

ДОСЛІДЖЕННЯ ВИКОРИСТАННЯ АЛГОРИТМІВ ДЛЯ ПЕРСОНАЛІЗОВАНИХ РЕКОМЕНДАЦІЙ НАСТІЛЬНИХ ІГОР

Цвик В.І., Посукан І.І., Ляпота В.М.

e-mail: vladyslav.tsvyk@nure.ua, inna.posukan@nure.ua,

vitaliy.lyapota@nure.ua

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

This work is devoted to the analysis of recommendation algorithms for board games based on the content-based approach. The study explores various recommendation methods, including vector-based similarity calculations and hybrid techniques to improve accuracy. The problem of cold start and data normalization is addressed, ensuring better personalization for users. Additionally, the effectiveness of the proposed system is assessed using precision, recall, and F1-score metrics. The research highlights the importance of combining multiple filtering techniques for enhancing recommendation relevance.

На сьогоднішній день сучасні програмні системи все частіше включають в себе рекомендаційні підсистеми для персоналізації контенту та підвищення залученості користувачів. Вони використовуються в різних сферах, зокрема в електронній комерції, пошукових сервісах, соціальних мережах та освітніх платформах. В останні роки зростає інтерес до впровадження рекомендаційних алгоритмів у сфері настільних ігор, що обумовлено значною різноманітністю ігрового контенту та необхідністю допомагати користувачам у виборі відповідних ігор. Традиційні методи пошуку нових настільних ігор, такі як рейтинги або відгуки, мають певні обмеження, оскільки вони не враховують індивідуальні вподобання гравців та їхній попередній досвід.

Основною проблемою у створенні такої системи є велика кількість параметрів, що характеризують ігри, а також необхідність ефективного аналізу вподобань користувачів. У цьому дослідженні розглядаються методи побудови рекомендаційних систем для настільних ігор, зокрема підходи, засновані на контентній фільтрації (Content-based Filtering) [1], а також можливості поєднання різних алгоритмів для підвищення точності рекомендацій.

Серед основних методів для побудови рекомендаційних систем можна виділити наступні:

– контентна фільтрація (Content-Based Filtering) – рекомендація об'єктів (настільних ігор) на основі їхніх характеристик та схожості з тим, що вже сподобалося користувачу;

– колаборативна фільтрація (Collaborative Filtering) – рекомендації, що базується на даних про взаємодію користувачів з об'єктами, не враховуючи їх характеристик;

– гібридні методи (Hybrid) – поєднання контентної та колаборативної фільтрації для забезпечення більшої релевантності рекомендації.

При відсутності великої кількості даних про взаємодію користувачів варто використовувати Content-Based Filtering [2], наприклад, у випадку запуску системи, де початковими даними є лише дані про настільні ігри та їх характеристики. У такій програмній системі варто передбачити можливість оцінювати об'єкти взаємодії користувачами. Отримавши такі дані можна застосувати методи Collaborative Filtering і об'єднавши їх з попередніми алгоритмами сформуванню гібридну рекомендаційну систему, яка б поєднувала обидва підходи.

При застосуванні Content-Based Filtering потрібно враховувати такі властивості настільних ігор, як популярність, складність, ігрові механіки, тематика, категорії, вікові обмеження, тривалість партії, кількість гравців, залежність від знання мови (кількість важливої текстової інформації). Також необхідно врахувати попередній досвідом гравця, а саме зі список ігор, які сподобалися користувачу. Одним з джерел даних, яке б містило може стати сервіс BoardGameGeek, де зібрано дані більше ніж про 20000 настільних ігор [3].

Кожна настільна гра може бути представлена у вигляді векторного профілю, що містить її характеристики. Основне завдання алгоритму – порівняти ці профілі з іграми, які вже сподобалися користувачу, та знайти найбільш схожі варіанти. Для цього спочатку будується матриця ознак, де кожен рядок відповідає певній грі, а стовпці містять значення таких параметрів, як складність, тематика, механіки, категорії, тривалість партії, популярність тощо.

