

## ДОДАТОК А

## Звіт результатів перевірки на унікальність тексту в базі ХНУРЕ



## Звіт подібності

## метадані

Назва організації  
**Kharkiv National University of Radio Electronics**  
 Заголовок  
**2025\_M\_ПІ\_ІПЗ-23-1\_Огу\_С\_І\_скорочений**  
 Автор Науковий керівник / Експерт  
**Огу Стефні ІфомаЄвген Кардаш**  
 підрозділ  
**каф. ПІ**

## Обсяг знайдених подібностей

Коефіцієнт подібності визначає, який відсоток тексту по відношенню до загального обсягу тексту було знайдено в різних джерелах. Зверніть увагу, що високі значення коефіцієнта не автоматично означають плагіат. Звіт має аналізувати компетентна / уповноважена особа.



## Тривога

У цьому розділі ви знайдете інформацію щодо текстових спотворень. Ці спотворення в тексті можуть говорити про МОЖЛИВІ маніпуляції в тексті. Спотворення в тексті можуть мати навмисний характер, але частіше характер технічних помилок при конвертації документа та його збереженні, тому ми рекомендуємо вам підходити до аналізу цього модуля відповідально. У разі виникнення запитань, просимо звертатися до нашої служби підтримки.

Заміна букв		2
Інтервали		0
Мікропробіли		0
Білі знаки		0
Парафрази (SmartMarks)		24

## Подібності за списком джерел

Нижче наведений список джерел. В цьому списку є джерела із різних баз даних. Колір тексту означає в якому джерелі він був знайдений. Ці джерела і значення Коефіцієнту Подібності не відображають прямого плагіату. Необхідно відкрити кожне джерело і проаналізувати зміст і правильність оформлення джерела.

10 найдовших фраз		Копію тексту
ПОРЯДКОВИЙ НОМЕР	НАЗВА ТА АДРЕСА ДЖЕРЕЛА URL (НАЗВА БАЗИ)	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)
1	МН_ПЗМД_ІСТ_КПІ_2025_Шерстюк І.О. 5/20/2025 National Technical University of Ukraine Igor Sikorskyi Kyiv Politech Institute (ФІОТ, К-ра інформаційних систем та технологій)	47 0.40 %
2	<a href="https://openarchive.nure.ua/bitstreams/c4504c2d-5929-4d9f-91df-48e3b1a5aae6/download">https://openarchive.nure.ua/bitstreams/c4504c2d-5929-4d9f-91df-48e3b1a5aae6/download</a>	26 0.22 %

