

НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ



ІНСТИТУТ ПРОБЛЕМ МОДЕЛЮВАННЯ
В ЕНЕРГЕТИЦІ ІМ. Г.С. ПУХОВА



**МАТЕРІАЛИ
У НАУКОВО-ПРАКТИЧНОЇ КОНФЕРЕНЦІЇ
«БЕЗПЕКА ЕНЕРГЕТИКИ В ЕПОХУ ЦИФРОВОЇ
ТРАНСФОРМАЦІЇ»**

22 листопада 2023 року

Київ – 2023

УДК [621.3+620.9]:[004[056.53+42+94] + 504.06]

ББК 31

Б-39

Рекомендовано до друку
Вченою радою Інституту
проблем моделювання в
енергетиці ім. Г.Є. Пухова НАН
України (протокол №10 від
30 листопада 2023 р.)

Б-39 **Безпека енергетики** в епоху цифрової трансформації, V науково-практична конференція Інституту проблем моделювання в енергетиці ім. Г.Є. Пухова Національної академії наук України : матеріали (Київ, 22 листопада 2023 р.). Київ : ІПМЕ ім. Г.Є.Пухова НАН України, 2023. 152 с.

В-39 **Energy security** in the digital transformation era, V scientific-practical conference of the G.E. Pukhov Institute for Modeling in Energy Engineering National Academy of Sciences of Ukraine : materials (Kyiv, November 22, 2023). Kyiv: PIMEE NAS of Ukraine, 2023. 152 p.

© Автори публікацій, 2023

© ІПМЕ ім. Г.Є.Пухова НАН України, 2023

**НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
ІНСТИТУТ ПРОБЛЕМ МОДЕЛЮВАННЯ В ЕНЕРГЕТИЦІ
ім. Г.Є. ПУХОВА НАН УКРАЇНИ**

**МАТЕРІАЛИ ІV НАУКОВО-ПРАКТИЧНОЇ
КОНФЕРЕНЦІЇ**

**БЕЗПЕКА ЕНЕРГЕТИКИ В ЕПОХУ ЦИФРОВОЇ
ТРАНСФОРМАЦІЇ**

**22 листопада 2023 року
м. Київ**

2023

Вельмишановний учасник _____

Запрошуємо Вас прийняти участь в роботі V науково-практичної конференції «Безпека енергетики в епоху цифрової трансформації», яка буде проходити 22 листопада 2023 року в Інституті проблем моделювання в енергетиці ім. Г.Є. Пухова Національної академії наук України (м. Київ).

ОРГАНІЗАТОРИ КОНФЕРЕНЦІЇ

Інститут проблем моделювання в енергетиці ім. Г.Є. Пухова НАН України
(м. Київ)

ПРОГРАМНИЙ КОМІТЕТ

Мохор Володимир Володимирович

член-кореспондент НАН України, доктор технічних наук, професор,
директор Інституту, голова програмного комітету

Чемерис Олександр Анатолійович

доктор технічних наук,
заступник директора з наукової роботи

Артемчук Володимир Олександрович

доктор технічних наук,
заступник директора з науково-організаційної роботи

Чьочь Вікторія Володимирівна

кандидат технічних наук,
т.в.о. заступника директора з науково-технічної роботи

ОРГАНІЗАЦІЙНИЙ КОМІТЕТ

Артемчук Володимир Олександрович

доктор технічних наук,
заступник директора з науково-організаційної роботи

Клименко Тетяна Михайлівна

завідувачка науково-організаційного відділу

Цуркан Оксана Володимирівна

молодший науковий співробітник

Потенко Олександр Сергійович

науковий співробітник

ПРОГНОЗУВАННЯ ПОВЕДІНКИ СПОЖИВАЧІВ З ВИКОРИСТАННЯМ ТЕХНОЛОГІЙ ШТУЧНОГО ІНТЕЛЕКТУ

У сучасному світі цифрової комерції, де споживачі зіткнулися з небувалою різноманітністю вибору, виникає складність у ефективному навігуванні по великих каталогах товарів. Це ставить перед користувачами задачу знаходити продукти, що відповідають їх індивідуальним потребам та вподобанням. І хоча ця різноманітність надає більше можливостей, вона також створює певні труднощі. Однією з ключових проблем є питання, як допомогти користувачам ефективно орієнтуватися у великих каталогах, щоб знайти те, що їм потрібно.

У відповідь на це, виникає необхідність у розробці систем рекомендацій, які виступають важливим рішенням цієї проблеми, збагачуючи досвід користувачів і одночасно оптимізуючи коефіцієнти конверсії продажів. Для вирішення цієї проблеми розглядається застосування системи рекомендацій, що використовує гібридний метод фільтрації. Цей підхід поєднує фільтрацію на основі вмісту, яка аналізує атрибути товарів, та колаборативну фільтрацію, яка враховує поведінку та вподобання користувачів, з метою створення більш точних та релевантних рекомендацій.

