

Quantify the Political Bias in News Edits: Experiments with Few-Shot Learners (Student Abstract)

Preetika Verma¹, Hansin Ahuja², Kokil Jaidka³

¹ Birla Institute of Technology and Science, Pilani

² Google India

³ National University of Singapore

preetikaverma08@gmail.com, jaidka@nus.edu.sg

Abstract

The rapid growth of information and communication technologies in recent years, and the different forms of digital connectivity, have profoundly affected how news is generated and consumed. Digital traces and computational methods offer new opportunities to model and track the provenance of news. This project is the first study to characterize and predict how prominent news outlets make edits to news frames and their implications for geopolitical relationships and attitudes. We evaluate the feasibility of training few-shot learners on the editing patterns of articles discussing different countries to understand their wider implications in preserving or damaging geopolitical relationships.

Introduction

Published news articles go through a series of updations in their life cycle - updating of figures, the addition of new facts, and correction of misinformation. Consequently, the article headline, whose purpose is to give an idea about the text below it, is also updated. Modeling the updates and tracing patterns in these revisions is useful for multiple tasks in journalism.

Beyond serving as a way to communicate facts, content edits serve an important purpose in setting the news agenda in society (Entman et al. 2004). When content is added to a news article, it can refocus audience attention to issues of salience, acknowledge the actors or context to an issue, or update an unfolding event (Tewksbury and Scheufele 2019). On the other hand, when content is removed, it can serve to simplify a complex and developing situation, remove public attention from a topic or point of view, and enforce a neutral point of view.

The NewsEdits dataset (Spangher et al. 2022) aims to explore text revision patterns and predictability of four article-level actions on sentences: additions, deletions, edits, and refactorizations. It consists of 1.2 million articles with 4.6 million versions from 22 media outlets. Headline updates have been observed in approximately 17% version pairs (376,944 after sampling 2 million pairs). For our study, we focus on 4 English language wires: Fox, New York Times, Washington Post, and Rebel.

Although the literature on news framing is many decades old (Tewksbury and Scheufele 2019), no study yet has leveraged the possibility of studying dynamic news framing practices through the provenance of news articles. We address this gap through a multidisciplinary approach that leverages deep learning to scale up a small-scale fine-grained content analysis of newspaper editing practices. Our project has implications for understanding the existing geopolitical biases that drive news publishing, with opportunities to relate them to the broader geopolitical attitudes of the general public.

We focus on quantifying the edit intentions of article headlines and modifying the edit taxonomy developed in prior work (Yang et al. 2016). Next, we evaluate the feasibility of training few-shot GPT-3 classifiers on a hand-annotated sample of news headlines from four news outlets. In the following paragraphs, we have discussed how we prepared and annotated the dataset, the performance of GPT-3 classifiers on detecting edits, and directions for future work.

Dataset Preparation

First, we retained edit pairs from headlines and articles with a maximum similarity threshold of 0.2 for average distance (Spangher et al. 2022). In this manner, we were left the dataset distribution reported in Table 1. Next, we used dictionary approaches to identify whether and where the news articles referred to six countries of interest with important geopolitical ties to the United States - China, Afghanistan, Syria, Iran, Israel, and France, and retain only the edited news headlines corresponding to those countries. The dictionaries comprised the names of the countries, major cities, and heads of state. Figure 1a reports the proportion of news headlines from different countries.

