Осипов Данило Олексійович
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 AN ONLINE PERFUME STORE

In modern e-commerce, recommender systems are an important tool for personalization and decision support. With the growth of data volumes, the search for relevant products becomes more difficult, especially in areas such as perfumery, where the choice depends on subjective factors. Traditional approaches often do not provide sufficient accuracy.

Modern artificial intelligence technologies, in particular embedding models, allow you to represent products in the form of vectors and determine their semantic similarity using metrics such as cosine. This allows you to better take into account the content of descriptions and hidden connections between products.

The relevance of the study is due to the need to implement effective intelligent methods to improve the quality of recommendations. Comparing embedding models is important for choosing the optimal solution taking into account accuracy and speed.

In the field of e-commerce, recommender systems are widely used to increase conversion and personalization. The most common are collaborative and content filtering. Collaborative filtering is based on the analysis of user behavior, but has the problems of "cold start" and data sparsity [1]. Content filtering uses product characteristics and allows you to recommend new objects, but limits the variety of recommendations and depends on the quality of metadata [2].

To overcome these shortcomings, hybrid approaches (weighted, cascaded, contextual) are used, which combine different methods [3]. However, traditional systems often do not take into account semantic relationships between objects.

Modern approaches are based on the use of embeddings and deep learning. The use of LLM and vector databases is relevant for complex domains, such as perfumery, where text descriptions and subjective factors play an important role.

The main problems remain the heterogeneity and sparsity of data, the complexity of modeling dynamic user preferences, as well as the high cost of implementing modern AI solutions.

Thus, the current task is to develop a recommendation system that combines accuracy, adaptability and efficiency when working with heterogeneous data in e-commerce environments.

The aim of the work is to study and compare methods for building recommendation systems based on embedding models for an online perfume store.

The research methods include domain analysis, comparative analysis of embedding models, vector and cosine similarity methods, as well as experimental evaluation of the effectiveness of recommendation systems.

To achieve the goal of creating a context-sensitive recommendation system for an online perfume store, the following steps were performed: combining all product information into a single text representation and converting it into vector representations using embeddings models [4].

Recommendations are formed by calculating the cosine similarity (1) between products and ranking the results in descending order of this score. Precision (2), Recall (3), and F1-score (4) metrics are used to assess effectiveness [5].

$$cos\,(\theta) = \frac{A*B}{\|A\|\|B\|} \tag{1}$$

where:

A and B are feature vectors of the two objects we are comparing. Vector A could represent the perfume selected by the user, and vector B could represent another fragrance from the catalog

$$Precision = \frac{\sum TP}{\sum (TP+FP)} \tag{2}$$

where:

TP - a true positive occurs when the system correctly predicts the positive class, meaning it successfully identifies a relevant item. FP - a false positive occurs when the system incorrectly predicts the positive class, suggesting an item that is actually irrelevant. FN -a false negative describes a situation where the system fails to recognize a positive class, effectively missing a relevant opportunity. TN - a true negative occurs when the system correctly predicts the negative class, filtering out irrelevant items (Fig. 1).

$$Recall = \frac{\sum TP}{\sum (TP+FN)} \tag{3}$$

where:

TP - a true positive. FP - a false positive. FN -a false negative. TN - a true *negative.*

F1-score – the harmonious average between Precision and Recall.

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

where:

Precision – the proportion of relevant recommendations among those suggested. Recall – the proportion of relevant items presented to the user.

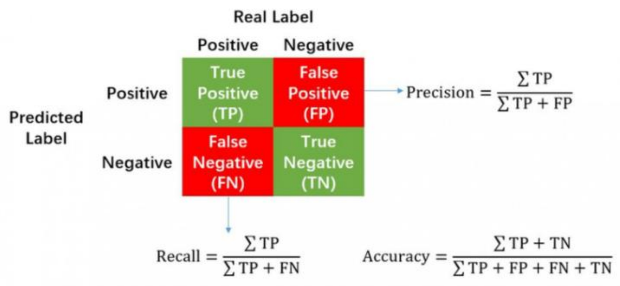

Fig. 1. **Confusion matrix and evaluation metrics in binary classification**

The system is implemented on the basis of an online store using the following stack: PHP (Laravel), MySQL and API embeddings models (OpenAI - text-embedding-3-small, Gemini - text-embedding-004, Ollama - nomic-embed-text).

