HiBias: A Nine-Faceted Bias Annotation Dataset for Media Bias Detection in Hindi

No Author Given

No Institute Given

Abstract Although automated detection, understanding and mitigation of media bias is highly required in Indian context, this requires curation of trustworthy annotated news articles datasets in Indian languages, such as, Hindi, that covers multiple facets of bias. In this paper, we introduce the first annotated dataset consisting of 400 unique articles from two leading Indian news media agencies in Hindi language. Our annotations include 9 different types of bias, that ensures exhaustiveness and additionally, include explanations from the annotators. The dataset and replication code are publicly available at [1] and [2].

Keywords: media bias dataset · framing bias · Indic languages · clickbait · linguistic bias

1 Introduction

With the era of digitization, users have continuous access to news which leads to severe competition among news media agencies to attract users and retain their attention [3, 2]. While users rely on the news media agencies for a fair and high quality reporting of news events, there has been several instances of deviation from journalistic news values in news reporting, such as, deliberately lying or leaving out context, not fact-checking sources, using clickbait, being biased, using politically aligned news reporting [16, **?**], etc., for audience attention. Subsequently, the continuous information overload from different sources can overwhelm the user, makes it difficult to differentiate between biased and unbiased news manually, and lead to biased news consumption [17, **?**,**?**]. Bias in news media reporting refers to non objective reporting of a news event through different ways, such as, omission or selective reporting of a news topic, slanted view (slanting can be in form of any kind of demographic group, such as, political leaning, gender [5], religion , location, ethnicity [7, 11], etc.,). In summary, media bias could be defined as any form of deviation from journalistic values by providing a false picture of the genuine information and has several [9, 10]. Therefore, irrespective of the type, news media bias can impact the public opinion adversely, spread misinformation, lead to formation of echo chambers and

[1] https://zenodo.org/records/18253271
[2] https://github.com/KushalTrivedi19032005/Indic-Media-Bias-Detection-Dataset

even polarization in society. In order to prevent these scenarios, there is a requirement for identification, detection and mitigation of bias to ensure informed readership [1].

This would further aid government agencies to implement automated monitoring systems and journalists to ensure objectivity, and fairness in reporting. However, the extensive information overload and continuous availability of news makes it increasingly difficult for users to manually ensure segregation of biased news from unbiased news in order to ensure responsible news consumption [13]. Additionally, research surveys indicate existence of political bias in Indian news articles irrespective of the type of bias [3]. India ranks poorly according to World Press Freedom Index 2025 [4] which further highlights the requirement of unbiased news production and consumption. This concern is particularly relevant in today's digital era, where a 2022 study by Oxford University Press (OUP) found that 54% of Indian youth turn to social media when seeking factual information [5]. However, automated bias detection is highly challenging as it requires identification and understanding of the underlying implicit attributes and its various forms, such as framing bias, selection bias, and presentation bias [11]. Each type requires specialized bias detection approaches [14, 15]. Current research primarily focuses on English-language media and often analyzes bias at the news source level, which, while useful, offers only generalized insights and fails to capture bias at the article level. Although some work exists on detecting different bias types, Indic languages remain largely unexplored in this domain. The high variance and richness of Indic language differences along with India's complex political landscape featuring a multi-party system unlike the largely binary political ideologies in other English speaking countries like the USA or UK makes direct application of bias detection approaches ineffective [6, ?]. Furthermore, translation-based approaches fail to account for culturally and linguistically specific nuances of Indian society. Thus, there is a pressing need to develop bias detection techniques tailored to Indic languages, and further mitigate bias in news media.

In order to build automated systems for media bias detection and mitigation for the Indian context and specifically, in Indic languages, such as, Hindi, there is a requirement for specific extensive datasets. Although there exists few research works to create annotated datasets for media bias detection, none of these are directly related to Indian society, specifically different non-English Indian languages. For example, [15] analyzed Indian English newspapers and their coverage of four large-scale news events to study information coverage bias and user reactions, [4] and [12] used sentiment-based analysis to understand bias

[3] https://maktoobmedia.com/india/lokniti-csds-study-80-journalists-believe-media-covers-modi-govt-too-favourably/

[4] https://www.outlookindia.com/national/india-ranks-151-out-of-180-countries-in-world-press-freedom-index-2025-rsf-calls-it-one-of-worlds-most-dangerous-countries

[5] https://www.livemint.com/news/india/despite-misinformation-concern-54-indians-derive-factual-information-from-social-media-reveals-oup-study-11656420664780.html