News outlet	No. of articles	No. of versions	No. of headline updates
Fox	78,566	117,171	5,155
Rebel	4344	19383	114
NY Times	87,556	395,643	59,750
WaPo	19,184	68,612	9,532

Table 1: Corpus statistics

	Accuracy	Precision	Recall	Macro-F1	Minority F1
Added - objective	0.64	0.34	0.34	0.54	0.34
Added - subjective	0.79	0.43	0.39	0.64	0.39
Removed - objective	0.89	0.20	0.10	0.54	0.13
Removed - subjective	0.87	0.10	0.15	0.52	0.12

Table 2: Predictive performance with GPT-3 classifiers

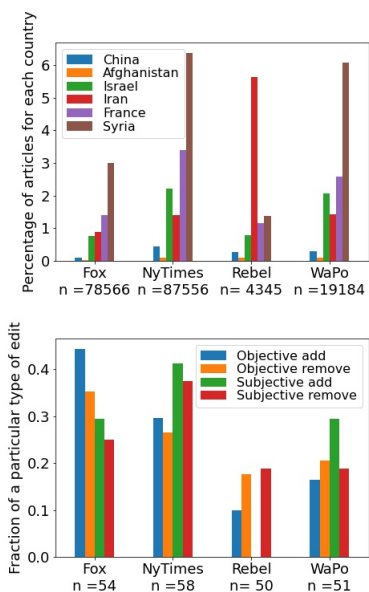


Figure 1: Proportion of news headlines in our corpus that (a) pertaining to different countries and (b) apply different edits.

Annotating Edit Intentions

We are mainly interested in *content edits*, which make substantive or developmental changes to the content. This is opposed to *copy-edits*, which characterize minor edits made to paraphrase or improve the readability of the content.

After a preliminary exploration of the Wikipedia edit taxonomy (Yang et al. 2016), a new set of edit intentions was created by the authors in discussion. This is because while some of the high-level goals of editing behavior are consistent between news- and Wikipedia editing (such as clarification, elaboration, specification, and enforcing a neutral point of view), others are not. News editing sometimes involves the addition of affective content or the removal of less critical details that may emerge from its agenda-setting and news-framing roles.

The authors reorganized the coding scheme to include seven categories. Two of them independently annotated a stratified random sample of 200 news headline pairs. They disagreed on 25 out of 200 examples, resulting in a macro-level inter-coder percentage agreement of 87.5%. Disagreements were resolved through discussion and applied to further refine the edit taxonomy.

We have adapted and reorganized the taxonomy of edit intentions evidenced in Wikipedia editing behavior. We con-

sider news editing to be a hierarchy of goals and content types. When the goal of editing is to add information, the information added could provide more information about the news item. Alternatively, it could provide affective information to guide the readers’ thought process. Therefore, our edit taxonomy comprises four categories - objective addition, objective removal, subjective addition, and subjective removal of information.

Results and Future Work

Table 2 contains the predictive performance when we trained GPT-3 classifiers on 200 observations with sparse data labels in a ten-fold cross-validation setup. The label distribution is reported in Figure 1b. The low Minority-F1 cases indicate just how challenging the problem of inferring news edit intentions really is, especially for edits involving information removal, which have a large class imbalance.

Through this project, we hope to shed light on how news publishers gatekeep geopolitical boundaries and set subjective agendas, even in an era of unending information. As the next steps, we will also explore the addition and removal of social identity categories such as gender, race, ethnicity, and religion, as a substantive body of work suggests that they are often used in headlines for agenda-setting (Price 1989). We will also experiment with classical approaches trained on the sentiment, semantic, and similarity-based differences between edited text pairs while continuing our experiments with few-shot learners.

References

- Entman, R. M.; et al. 2004. *Projections of power: Framing news, public opinion, and US foreign policy*. University of Chicago Press.
- Price, V. 1989. Social identification and public opinion: Effects of communicating group conflict. *Public Opinion Quarterly*, 53(2): 197–224.
- Spangher, A.; Ren, X.; May, J.; and Peng, N. 2022. NewsEdits: A News Article Revision Dataset and a Novel Document-Level Reasoning Challenge. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 127–157.
- Tewksbury, D.; and Scheufele, D. A. 2019. News framing theory and research. In *Media effects*, 51–68. Routledge.
- Yang, D.; Halfaker, A.; Kraut, R.; and Hovy, E. 2016. Edit categories and editor role identification in Wikipedia. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC’16)*, 1295–1299.