

УДК 004.9:339

ДОСЛІДЖЕННЯ МЕТОДІВ ПРОЄКТУВАННЯ СИСТЕМ РЕКОМЕНДАЦІЇ ТОВАРІВ

Новіков М. В.

Науковий керівник – к.т.н., доц. Міщеряков Ю. В.

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

e-mail: mykyta.novikov@nure.ua

The work provides a definition of the concept of a recommendation system, discusses the problem of increasing the efficiency of recommendation systems, examines the main criteria of efficiency, and investigates the main classes of recommendation systems. The concept of hybrid recommendation systems is introduced, detailing their features and advantages, discussing the main tasks of designing hybrid systems, conducting an analysis of the task of improving speed, and addressing the cold start problem.

В контексті електронної комерції рекомендаційна система – інструмент, можливості котрого полягають в аналізі великих масивів даних з метою надання клієнтам персональних рекомендацій. Рекомендаційна система дозволяє на основі інформації про об'єкти, суб'єкти та результати діяльності підприємства робити певні прогнози з вподобань для конкретних клієнтів. Система аналізує попередню взаємодію між клієнтом та товаром – та прогнозує показники взаємодії в майбутньому [1]. Дані системи можуть бути реалізовані за різними принципами та алгоритмами, або ж поєднувати в собі більш ніж один з метою досягнення більш високих показників.

Для рекомендаційної системи основними критеріями ефективності є точність та швидкодія. Швидкість генерації рекомендацій зменшується зі зростанням кількості користувачів й обсягу даних, та в умовах експлуатації може досягати неприпустимих значень. З цієї причини проєктування системи в даній предметній області повинно виконуватися з орієнтацією на високий показник швидкості системи, з урахуванням факторів масштабованості та високонавантаженості. Дані питання у свою чергу передбачають оптимізацію не лише алгоритму, а й вирішення таких проблем, як висока розрідженість даних, зберігання допоміжних та проміжних даних для прискорення обчислень, тощо.

Точність у свою чергу є більш комплексною проблемою, що має в собі велику кількість підпроблем, однією з котрих є холодний старт – прогнозування в умовах відсутності достатньої кількості даних. Формування ж пропозицій у даних умовах у свою чергу може дуже негативно вплинути на ефективність системи при експлуатації. Більш того, проблема холодного старту є особливо характерною для предметної області продажу товарів – переважна більшість рекомендацій в даному

середовищі відбувається в умовах повного знаходження або близькості до холодного старту для користувача.

В області рекомендаційних систем можна виділити декілька різних за принципом класів простих систем, кожна має різні властивості та по-різному демонструє себе у різних умовах таких як холодний старт користувача, вкрай великі обсяги даних, постійно оновлення переліку товарів, тощо. Нижче наведені основні зі згаданих класів систем.

- Колаборативна фільтрація – формування рекомендацій на основі історії уподобань «схожих» користувачів, іншими словами аналіз взаємодії користувачів з товарами;

- Контентно-орієнтовані рекомендації – використання інформації про зміст товару, тобто певний набір його характеристик, з метою відбору прийнятних для профілю користувача з можливістю урахування певного характеристико-орієнтованого контексту;

- Методи на основі знань – використання експертних знань з предметної області з метою побудови з їх використанням рекомендацій на основі даних про покупців та товари.

Реалізація кожного з даних підходів може відрізнятись в залежності від інтерпретації, котра обрана при проєктуванні. Так, наприклад, система може бути представлена у вигляді довільного покрокового алгоритму, операцій над матрицями, класифікацією з використанням машинного навчання, тощо.

