Can Large Language Models Support Editors Pick Related News Articles?

Bilal Mahmood (University of Bergen, Norway)
Mehdi Elahi (University of Bergen, Norway)
Samia Touileb (University of Bergen, Norway)
Lubos Steskal (TV 2 Norway)

Media Futures

Abstract

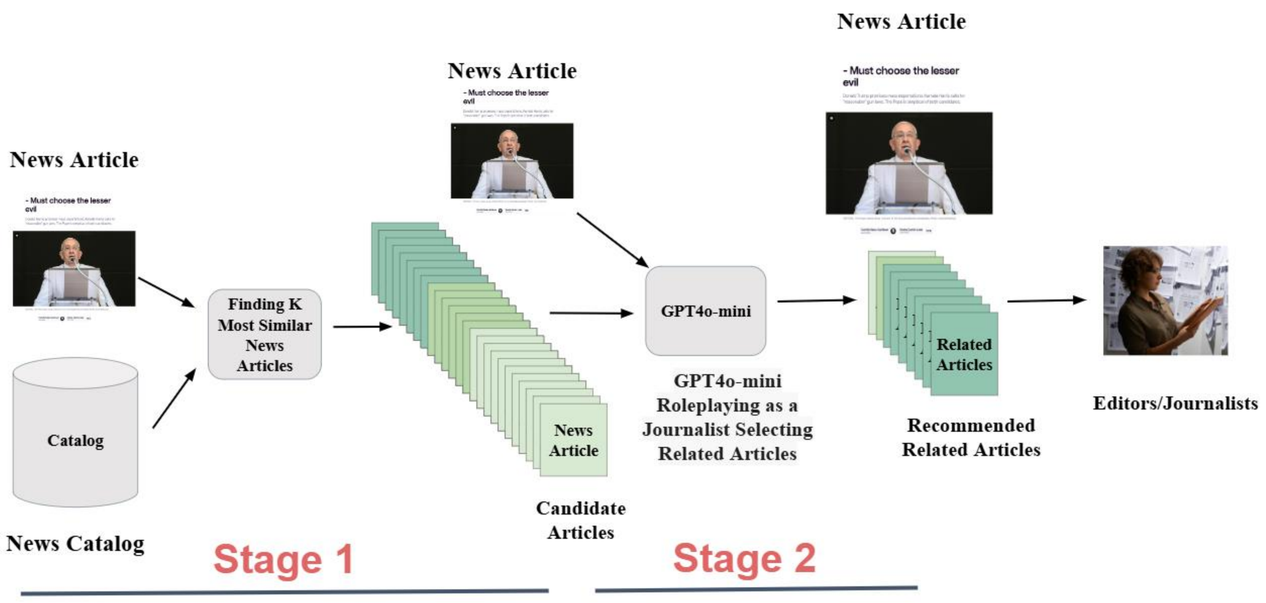
Editors and journalists are essential to news platforms, creating credible stories, choosing front-page content, and hand-selecting related articles for readers. This process, often done **manually** using an internal search tool, can be **time-intensive**. To address this, the work proposes an **automatic** system using a Large Language Model, **GPT4o-mini**, to generate recommended related articles for editors' review.

Research question

RQ: How can a recommendation mechanism based on Large Language Models (LLMs) be effectively utilized to support editors and journalists in their task of selecting related news articles?

Approach

Overall schematic view of our 2-stage approach to support journalists and editors in their task of selecting related articles for a given news article at TV 2.



Role-playing system prompt given to the LLM, **GPT4o-mini**, asking it to give **relatedness score** to the potential article as well as the **explanation** as if it were a journalist.

System Prompt

You are a journalist and want to select the related article. I will give you two articles, one main article and the other potential related article and you have to tell me whether that potential article is related or not by providing a score on a scale from 1 to 10. Return the result in a JSON format with related score, and explanations describing the reasoning behind the score as keys. The explanations should be less than 50 words. Start the explanation with the main article and then the potential article.

Prompt

Main article: {main_article} Potential article: {potential_article}. Answer:

Results

Evaluating this approach on **236** test news articles from TV 2, a major Norwegian news platform, showed promising results in supporting editors and journalists in selecting related articles. Recommending Top 5 most related articles using our proposed approach achieved a **Recall@5** score of **56%.**

Approach	CandSize	Recall		Precision		MAP	
		@5	@CandSize	@5	@CandSize	@5	@CandSize
KNN (baseline)	_	0.436	0.436	0.107	**0.107**	0.366	0.366
KNN+GPT	5	0.436	0.436	0.107	**0.107**	0.398	0.366
KNN+GPT	10	0.488	0.520	0.121	0.065	0.422	**0.370**
KNN+GPT	20	0.554	0.606	**0.136**	0.039	**0.456**	0.369
KNN+GPT	50	**0.559**	**0.656**	**0.136**	0.017	0.448	0.367

Conclusion

In an offline setting, we demonstrated that our approach effectively recommended related articles to TV2's journalists and editors. Our method achieved a **Recall@5** score of **56%**, meaning that more than half of the articles selected by journalists and editors were included in our recommendations. For future work, we aim to experiment with various prompts, test LLMs of different sizes, and eventually create a **practical tool** to integrate into TV2's **newsroom**, enhancing productivity for journalists and editors.

PARTNERS

amedia Schibsted 2 NRK Universitas Oslo Bergens Tidende nb.no VIZRT WOLFTECH NORCE UiS Faktisk.

HOST

University of Bergen

FUNDED BY

This research is funded by SFI MediaFutures partners and the Research Council of Norway (grant number 309339).

Forskningsrådet Norwegian Centre for Research-based Innovation