3	Дослідження методів створення персонального дашборду із впровадженням рекомендаційної системи.pdf 12/24/2023 Odessa National Polytechnic University (ІПТДМ, Каф. ІТПД)	25 0.21 %
4	МН_ПЗМД_ІСТ_КПІ_2025_Шерстюк І.О. 5/20/2025 National Technical University of Ukraine Igor Sikorskyi Kyiv Politech Institute (ФІОТ, К-ра інформаційних систем та технологій)	22 0.19 %
5	<a href="https://openarchive.nure.ua/bitstreams/c4504c2d-5929-4d9f-91df-48e3b1a5aae6/download">https://openarchive.nure.ua/bitstreams/c4504c2d-5929-4d9f-91df-48e3b1a5aae6/download</a>	19 0.16 %
6	МН_ПЗМД_ІСТ_КПІ_2025_Шерстюк І.О. 5/20/2025 National Technical University of Ukraine Igor Sikorskyi Kyiv Politech Institute (ФІОТ, К-ра інформаційних систем та технологій)	18 0.15 %
7	<a href="https://krs.chmnu.edu.ua/jspui/bitstream/123456789/3758/1/%D0%9A%D0%B2%D0%B0%D0%BB%D1%96%D1%84%D1%96%D0%BA%D0%B0%D1%86%D1%96%D0%B9%D0%BD%D0%B0%20%D1%80%D0%BE%D0%B1%D0%BE%D1%82%D0%B0%20%D0%84%D0%B2%D0%B4%D0%BE%D0%BA%D1%96%D0%BC%D0%BE%D0%B2.pdf">https://krs.chmnu.edu.ua/jspui/bitstream/123456789/3758/1/%D0%9A%D0%B2%D0%B0%D0%BB%D1%96%D1%84%D1%96%D0%BA%D0%B0%D1%86%D1%96%D0%B9%D0%BD%D0%B0%20%D1%80%D0%BE%D0%B1%D0%BE%D1%82%D0%B0%20%D0%84%D0%B2%D0%B4%D0%BE%D0%BA%D1%96%D0%BC%D0%BE%D0%B2.pdf</a>	17 0.14 %
8	2019_81260001_Antonyshchak_Serhii_Andriiovych_57720 10/25/2024 National University "Lviv Politechnika" (National University Lviv Politechnika)	15 0.13 %
9	<a href="https://krs.chmnu.edu.ua/jspui/bitstream/123456789/3758/1/%D0%9A%D0%B2%D0%B0%D0%BB%D1%96%D1%84%D1%96%D0%BA%D0%B0%D1%86%D1%96%D0%B9%D0%BD%D0%B0%20%D1%80%D0%BE%D0%B1%D0%BE%D1%82%D0%B0%20%D0%84%D0%B2%D0%B4%D0%BE%D0%BA%D1%96%D0%BC%D0%BE%D0%B2.pdf">https://krs.chmnu.edu.ua/jspui/bitstream/123456789/3758/1/%D0%9A%D0%B2%D0%B0%D0%BB%D1%96%D1%84%D1%96%D0%BA%D0%B0%D1%86%D1%96%D0%B9%D0%BD%D0%B0%20%D1%80%D0%BE%D0%B1%D0%BE%D1%82%D0%B0%20%D0%84%D0%B2%D0%B4%D0%BE%D0%BA%D1%96%D0%BC%D0%BE%D0%B2.pdf</a>	15 0.13 %
10	<a href="https://openarchive.nure.ua/server/api/core/bitstreams/dd5ba248-fca4-45d9-a0ac-0c0964a6f0f4/content">https://openarchive.nure.ua/server/api/core/bitstreams/dd5ba248-fca4-45d9-a0ac-0c0964a6f0f4/content</a>	15 0.13 %
<b>з бази даних RefBooks (0.00 %)</b>		
ПОРЯДКОВИЙ НОМЕР	ЗАГОЛОВОК	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)
<b>з домашньої бази даних (0.11 %)</b>		
ПОРЯДКОВИЙ НОМЕР	ЗАГОЛОВОК	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)
1	2021_М_ПІ_Байдак_В_Є_ 5/31/2024 Kharkiv National University of Radio Electronics (Kharkiv National University of Radio Electronics)	13 (1) 0.11 %
<b>з програми обміну базами даних (1.96 %)</b>		
ПОРЯДКОВИЙ НОМЕР	ЗАГОЛОВОК	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)
1	МН_ПЗМД_ІСТ_КПІ_2025_Шерстюк І.О. 5/20/2025 <b>National Technical University of Ukraine Igor Sikorskyi Kyiv Politech Institute (ФІОТ, К-ра інформаційних систем та технологій)</b>	<b>128 (8) 1.08 %</b>
2	Дослідження методів створення персонального дашборду із впровадженням рекомендаційної системи.pdf 12/24/2023 <b>Odessa National Polytechnic University (ІПТДМ, Каф. ІТПД)</b>	<b>50 (4) 0.42 %</b>
3	122-Марченко_Микола_Миколайович-2021 6/5/2024 Donetsk National University named after V. Stus (Donetsk National University named after V. Stus)	23 (3) 0.19 %

4	2019_81260001_Antonyshchak_Serhii_Andriiovych_57720 10/25/2024 National University "Lviv Politechnika" (National University Lviv Politechnika)	15 (1) 0.13 %
5	Побудова інформаційної інфраструктури міста за допомогою рекомендаційних систем 3/16/2025 National Technical University of Ukraine Igor Sikorskyi Kyiv Politech Institute (National Technical University of Ukraine Igor Sikorskyi Kyiv Politech Institute)	11 (1) 0.09 %
6	Рекомендаційні системи на основі навчання з підкріпленням 3/16/2025 National Technical University of Ukraine Igor Sikorskyi Kyiv Politech Institute (National Technical University of Ukraine Igor Sikorskyi Kyiv Politech Institute)	5 (1) 0.04 %