Кожен з описаних методів має свої особливості реалізації, власні недоліки та переваги при задіянні у різних умовах. Тому при проєктуванні рекомендаційних систем найбільш поширеною практикою є їх поєднання. У свою чергу системи, з довільною моделлю, що задіють більш ніж один алгоритм побудови рекомендацій, виділяються в окремий клас гібридних рекoмендаторів. Такий підхід дозволяє задіяти при формуванні рекомендацій найбільш широкий перелік даних різного характеру для більш глибокого аналізу та, відповідно, точніших рекомендацій. Задача ж проєктування гібридної системи полягає в обґрунтованому та ефективному поєднанні методів та практик в єдину модель з метою досягнення підвищення її загальних характеристик, досягнувши при цьому сприйнятливих показників оптимізації.

При проєктуванні такої моделі одним з основних питань є спосіб, у який поєднуються алгоритми. Гібридні рекомендаційні системи поділяються на класи в залежності від способу об'єднання окремих її компонентів. Так композиція компонентів може бути зваженою, змішаною, каскадною, з поєднанням ознак, з перемиканням, тощо [2]. Тому під час визначення математичної моделі прийнято звертати увагу ще й на складність того чи іншого компоненту схеми, на ступінь складності обчислень порівняльно до результату, що система отримує від застосування.

Головним же питанням, котре вирішується при проектуванні гібридних систем, є вибір алгоритмів. Необхідно визначити, які алгоритми слід комбінувати та як їх найкраще інтегрувати для досягнення оптимальних результатів. В контексті даної задачі не існує універсального рішення або методів, оскільки кожен набір даних та кожен контекст вимагають індивідуального підходу, а спрогнозувати ефективність роботи окремих алгоритмів як єдиної взаємопов'язаної системи неможливо без проведення відповідних експериментів. Стосовно алгоритмів формування рекомендацій, на даний момент відбувається зріст методів з використанням нейронних мереж. Зокрема, можна виділити такі алгоритми як «Wide & Deep», «DLRM», «DCN», «GNN», «DeepFM», «CVAE». Дані підходи задіють можливості машинного навчання, такі як стійкість до великих обсягів даних та вирішення комплексних проблем, та поєднують їх з класичними алгоритмами, завдяки чому система може ідентифікувати в значній мірі комплексні взаємозв'язки у взаємодії користувача з товаром [3], при цьому значно зменшується складність системи [4]. Описані властивості даних алгоритмів у результаті дають змогу досягнути відносно високих показників точності та дозволяє знизити вплив холодного старту на ефективність рекомендаційної мережі.

Проведені дослідження виявили, що найбільш доцільним є використання контентно-колаборативного гібриду з персоналізацією на основі контекстуальних ознак, за змішаним типом архітектури. Для рішення проблеми розріджених даних найбільш продуктивним виявилися такі підходи як матрична факторизація та алгоритм на основі механізму уваги. В умовах холодного старту найбільші показники має підхід на основі глибинного навчання. Методи на основі моделі демонструють меншу похибку в умовах холодного старту, ніж традиційні алгоритми, та в силу великих обсягів розріджених даних займають значно менше часу для обчислення.

Список використаних джерел:

1. Aggarwal C. C. *Recommender Systems*. Cham : Springer International Publishing, 2016. URL: <https://doi.org/10.1007/978-3-319-29659-3> (date of access: 04.03.2024).
2. Burke R. *User Modeling and User-Adapted Interaction*. 2002. Vol. 12, no. 4. P. 331–370. URL: <https://doi.org/10.1023/a:1021240730564> (date of access: 04.03.2024).
3. *Deep Learning Based Recommender System* / S. Zhang et al. *ACM Computing Surveys*. 2019. Vol. 52, no. 1. P. 1–38. URL: <https://doi.org/10.1145/3285029> (date of access: 04.03.2024).
4. Гребенюк М., Ситнікова П. Е. Компактна гібридна модель користувача для покращення рекомендаційних систем // Перспективні напрямки сучасної електроніки, інформаційних і комп'ютерних систем (MEICS-2023) : Тези доповідей на VIII Всеукраїнській науково-практичній конференції (22–24 листопада 2023 р.). Дніпро, 2023. С. 106-107.