The experimental study was conducted on a dataset of over 8,000 products. The data underwent pre-processing and were converted into vectors via the corresponding APIs. Based on the calculation of cosine similarity, ranked lists of recommendations were formed, which were evaluated according to the specified metrics. The response is the average time it takes to process a recommendation request.

Figure 2 shows the comparative performance of the OpenAI, Gemini, and Ollama embeddings based on three evaluation metrics: Accuracy, Congruence, and F1-score (Fig. 2).

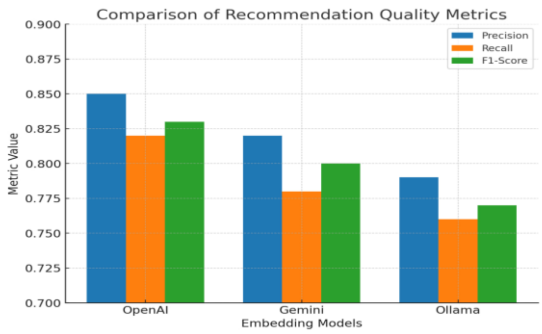

Fig. 2. **Comparison of recommendation quality indicators in different embedding models**

Conclusion. The subject area analysis was conducted. It was found that perfume products are characterized by a high level of subjectivity of choice and complexity of formalization of characteristics.

The architecture of the recommender system was developed and implemented. A hybrid mechanism was implemented that combines semantic proximity of products (Content-Based) with heuristic reinforcement (Boosting) based on brand and category. Integration with three vector providers was provided: OpenAI, Google Gemini and the local Ollama solution.

An experimental study of the effectiveness of the models was performed. A comparative analysis of three embedding models on a sample of 8000 products showed the following results:

OpenAI demonstrated the best accuracy and completeness indicators. The model most effectively takes into account the semantic context and features of perfume descriptions, however, it is characterized by higher API costs and longer query processing time.

Gemini showed balanced performance. It is slightly inferior to OpenAI in accuracy, but provides faster response times, making it an effective compromise solution for medium-scale systems.

Ollama provided the best performance (the lowest response time) and zero usage costs. Despite the slightly lower accuracy of recommendations, this model is optimal for highly loaded systems or conditions of limited computing resources.

References:

1. Zhang, Y., Feng, F., Zhang, J., Bao, K., Wang, Q., & He, X. (2025) CoLLM: Integrating Collaborative Embeddings Into Large Language Models for Recommendation. IEEE Transactions on Knowledge and Data Engineering, 37(5), 2329–2340. Retrieved from: https://doi.org/10.1109/TKDE.2025.3540912.
2. Li, X. (2025) Design of a Web Content Personalized Recommendation System Based on Collaborative Filtering Improved by Combining k-Means and LightGBM. Journal of Web Engineering, 24(2), 267–290. Retrieved from: https://doi.org/10.13052/jwe1540-9589.2425.
3. Çano, E., & Morisio, M. (2019) Hybrid recommender systems: A systematic literature review. Intelligent Data Analysis, 21, 1487–1524. Retrieved from: https://doi.org/10.3233/IDA-163209.
4. Smelyakov, K., Karachevtsev, D., Kulemza, D., Samoilenko, Y., Patlan, O., & Chupryna, A. (2020) Effectiveness of Preprocessing Algorithms for Natural Language Processing Applications. 2020 IEEE International Conference on Problems of Infocommunications. Science and Technology (PIC S&T), 187–191. Retrieved from: https://doi.org/10.1109/PICST51311.2020.9467919.
5. Xia, P., Zhang, L., & Li, F.-Z. (2015) Learning similarity with cosine similarity ensemble. Information Sciences, 307, 39–52. Retrieved from: https://doi.org/10.1016/J.INS.2015.02.024.