2023/12/27/in-india-narendra-modi-s-grip-on-the-supreme-court_6378593_4.html

Gennaro, G., & Ash, E. (2021). Emotion and Reason in Political Language. *The Economic Journal*, 132(643). <https://doi.org/10.1093/ej/ueab104>

Jessee, S., Malhotra, N., & Sen, M. (2022). A decade-long longitudinal survey shows that the Supreme Court is now much more conservative than the public. *Proceedings of the National Academy of Sciences*, 119(24). <https://doi.org/10.1073/pnas.2120284119>

Jha, A. (2024, January 20). From the India Today archives (2019) How Supreme Court finely balanced its Ram temple verdict. *India Today*. <https://www.indiatoday.in/ram-temple-verdict>

Malik, V., Sanjay, R., Nigam, S. K., Ghosh, K., Guha, S. K., Bhattacharya, A., & Modi, A. (2021). ILDC for CJPE: Indian Legal Documents Corpus for Court Judgment Prediction and Explanation. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* (pp. 4046–4062). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.acl-long.313>

Orentlicher, D. (2018). Politics and the Supreme Court: The Need for Ideological Balance. *University of Pittsburgh Law Review*, 79(3). <https://doi.org/10.5195/lawreview.2018.565>

Rheault, L., & Cochrane, C. (2019). Word Embeddings for the Analysis of Ideological Placement in Parliamentary Corpora. *Political Analysis*, 28(1), 112–133. <https://doi.org/10.1017/pan.2019.26>