Table 1: Annotation Guidelines for Bias Classification

SNO	Bias Labels	Description
1	**Selection Bias**	**Rule:** Focus on what is included or, more importantly, excluded from the story or the entire news cycle. • **Inclusion/Exclusion of Sources:** A story predominantly features sources supporting one side (e.g., only Party X spokespersons) while ignoring or minimizing opposing voices (e.g., opposition leaders). • **Topic Filtering:** One key political event, scandal, or achievement related to one party is given zero coverage or is buried deep, while similar news about another party is highly publicized. • **Data Cherry-picking:** Using only the specific statistics or facts (e.g., economic growth rates, poll numbers) that favor a particular narrative and omitting inconvenient data.
2	**Framing Bias**	**Rule:** Focus on how the story is packaged, emphasizing certain aspects to influence interpretation. • **Choice of Metaphor/Analogy:** Using loaded terms or imagery to describe a party or event (e.g., describing a protest as 'chaos/anarchy' (अराजकता) vs. 'democratic expression' (लोकतांत्रिक अभिव्यक्ति)). • **Order of Information:** Placing the preferred narrative's conclusion or preferred facts (e.g., "The government succeeded...") at the very start of the headline/story. • **Focus of Responsibility:** Attributing success/positive outcomes to one party's 'vision' (दृष्टि) or 'leadership' (नेतृत्व), but blaming negative outcomes on external factors or the 'past' (अतीत).
3	**Linguistic Bias**	**Rule:** Focus on the specific words, adjectives, and nouns used to describe people, groups, or actions. (Applicable to Hindi vocabulary). • **Loaded Adjectives:** Describing an action using emotionally charged words (e.g., a leader's firm decision is called 'bold' vs. 'stubborn'). • **Labeling:** Using different labels for similar groups (e.g., calling one group's supporters 'patriots' and an opposing group's supporters 'agitators'). • **Euphemisms/Dysphemisms:** Softening negative actions (e.g., calling a major financial collapse a 'minor adjustment') or hardening neutral ones.
4	**Ideology Bias**	**Rule:** Focus on how the news story aligns with or subtly promotes a specific political, economic, or social doctrine. • **Normalization of Views:** Presenting a partisan or extreme ideological position (e.g., hyper-nationalism or far-left policy) as the only acceptable or mainstream viewpoint. • **Policy Critique:** Consistent, unquestioning support for or dismissal of policies based solely on the party's ideological stance. • **Historical Interpretation:** Re-interpreting historical events or figures in a way that aligns with the favored ideological narrative.
5	**Stance Bias**	**Rule:** Focus on the reporter, anchor, or publication's express or implied approval or disapproval of a political actor or event. • **Approval/Disapproval in Tone:** The overall tone of the narration is visibly fawning/respectful when discussing one side and skeptical/derisive when discussing the other.
6	**Omission Bias**	**Rule:** Focus on the intentional leaving out of crucial context or necessary facts that would alter the story's meaning. This includes **Missing Context**, i.e., reporting a politician's controversial statement without including the preceding question or the full sequence of events that provoked the comment or **Withholding Key Facts** or **Incomplete Follow-up**
7	**Fact Bias**	**Rule:** Focus on the use of verifiable information, particularly where a genuine fact is mixed with opinion or presented as fact, such as opinion as fact, misrepresentation or over-simplification.
8	**Emotional Bias**	**Rule:** Focus on language and presentation designed to elicit a strong emotional response (anger, fear, solidarity) rather than rational thought through appeal to fear, anger, sympathy, etc. and usage of highly emotional storyline.
9	**Clickbait Bias**	**Rule:** Focus on the headline. Assign this label if the headline is designed to attract readers through sensational, controversial, or misleading statements.

in Indian English news sources. While [8] provides a detailed understanding of bias in Indian news through the lens of different events. However, they focus on understanding bias only at the sentence level for English news articles, whereas we also intend to understand the implicit contextual bias at the news article level and for Indic languages. Therefore, in this paper, we introduce the first Hindi media political bias benchmark dataset, covering nine distinct types of bias based on both story selection and linguistic framing. To the best of our knowledge, this is the first dataset in the Indian context that covers an exhaustive types of biases. However this is a preliminary work, we are working on it to make it more extensive and exhaustive with respect to different news media agencies, demographic biases, different Indian languages. This dataset aims to support the research community in developing robust and trustworthy automated systems for monitoring and mitigating news media bias in Indian context.

