

ДОСЛІДЖЕННЯ І РОЗРОБКА МЕТОДІВ ТА АЛГОРИТМІВ РЕКОМЕНДАЦІЙНИХ СИСТЕМ

Антоненко О. С., Смілянець В. Р.

Одеський національний університет І.І. Мечникова

Ключові слова: рекомендаційні системи, глибоке навчання, холодний старт, колаборативна фільтрація, фільтрація на основі вмісту

В останні роки використання систем рекомендацій стає все більш популярним, особливо в сфері електронної комерції та онлайн-платформ. Такі системи представляють собою тип штучного інтелекту, який надає рекомендації користувачам на основі їх попередньої поведінки, уподобань та інтересів. Однак ефективність рекомендаційних систем залежить від методів і алгоритмів, які використовуються для їх розробки, тому дослідження та розробка методів і алгоритмів рекомендаційних систем стали популярною та важливою темою в галузі штучного інтелекту [1].

Системи рекомендацій повсюдно використовуються різноманітними додатками, тому наприклад платформи електронної комерції, такі як Amazon, eBay і Alibaba, значною мірою покладаються на рекомендаційні системи, щоб пропонувати товари своїм користувачам; платформи соціальних медіа, такі як Facebook, Instagram, Twitter, LinkedIn – щоб показувати користувачам відповідні пости та друзів; розважальні платформи, такі як Netflix, Spotify і YouTube – щоб пропонувати своїм користувачам фільми, пісні та відео; освітні платформи, такі як Coursera, edX і Udemy – щоб рекомендувати курси своїм учням [1, 3].

Існують різні підходи та методи проектування рекомендаційної системи. Наприклад колаборативна фільтрація передбачає використання минулої поведінки користувачів, такі як оцінки або історія покупок. Фільтрація на основі вмісту використовує атрибути елемента, такі як жанр або ключові слова. Гібридні підходи поєднують попередні підходи. Методи матричної факторизації також часто використовуються для виявлення прихованих шаблонів у взаємодії між користувачем і елементом. А також нещодавні досягнення в глибокому навчанні, такі як нейронні мережі були застосовані до систем рекомендацій для покращення продуктивності [2].

Використання рекомендаційних систем перетинається з різними проблемами, які впливають на точність отриманого результату. Наприклад

труднощі пов'язані з «холодним стартом» або роботу з розрідженими та шумними даними [3]. Якість результату роботи рекомендаційних систем напряму тісно пов'язані з рівнем продажів платформи, а також задоволення користувача. Тому метою роботи є покращення ефективності рекомендаційної системи для конкретного набору даних (наприклад MovieLens [4] та MyAnimeList [5]). Набори містять інформацію про фільми/аніме (жанри, студії випуску, тривалість та ін.), які будуть застосовуватись методами на основі вмісту; інформацію про користувачів (місце проживання, вік та ін.), які будуть корисні для знаходження схожих користувачів; а також найважливіша інформація - попередній досвід взаємодії користувачів з продуктами. Такі дані можуть бути представлені у вигляді матриці користувач-продукт зі значеннями оцінки. Також для побудування більш детальних рекомендацій записи досвіду містять дату та інші контексті дані. Тут і трапляються описані труднощами розрідженості та шумності, адже більшість значень матриці будуть пустими, а також вони не в повній мірі можуть представляти реальний досвід: користувач, може переглянути фільм, але не залишити відгук.

Об'єктом дослідження є рекомендаційні системи, а предметом – методи та алгоритми їх розробки та вдосконалення. У дослідженні будуть використані та порівняні різні методи. Дослідження також використовуватиме різні показники оцінки, такі як precision, recall, F1-score та RMSE для вимірювання ефективності моделей.

Література

1. Dietmar Jannach, Measuring the Business Value of Recommender Systems / Dietmar Jannach, Michael Jugovac // ACM Transactions on Management Information Systems, Vol. 10, No. 4, Article 1. Publication date: December 2019. – Access mode: <https://arxiv.org/pdf/1908.08328.pdf>
2. Yehuda Koren, Matrix factorization techniques for recommender systems / Yehuda Koren, Robert Belz, Chris Volinsky / IEEE Computer Society. August 2009 – Access mode: [https://datajobs.com/data-science-repo/Recommender-Systems-\[Netflix\].pdf](https://datajobs.com/data-science-repo/Recommender-Systems-[Netflix].pdf)
3. Dietmar Jannach, Recommender Systems. An Introduction / Dietmar Jannach, Markus Zanker, Alexander Felfernig, Gerhard Friedrich // Cambridge University Press, 2011 – Access mode: <https://arxiv.org/pdf/1908.08328.pdf>
4. MovieLens 100K Dataset | Kaggle [Електронний ресурс] - <https://www.kaggle.com/datasets/prajitdatta/movielens-100k-dataset>
5. MyAnimeList Dataset | Kaggle [Електронний ресурс] - <https://www.kaggle.com/datasets/azathoth42/myanimelist>