У рамках проекту здійснюється збір та попередня обробка даних, що включають інформацію про взаємодію користувачів із товарами, а також атрибути самих товарів. Використовуються техніки, такі як косинусна подібність для визначення схожості між товарами та сингулярний розклад матриць для аналізу користувацьких переваг. Особливу увагу приділено механізму «багаторукого бандита», який дозволяє експериментувати з різними варіантами рекомендацій та вибирати найефективніші, тим самим покращуючи точність рекомендацій та загальну корисність системи для користувачів.

Структура запропонованої гібридної системи рекомендацій

Структура гібридної системи рекомендацій (рис. 1) може бути описана таким чином:

1. збір даних: збір інформації про взаємодії користувачів та їх оцінки;
2. обробка даних: перетворення зібраних даних у вектори ознак для елементів та профілів користувачів;
3. контент-орієнтована фільтрація: аналіз властивостей елементів і порівняння їх з інтересами користувачів;
4. колаборативна фільтрація: аналіз взаємодій між користувачами та елементами для знаходження схожих зацікавлень;
5. інтеграція та комбінування: об'єднання результатів з обох підходів за допомогою важелевого комбінування;

6. видача рекомендацій: надання кінцевих рекомендованих елементів користувачам;

7. використання стратегії багаторукого бандита для оптимізації вагів.



Рисунок 1 – Загальна схема роботи гібридної системи з впровадженням стратегії багаторукого бандита

Для першої частини гібридної системи використовується фільтрація вмісту. Механізм рекомендацій працює шляхом порівняння характеристик елементів із профілем користувача та ранжування елементів на основі їхньої схожості з уподобаннями користувача. Кінцевим результатом є список рекомендованих елементів, які, швидше за все, зацікавлять користувача. Одним із основних способів досягти цього, особливо з текстовими даними, є косинусна подібність, яку часто отримують із представлень вмісту TF-IDF [1].

Основою для цього підходу є метрика TF-IDF (1) та (2). Вона кількісно визначає важливість слова в конкретному документі порівняно з його поширеністю в колекції документів. Обчислення складається з двох частин:

$$TF = \frac{n_i}{\sum_k n_k}, \quad (1)$$

де n_i – число входжень слова в документ;

n_k – кількість слів в документ.

$$IDF = \log \frac{|D|}{|(d_i \ni t_i)|}, \quad (2)$$

де $|D|$ – кількість документів колекції;

$|(d_i \ni t_i)|$ – кількість документів, в яких зустрічається слово t_i .

Після визначення ваги TF-IDF для всіх термінів у всіх документах кожен документ можна представити як вектор у багатовимірному просторі, де кожен вимір відповідає вазі TF-IDF терміна [1].

Далі, щоб оцінити подібність між будь-якими двома документами, використовується міра косинусної подібності. Ця метрика обчислює косинус кута між двома векторами (3).

$$\cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|}, \quad (3)$$

де $A \cdot B$ – скалярний добуток векторів;

$\|A\|$ та $\|B\|$ – скалярні значення цих векторів [2].

Для реалізації фільтрації вмісту використовується бібліотека `scikit-learn` вона може аналізувати текстові описи книг та обчислювати схожості між ними. Бібліотека використовує інструменти `TfidfVectorizer` для створення матриці TF-IDF на основі текстових описів книг та `linear_kernel` для обчислення косинусної подібності між книгами на основі цієї матриці.

Другою важливою частиною гібридної системи є колаборативна фільтрація яка базується на ідеї, що користувачі, які мають схожі зацікавлення чи вподобання в минулому, ймовірно, матимуть схожі вподобання й у майбутньому. Цей метод працює на основі аналізу взаємодії користувачів з продуктами, для цього зручно використовувати сингулярне розкладання матриці [3].

Припустимо, що є матриця R (4):

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \dots & r_{mn} \end{bmatrix}, \quad (4)$$

Де R – матриця з розміром $m \times n$, де m це кількість користувачів, n це кількість ознак товару;

r_{ij} – оцінка, яку користувач i дав товару j .

Декомпозицію матриці R можна представити наступним чином (5):

$$R = U \Sigma V^T, \quad (5)$$

де U – ортогональна матриця розміром $m \times m$, що представляє вектори ознак, які відповідають користувачам у прихованому просторі ознак;

Σ – це діагональна матриця розміром $m \times n$ з невід’ємними дійсними числами, відомими як сингулярні значення;

V^T – ортогональна матриця розміром $n \times n$, що представляє вектори ознак, що відповідають елементам у прихованому просторі ознак.