2 Dataset Details and Observations

In this Section, we discuss dataset collection, description and annotation, respectively.

Dataset Collection : We have considered two leading Indian News Media agencies for Hindi Language, i.e., *Aaj Tak*[6] and *Live Hindustan*[7], respectively. Our dataset comprises of 400 news articles in which 200 were collected from *Live Hindustan* and 200 were collected from *Aaj Tak*. Out of the *Live Hindustan* there 118 articles which were biased and 82 as unbiased and *Aaj Tak*, 64 articles were found to be biased and 36 to be unbiased. After dataset collection, we organize our dataset into the following fields, *news_publisher*, *article_date*, *article_links*, *Bias Label* which comprises of binary classes, i.e., either Yes or No, *Selection Bias Label*, *Framing Bias Label*, *Linguistic Bias Label*, *Ideology Bias Label*, *Selection Bias Label*, *Stance Bias Label*, *Omission Bias Label*, *Fact Bias Label* and *Clickbait Bias Label*. For example, if an annotator identifies selection bias, stance bias, and clickbait bias in an article, the corresponding annotation labels would be 1, 0, 0, 0, 1, 0, 0, 1 for the nine bias labels, respectively. Finally, our dataset includes *rationale* provided by an annotator as explanations that provide reasoning of their decision. This further ensures the interpretability of the automated system.

Dataset Annotation : For annotation procedure, we provided the annotators with detailed overview of the task with detailed examples, a snippet is shown in Table 1. Every news article was annotated by 2 annotators. An article can have one or multiple of these biases. Our Cohen's Kappa [18] score is 0.81 which shows high trustworthiness in annotations for whether an article is biased or unbiased. However, we observed that understanding the type of bias has an average Cohen's Kappa score of 0.65 which highlights the difficulty in understanding type of bias and further, requirement of this nuanced dataset.

[6] https://www.aajtak.in/
[7] https://www.livehindustan.com/

Observations : Our preliminary observations (as shown in Fig 1) indicate that a high fraction of Hindi political news articles exhibit bias from both *Aaj Tak* and *Live Hindustan*, i.e., around 64% and 59% (as per the dataset). Although this is based on only a sample of news articles from 2 news media agencies, it highlights presence of news media bias in Hindi news across different media organizations and the need to create a large-scale Hindi media bias dataset. Additionally, we observe that selection, framing, and linguistic biases were the most prevalent forms in this dataset, comprising of 42%, 30%, and 28%, respectively.

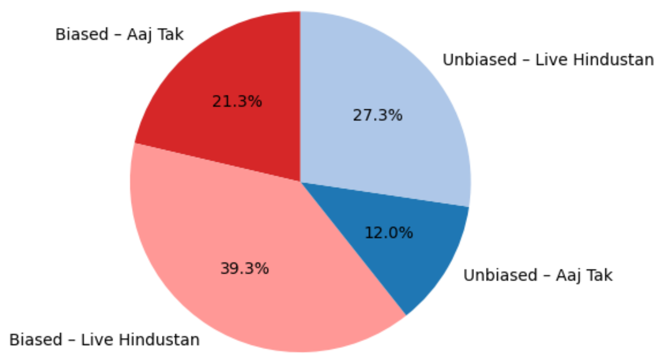

Figure 1: Distribution of biased and unbiased news articles across agencies.

3 Conclusions

Media Bias can severely impact public perception of the event, form echo chambers, and cause segregation and polarization in society. Therefore, identification of different forms of media bias and further mitigation and handling is severely required. Developing automated approaches for low-resource domains can be challenging due to the non-availability of task and domain-specific data. Additionally, understanding the implicit factors of Indian political context, specifically with the high variance across different languages is almost impossible. Furthermore, the common masses are more likely to read news in Indian languages, such as Hindi, specific to their region rather than English, which makes understanding of media bias detection and mitigation specifically for Hindi highly relevant. Therefore, we propose to create Indian language-specific, Hindi based, media bias benchmark datasets, that cover multiple facets of bias considering two major newspapers and 400 news articles. Although this is a preliminary work, our initial observations suggest high forms of bias while ensuring trustworthy annotations. Our future directions include extension of this dataset to large-scale by increasing number of news articles and news media agencies followed by more Indic languages.

