

## SECTION 13.

### COMPUTER AND SOFTWARE ENGINEERING

---

**Логвінова Олена Олегівна**

здобувач вищої освіти факультету комп'ютерних наук  
*Харківський національний університет радіоелектроніки, Україна*

**Науковий керівник: Смеляков Кирило Сергійович** 

д-р. техн. наук, професор  
*Харківський національний університет радіоелектроніки, Україна*

---

## COMPARATIVE STUDY OF EMBEDDING-BASED RECOMMENDATION METHODS FOR PERSONALIZED NEWS

The rapid expansion of digital news platforms has significantly increased the volume of available information, creating challenges for users in identifying relevant content. In this context, recommender systems have become essential tools for filtering information and delivering personalized news feeds. However, the news domain presents unique difficulties, including high content turnover, short relevance periods, and constantly evolving user interests.

Conventional recommendation techniques, such as collaborative filtering and content-based filtering, are widely applied but demonstrate limited effectiveness in dynamic environments [1]. Collaborative filtering relies on historical user interactions and often fails when dealing with new users or newly published articles [2]. Content-based methods focus on item attributes but typically lack the ability to capture deeper semantic relationships within textual data [3]. Although hybrid approaches attempt to mitigate these issues, they tend to increase system complexity [4].

Recent developments in natural language processing have enabled the application of embedding models, which encode textual data into dense vector representations. These representations allow systems to evaluate semantic similarity between documents more effectively, making them particularly suitable for news recommendation tasks [5].

The purpose of this study is to investigate and compare embedding-based approaches for generating personalized news recommendations, with a focus on both recommendation quality and system performance.

The recommendation mechanism is based on cosine similarity (1), which quantifies the similarity between two vectors in the embedding space. To assess the performance of the recommendation system, standard evaluation metrics are applied, such as Precision (2), Recall (3) and F1 (4).

$$\cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} \quad (1)$$

where:

*A* and *B* denote vector representations of news articles in the embedding space; the numerator represents the scalar (dot) product of the vectors, reflecting their directional alignment;  $\|A\|$  and  $\|B\|$  denote the Euclidean norms (magnitudes) of the corresponding vectors;  $\theta$  is the angle between the vectors in the high-dimensional space; the cosine similarity value lies within the interval  $[-1, 1]$ , where values closer to 1 indicate a higher degree of semantic similarity between articles.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

where:

*TP* (True Positive) denotes the number of relevant articles correctly identified by the recommendation system; *FP* (False Positive) represents the number of irrelevant articles that were incorrectly recommended; Precision reflects the proportion of relevant items among all recommended articles and characterizes the accuracy of the system.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

where:

*TP* (True Positive) denotes correctly recommended relevant articles; *FN* (False Negative) corresponds to relevant articles that were not retrieved by the system; Recall measures the ability of the system to identify all relevant items within the dataset.

$$F1 = 2 * \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

where:

Precision evaluates the correctness of the recommended results; Recall reflects the completeness of relevant item retrieval; the F1-score represents the harmonic mean of Precision and Recall and provides a balanced assessment of the system's performance, especially in cases where class distribution is uneven.

The experimental evaluation was conducted on a dataset consisting of over

3,000 news articles covering multiple thematic categories, including politics, technology, sports, culture, and business. To ensure reliable assessment of recommendation quality, a subset of 300 article pairs was manually annotated by independent evaluators. The annotation process was based on semantic, contextual, and stylistic relevance, and the final ground truth labels were assigned using majority voting, which increased the robustness of the evaluation dataset.

Three embedding-based approaches were investigated in this study: OpenAI Embeddings, Sentence Transformers (all-MiniLM-L6-v2), and Gemini Embeddings (text-embedding-001). All models were integrated into a unified recommendation pipeline in order to eliminate potential architectural biases and ensure a fair and consistent comparison. The same preprocessing steps, similarity computation method (cosine similarity), and evaluation protocol were applied across all experiments.

The comparative results of the evaluated embedding models are illustrated in Figure 1, where both accuracy metrics and response time are presented in a unified visualization format.



Fig. 1. Performance comparison of embedding models across accuracy metrics and response time in milliseconds

**Conclusion.** The findings of this study emphasize that the choice of an embedding model for personalized news recommendation systems should be determined by the specific operational requirements and constraints of the platform. The experimental results demonstrate clear trade-offs between recommendation accuracy, response time, and computational efficiency across the evaluated approaches.

OpenAI Embeddings achieved the highest performance in terms of Precision, Recall, and F1-score, confirming their strong capability to capture nuanced semantic

relationships within news content. This makes them particularly suitable for applications where recommendation quality and deep personalization are the primary objectives, such as large-scale news platforms with diverse and extensive content collections. However, their reliance on external API access introduces additional latency and operational costs.

Sentence Transformers provide a well-balanced solution, combining competitive accuracy with faster response times and the advantage of local deployment. Their relatively lightweight architecture and independence from external services make them an effective choice for systems that require scalability, cost-efficiency, and stable performance under varying workloads.

Gemini Embeddings demonstrated the lowest response time, highlighting their suitability for real-time and high-load environments where rapid content delivery is critical. At the same time, their lower accuracy metrics indicate a trade-off, suggesting that such models are more appropriate for scenarios where speed is prioritized over precise personalization.

Overall, the study confirms that embedding-based approaches significantly enhance the effectiveness of news recommendation systems compared to traditional methods. The results also highlight the importance of aligning model selection with application-specific priorities, whether they involve accuracy, efficiency, or latency.

Future research may focus on the development of hybrid recommendation strategies that combine embedding-based methods with collaborative filtering or graph-based techniques, as well as the integration of multilingual and multimodal data to further improve recommendation quality in modern digital news ecosystems.

## References:

1. H. B. Dergi and M. B. Akgün, "Social and Categorical Signals in Contrastive Learning for Recommendation Systems," in 2023 31st Signal Processing and Communications Applications Conf. (SIU), Istanbul, Turkiye, 2023, pp. 1–4. URL: <https://doi.org/10.1109/SIU59756.2023.10223958>
2. J. Yu et al., "Collaborative Filtering Recommendation with Fluctuations of User' Preference," in 2021 IEEE Int. Conf. on Information Communication and Software Engineering (ICICSE), Chengdu, China, 2021, pp. 222–226. URL: <https://doi.org/10.1109/ICICSE52190.2021.9404120>
3. J. Neve, "Content-Based Filtering," in Reciprocal Recommender Systems, SpringerBriefs in Computer Science. Cham: Springer, 2025. URL: [https://doi.org/10.1007/978-3-031-85103-2\\_4](https://doi.org/10.1007/978-3-031-85103-2_4)
4. Y. Chen et al., "Contrastive Graph Learning for Session-based Recommendation," in 2022 4th Int. Conf. on Artificial Intelligence and Advanced Manufacturing, Hamburg, Germany, 2022, pp. 139–142. URL: <https://doi.org/10.1109/AIAM57466.2022.00033>
5. K. Smelyakov, A. Chupryna, D. Darahan, and S. Midina, "Effectiveness of modern text recognition solutions and tools for common data sources," in CEUR Workshop Proc., vol. 2870, pp. 154–165, 2021. [Online]. URL: <https://ceur-ws.org/Vol-2870/>