

УДК 004.94

ГІБРИДНИЙ МЕТОД ПЕРСОНАЛІЗАЦІЇ РЕКОМЕНДАЦІЙ НА ОСНОВІ ПОВЕДІНКОВИХ ДАНИХ КОРИСТУВАЧІВ

Синьова В.О.

e-mail: viktoriiia.synova@nure.ua

Науковий керівник – к.т.н., доц. Білова Т.Г.

Харківський національний університет радіоелектроніки, каф. ІУС
м. Харків, Україна

The report considers a hybrid method for personalizing recommendations based on user behavioral data in e-commerce systems. The proposed approach combines collaborative filtering, content-based filtering and behavioral data analysis to improve the relevance of recommendations. Special attention is paid to the use of implicit user interaction signals and the dynamic adjustment of algorithm weights depending on the stage of the IT project. The results demonstrate the potential of the hybrid approach to increase recommendation accuracy and improve user experience in digital commerce platforms.

Сьогодні інтернет-магазини працюють із величезним асортиментом товарів і користувачам дедалі важче орієнтуватися в цьому потоці. Інформаційне перевантаження стає серйозною проблемою, тому що людина може просто не знайти потрібний товар, навіть якщо він є в каталозі. У таких умовах персоналізовані рекомендації перетворюються з приємної опції на необхідний інструмент, який допомагає утримувати увагу клієнтів, підвищувати конверсію та формувати лояльність до платформи. Однак традиційні методи, такі як колаборативна фільтрація або контентний підхід, окремо вже не дають тієї точності, яка потрібна бізнесу сьогодні [1]. Вони добре працюють в ідеальних умовах, але на практиці, особливо на ранніх етапах життя проєкту, наприклад, під час запуску мінімально життєздатного продукту, стикаються з тим, що даних мало, вони розріджені, а нові користувачі або товари не мають історії взаємодій, тобто має місце так звана проблема «холодного старту».

Одним із найперспективніших способів підвищити якість рекомендацій є використання гібридних методів, які поєднують сильні сторони різних алгоритмів. Завдяки цьому можна компенсувати недоліки кожного з них і досягти кращого результату.

У даному дослідженні робиться акцент на поведінкових даних користувачів. Це не тільки явні дії, такі як оцінки, відгуки та покупки, а й неявні сигнали, тобто, які сторінки людина відкриває, скільки часу проводить на картці товару, що шукає, у якому порядку переглядає категорії. Саме такі неявні дані часто дають більш об'єктивну картину інтересів, ніж явні оцінки, які користувачі ставлять неохоче [2].

У даній роботі розглядається гібридний метод, який має бути позбавлений недоліків «чистих» алгоритмів і підходити для різних етапів

життєвого циклу продукту. Він поєднує три компоненти: матрицю поведінкових, тобто неявних даних; контентну подібність у вигляді метаданих товарів; колаборативну фільтрацію. Головною особливістю є динамічне зважування цих компонент, тобто використовуються не фіксовані ваги, а динамічне налаштування залежно від того, на якій стадії розвитку перебуває проєкт і скільки даних уже накопичено про конкретного користувача. У таблиці 1 показано, як змінюється роль кожного компонента залежно від стадії IT-проєкту.

Таблиця 1 – Використання компонентів гібридного методу на різних стадіях IT-проєкту

Стадія IT-проєкту	Доступність поведінкових даних	Основний метод рекомендацій	Додаткові джерела даних
Ініціація (MVP)	Низька	Контентно-орієнтована фільтрація	Метадані товарів, експертні правила
Проектування та розробка	Обмежена	Комбінування контентної та поведінкової інформації	Історія переглядів, кліки
Впровадження	Середня	Гібридний підхід	Поведінкові патерни користувачів
Експлуатація	Висока	Колаборативна фільтрація та факторизація матриць	Повна історія взаємодій користувачів

На етапі запуску платформи, коли проєкт тільки виходить на ринок, поведінкових даних ще обмаль, тому головну роль відіграють контентна фільтрація та експертні правила, наприклад, рекомендації на основі популярності або сезонності. Коли проєкт переходить у фазу зростання і накопичується достатня історія взаємодій, метод автоматично збільшує вагу колаборативної складової, використовуючи алгоритми факторизації матриць, які добре виявляють приховані зв'язки між користувачами та товарами.

Щодо обробки поведінкових патернів, то окрім «грубих» сигналів, таких як купівля та клік, даний аналіз враховує послідовність дій, тобто навігаційні ланцюжки, та час утримання на картці товару, що дозволяє точніше зрозуміти, що саме цікавить людину. Наприклад, якщо користувач довго вивчає дані про книгу, але не додає її в кошик, система інтерпретує це як високу зацікавленість, але з ціновим бар'єром.

Для таких ситуацій закладено логіку «відкладеного попиту», тобто рекомендації зі схожої категорії, але з товарами нижчого цінового сегмента, з'являються не одразу, а через деякий час, щоб не дратувати користувача нав'язливістю, але водночас не втратити його потенційний інтерес.

Також розроблено архітектуру модуля рекомендацій, який можна додати до вже існуючого IT-проєкту без кардинальної зміни бекенда. Його

спроєктовано як мікросервіс із власною базою даних для зберігання агрегованих профілів користувачів. Таке рішення дозволяє горизонтально масштабувати систему, коли проєкт виходить на стадію експлуатації й навантаження зростає. Крім того, сервіс підтримує потокову обробку даних, щоб рекомендації могли оновлюватися в реальному часі, тобто одразу після того, як користувач виконав якусь дію на сайті.

Ефективність методу перевірено на даних умовного інтернет-магазину, використовуючи журнали взаємодій користувачів за півроку. Результати показали, що даний гібридний підхід із динамічним зважуванням підвищує точність precision@5 приблизно на 12–15% порівняно з використанням лише колаборативної фільтрації. Найбільший вигреш метод дав для «холодних» товарів, тобто тих, що не мають історії покупок [3].

У даному випадку precision@5 є часткою релевантних рекомендацій серед перших п'яти, що запропоновані користувачеві [4]. Іншими словами, даний метод у середньому збільшив кількість релевантних товарів у топ-5 на 12–15 відсоткових пунктів.

Таким чином, поєднання різних методів із врахуванням детальної поведінки користувачів є перспективним шляхом. Запропонований метод не тільки робить рекомендації точнішими, а й дозволяє системі гнучко реагувати на зміни в поведінці користувачів і бізнес-цілях проєкту на різних стадіях. У майбутньому варто додати механізми онлайн-навчання, щоб модель могла адаптуватися до нових трендів без необхідності щодобового перетренування.

Список використаних джерел:

1. Zhou K., Wang, T. Personalized Interiors at Scale: Leveraging AI for Efficient and Customizable Design Solutions // Human-Computer Interaction. 2024. No. 1. P. 1–18. DOI: <https://doi.org/10.48550/arXiv.2405.19188>.
2. Петров К., Кобзев І. Метод визначення переважності споживчих характеристик товару на основі інформації про його вибір покупцем. Системи управління, навігації та зв'язку. Збірник наукових праць. 2022. Т. 2, № 68. С. 68–74. DOI: <https://doi.org/10.26906/sunz.2022.2.068>.
3. Ms. Tejashri Sharad Phalle. Content Based Filtering And Collaborative Filtering: A Comparative Study. Journal of Advanced Zoology. 2024. Vol. 45. P. 96–100.
4. Ricci F., Rokach L., Shapira B. Recommender Systems Handbook. 3rd ed. New York, NY : Springer, 2022. 1071 p.