References

1. Arifah, I.D.C., Maureen, I.Y., Rofik, A., Puspila, N.K.W., Erifiawan, H., et al.: Social media platforms in managing polarization, echo chambers, and misinformation risk in interreligious dialogue among young generation. Journal of Social Innovation and Knowledge **1**(aop), 1–33 (2025)
2. Chakraborty, R., Bhandari, S., Chakraborty, N., Das, R.: Eve2sign: Creating signed networks of news events. In: Text2Story@ ECIR. pp. 79–87 (2020)
3. Chakraborty, R., Bhavsar, M., Dandapat, S.K., Chandra, J.: Tweet summarization of news articles: An objective ordering-based perspective. IEEE Transactions on Computational Social Systems **6**(4), 761–777 (2019)
4. Chauhan, S., Verma, S., Pandey, A., Srivastava, A., Malik, P., Garg, N., Singh, A.K.: Media bias monitor: Quantifying biases of indian print media. In: 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT). pp. 1–7. IEEE (2024)
5. Cheng, M., Durmus, E., Jurafsky, D.: Marked personas: Using natural language prompts to measure stereotypes in language models. In: Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). pp. 1504–1532 (2023)
6. Horych, T., Mandl, C., Ruas, T., Greiner-Petter, A., Gipp, B., Aizawa, A., Spinde, T.: The promises and pitfalls of llm annotations in dataset labeling: A case study on media bias detection. In: Findings of the Association for Computational Linguistics: NAACL 2025. pp. 1370–1386 (2025)
7. Kambhatla, G., Stewart, I., Mihalcea, R.: Surfacing racial stereotypes through identity portrayal. In: Proceedings of the 2022 ACM conference on fairness, accountability, and transparency. pp. 1604–1615 (2022)
8. Kuila, A., Jena, S., Sarkar, S., Chakrabarti, P.P.: "news has various shades": Quantifying and analyzing media bias at aspect-level granularity. ACM Journal on Computing and Sustainable Societies **2**(4), 1–51 (2024)
9. Kuuse, S.F., Kangur, U., Chakraborty, R., Sharma, R.: Crowdsourced fact-checking or biased commentary? analyzing political bias in twitter's community notes. In: Companion Proceedings of the ACM on Web Conference 2025. pp. 2661–2669 (2025)
10. Lüüsi, L., Kangur, U., Chakraborty, R., Sharma, R.: Political stance detection in estonian news media. In: Proceedings of the 9th International Workshop on Computational Linguistics for Uralic Languages. pp. 12–28 (2024)
11. Navon, I.R.L., Reingold, O., Shen, J.H.: Fairness with respect to stereotype predictors: Impossibilities and best practices. Transactions on Machine Learning Research (2025)
12. Rawat, S., Vadivu, G.: Media bias detection using sentimental analysis and clustering algorithms. In: Proceedings of international conference on deep learning, computing and intelligence: ICDCI 2021. pp. 485–494. Springer (2022)
13. Rodrigo-Ginés, F.J., Carrillo-de Albornoz, J., Plaza, L.: A systematic review on media bias detection: What is media bias, how it is expressed, and how to detect it. Expert Systems with Applications **237**, 121641 (2024)

14. Sabane, M., Litake, O., Chadha, A.: Breaking language barriers: A question answering dataset for hindi and marathi. arXiv preprint arXiv:2308.09862 (2023)
15. Sen, A., Ghatak, D., Khanuja, G., Rekha, K., Gupta, M., Dhakate, S., Sharma, K., Seth, A.: Analysis of media bias in policy discourse in india. In: Proceedings of the 5th ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies. pp. 57–77 (2022)
16. Spinde, T., Rudnitckaia, L., Sinha, K., Hamborg, F., Gipp, B., Donnay, K.: Mbic – a media bias annotation dataset including annotator characteristics (2021), https://arxiv.org/abs/2105.11910
17. Spinde, T., Plank, M., Krieger, J.D., Ruas, T., Gipp, B., Aizawa, A.: Neural media bias detection using distant supervision with BABE - bias annotations by experts. In: Moens, M.F., Huang, X., Specia, L., Yih, S.W.t. (eds.) Findings of the Association for Computational Linguistics: EMNLP 2021. pp. 1166–1177. Association for Computational Linguistics, Punta Cana, Dominican Republic (Nov 2021). https://doi.org/10.18653/v1/2021.findings-emnlp.101, https://aclanthology.org/2021.findings-emnlp.101/
18. Warrens, M.J.: Five ways to look at cohen's kappa. Journal of Psychology & Psychotherapy **5** (2015)