Після декомпозиції, можна виділити рейтинг \hat{r}_{ij} (6) користувача i , товару j :

$$\hat{r}_{ij} = \sum_{k=1}^K u_{ik} \sigma_k u_{jk}, \quad (6)$$

де k – кількість сингулярних значень, які необхідно зберегти [3].

Для колаборативної фільтрації використовується бібліотеки `scikit-learn` і `pandas`. Вона використовує метод головних компонент для зменшення розмірності матриці, обчислює схожість користувачів за допомогою `cosine_similarity` та рекомендує книги на основі замовлень схожих користувачів.

Ключовим етапом є об'єднання обох рекомендацій, що дозволяє вирішити обмеження, характерні для окремих типів систем, забезпечуючи більш високу якість та точність рекомендацій. Для гібридної системи було обрано важелеве комбінування результатів. Це означає, що якщо одна модель має більшу вагу, інша модель відповідно матиме меншу вагу [4].

Комбінований рейтинг R для колаборативної фільтрації та фільтрація на основі вмісту може бути обчислений за формулою:

$$R = w_1 R_{colab} + w_2 R_{content}, \quad (7)$$

де w_1 – вага колаборативного метода;

R_{colab} – рекомендаційний рейтинг колаборативної фільтрації;

w_2 – вага фільтрації на основі вмісту;

$R_{content}$ – рекомендаційний рейтинг фільтрації на основі вмісту;

Для забезпечення коректності, зазвичай ваги нормалізують, тобто: $w_1 + w_2 = 1$ [4].

Але в гібридній системі рекомендацій, досить складно обрати правильні ваги, тому використовується ідея «багатурукого бандита», де кожна «рука» відповідає певній комбінації ваг, що визначає баланс між колаборативними та контентними рекомендаціями. Для оптимального вибору цих ваг використовується стратегія верхньої довірчої межі, яка адаптивно регулює ці ваги залежно від накопиченого досвіду [5].

$$R_{ui}^{(s)} = w_1^{(s)} R_{colab,ui} + w_2^{(s)} R_{content,ui} + P(u)\lambda, \quad (8)$$

де $w_1^{(s)}, w_2^{(s)}$ представляють коефіцієнти користувачів u та товару i для вибраної руки s [5].

Стратегію багатурукого бандита застосовано до сайту, що є макетом інтернет-книгарні. В якості серверної частини використано Flask, досить потужний вебфреймворк. Для тестування, створено користувача, який купує різні книги, на основі яких йому пропонуються схожі товари. Приклад рекомендацій зображено на рисунку 2.

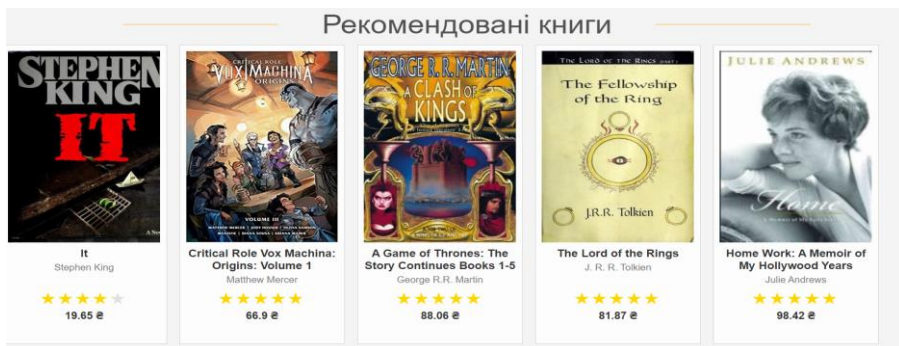


Рисунок 2 – Результат роботи рекомендаційної системи

1. Recommender System Series, Part 1: Introduction & Content-Based Filtering [Электронный ресурс]. – Режим доступа: <https://medium.com/encora-technology-practices/recommender-system-series-part-1-introduction-content-based-filtering-785ade61d2dc>.
2. Understanding Cosine Similarity in Recommendation Engines [Электронный ресурс]. – Режим доступа: <https://medium.com/@TechTalkWithAlex/understanding-cosine-similarity-in-recommendation-engines-3d98068f2b73>.
3. <https://medium.com/hackernoon/introduction-to-recommender-system-part-1-collaborative-filtering-singular-value-decomposition-44c9659c5e75>.
4. 7 Types of Hybrid Recommendation System [Электронный ресурс]. – Режим доступа: <https://medium.com/analytics-vidhya/7-types-of-hybrid-recommendation-system-3e4f78266ad8>.
5. Know all About UCB Algorithm in Reinforcement Learning [Электронный ресурс]. – Режим доступа: <https://www.turing.com/kb/guide-on-upper-confidence-bound-algorithm-in-reinforced-learning>.