

# ABSTRACT

Personalized news recommender systems play a vital role in addressing information overload by delivering relevant and up-to-date news content to users. However, most previous studies focus on only one aspect: approaches that rely solely on semantic relevance often fail to account for content freshness, while methods based solely on recency tend to recommend news articles that lack contextual relevance. To address this limitation, we propose an innovative news recommendation framework that integrates the RoBERTa-Large transformer model with both textual features (titles and abstracts) and temporal features (publication dates). Our approach introduces a modular scoring mechanism based on late fusion, which adaptively combines relevance scores, and time decay scores derived from Unix timestamps. This enables the system to flexibly balance content relevance and recency at inference time. To enhance temporal diversity in the data, we utilize the Microsoft News Dataset (MIND), a large-scale dataset collected from MSN News, which we further enriched by scraping publication dates using the web scraping tool Apify. Experimental results on MIND, averaged over 10 runs, show that the proposed RoBERTa-Large model consistently outperforms baseline models such as BERT and DeBERTaV3, achieving an AUC of 0.7813. Despite its computational demands, our method significantly enhances personalized news recommendation by effectively capturing both semantic relevance and temporal dynamics.

**Keywords:** Personalized news recommendation, RoBERTa-Large, Temporal features, Time decay, News relevance, Microsoft News Dataset