Одним із найпоширеніших способів обчислення схожості між іграми є використання косинусної подібності (cosine similarity) [4] між їхніми векторними представленнями. Косинусна подібність визначає ступінь схожості між двома векторами за допомогою косинуса кута між ними. Чим ближче значення до 1, тим більше схожі ігри. Інший підхід – евклідова відстань, яка визначає наскільки близько розташовані дві гри у багатовимірному просторі ознак. Для категоріальних ознак (наприклад, механіки, тематика, жанри) можна використовувати методи One-Hot Encoding або TF-IDF для текстових описів, що дозволяє будувати багатовимірне представлення гри та порівнювати їх між собою.

Важливим аспектом є нормалізація числових параметрів, таких як рейтинг, складність чи тривалість партії, оскільки вони мають різні масштаби. Наприклад, рейтинг може мати значення від 1 до 10, тоді як тривалість партії – від 5 до 300 хвилин. Щоб уникнути домінування окремих параметрів, застосовується min-max нормалізація або z-score стандартизація.

Одним із викликів цього підходу є проблема холодного старту, коли у нового користувача немає історії вподобань, і система не має достатньо інформації для формування рекомендацій. У таких випадках можна

комбінувати Content-Based Filtering з іншими методами, наприклад, використовувати популярні ігри з високими рейтингами або враховувати схожість з профілями інших користувачів (елементи колаборативної фільтрації).

Після впровадження рекомендаційної системи варто визначити, наскільки добре розроблений алгоритм пропонує користувачам релевантні ігри. Для цього можна використовувати такі метрики, як Precision, Recall та F1-score, які аналізують, яка частка рекомендованих ігор дійсно цікава користувачу (Precision) і наскільки повно система покриває всі можливі релевантні варіанти (Recall). F1-score комбінує ці два показники, дозволяючи збалансовано оцінити модель.

Для реалізації описаних вище алгоритмів Content-Based Filtering існує декілька підходів і технологій які можуть бути використані.

По-перше, можна використовувати мову Python, бібліотеки для машинного навчання та роботи з аналізом даних, такі як Scikit-learn, Pandas та NumPy.

По-друге, для платформи .NET існують бібліотеки ML.NET та MathNet для роботи з машинним навчанням та великими числовими даними.

По-третє, існують хмарні (cloud) рішення, такі як Google Cloud AI чи Azure AI для розробки та навчання моделей машинного навчання [5].

Отже, при проектуванні програмної системи рекомендації настільних ігор варто ретельно підійти до проектування та вибору алгоритмів. У випадку відсутності даних про взаємодію користувача з об'єктами варто використовувати Content-Based Filtering, який у процесі розвитку системи можна скомбінувати з Collaborative Filtering.

Список використаних джерел:

1. Ms. Tejashri Sharad Phalle. Content Based Filtering And Collaborative Filtering: A Comparative Study. *Journal of Advanced Zoology*. 2024. Vol. 45. P. 96–100.
2. Marwa Hussien Mohamed, Mohamed Khafagy, Mohamed Hasan Ibrahim. Recommender Systems Challenges and Solutions Survey. *International Conference on Innovative Trends in Computer Engineering*. URL: <http://dx.doi.org/10.1109/ITCE.2019.8646645> (date of access: 04.03.2025).
3. Phil Woodward, Sam Woodward. Mining the Boardgamegeek. *Significance*. 2019. Vol. 16, Issue 5. P. 24–29.
4. Muhammad Falah, Dewi Handayani. Content-based filtering using cosine similarity algorithm for alternative selection on training programs. *Journal of Soft Computing Exploration*. 2023. Vol. 4. P. 204 – 212.
5. Макеєв О., Кравець Н. Дослідження методів створення сервісно-орієнтованих програмних систем у Azure. *Computer Systems and Information Technologies*. 2023. №2(11). С. 38–47.