### з Інтернету (1.33 %)

ПОРЯДКОВИЙ НОМЕР	ДЖЕРЕЛО URL	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)
1	<a href="https://openarchive.nure.ua/bitstreams/c4504c2d-5929-4d9f-91df-48e3b1a5aae6/download">https://openarchive.nure.ua/bitstreams/c4504c2d-5929-4d9f-91df-48e3b1a5aae6/download</a>	89 (7) 0.75 %
2	<a href="https://krs.chmnu.edu.ua/jspui/bitstream/123456789/3758/1/%D0%9A%D0%B2%D0%B0%D0%BB%D1%96%D1%84%D1%96%D0%BA%D0%B0%D1%86%D1%96%D0%B9%D0%BD%D0%B0%20%D1%80%D0%BE%D0%B1%D0%BE%D1%82%D0%B0%20%D0%84%D0%B2%D0%B4%D0%BE%D0%BA%D1%96%D0%BC%D0%BE%D0%B2.pdf">https://krs.chmnu.edu.ua/jspui/bitstream/123456789/3758/1/%D0%9A%D0%B2%D0%B0%D0%BB%D1%96%D1%84%D1%96%D0%BA%D0%B0%D1%86%D1%96%D0%B9%D0%BD%D0%B0%20%D1%80%D0%BE%D0%B1%D0%BE%D1%82%D0%B0%20%D0%84%D0%B2%D0%B4%D0%BE%D0%BA%D1%96%D0%BC%D0%BE%D0%B2.pdf</a>	53 (5) 0.45 %
3	<a href="https://openarchive.nure.ua/server/api/core/bitstreams/dd5ba248-fca4-45d9-a0ac-0c0964a6f0f4/content">https://openarchive.nure.ua/server/api/core/bitstreams/dd5ba248-fca4-45d9-a0ac-0c0964a6f0f4/content</a>	15 (1) 0.13 %

### Список прийнятих фрагментів (немає прийнятих фрагментів)

ПОРЯДКОВИЙ НОМЕР	ЗМІСТ	КІЛЬКІСТЬ ОДНАКОВИХ СЛІВ (ФРАГМЕНТІВ)
------------------	-------	---------------------------------------

## ДОДАТОК Б

### СЛАЙДИ ПРЕЗЕНТАЦІЇ

МІНІСТЕРСТВО  
ОСВІТИ І НАУКИ  
УКРАЇНИ

НАЦІОНАЛЬНИЙ  
УНІВЕРСИТЕТ  
РАДІОЕЛЕКТРОНИКИ  
NURE

Комплексний курсовий проєкт

# Дослідження рекомендацій для довгострокового задоволення користувачів на прикладі сучасних платформ персоналізації контенту

SE  
software  
engineering

Огу Стефні Іфома, ІПЗм-23-1  
Науковий керівник: доц. Вленда Наталя Анатоліївна

23 червня, 2025

## Дослідження

### Актуальність та стан розвитку галузі

Проблема інформаційного перевантаження користувачів

Глобальні дані, що генеруються щорічно

Згенеровані дані (екстабайти)

402.74 мільйони ТБ  
КОЖЕН ДЕНЬ

Рік	Згенеровані дані (екстабайти)
2010	~1
2011	~2
2012	~4
2013	~8
2014	~16
2015	~32
2016	~64
2017	~128
2018	~256
2019	~512
2020	~1024
2021	~2048
2022	~4096
2023	~8192
2024	~16384
2025	181 36

147 36

Рік

За джерелом [ExplodingTopics](#)

SE  
software  
engineering

2

## Дослідження

### Напрямок дослідження



Аналіз підходів розробки рекомендаційних систем (за вмістом, колаборативна фільтрація, гібридна) та ідентифікація актуальних проблем галузі



Створення системи рекомендацій, яка орієнтована на **довгострокове задоволення** користувача



## Дослідження

### Об'єкт дослідження

**Об'єктом дослідження** є сучасні **рекомендаційні системи** на прикладі платформ, які використовують алгоритми персоналізації контенту. Особливу увагу приділено підходам, які враховують **поведінку користувача в динаміці та довготривалі метрики**.



# Огляд літератури

## Перелік джерел у галузі

Перелік літератури			
Публікації конференцій та тез	<ol style="list-style-type: none"> <li>1) RecSys '09: Proceedings of the third ACM conference on Recommender systems</li> <li>2) RecSys '11: Proceedings of the 10th ACM conference on Recommender systems</li> <li>3) RecSys '18: Proceedings of the 12th ACM Conference on Recommender Systems</li> <li>4) RecSys '20: Proceedings of the 14th ACM Conference on Recommender Systems</li> <li>5) 1st ACM Conference on Electronic Commerce, Denver, Colorado, United States</li> <li>6) 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining</li> <li>7) 17th ACM Conference on Recommender Systems</li> <li>8) CHI '02: Human Factors in Computing Systems</li> <li>9) CHI '05: Extended Abstracts on Human Factors in Computing Systems</li> <li>10) UICERSTG Workshop at the 5th ACM Conference on Recommender Systems</li> <li>11) Information Science and Applications (ICISA) 2016</li> <li>12) IUI '02: Proceedings of the 7th international conference on Intelligent user interfaces</li> <li>13) Fifth international conference on computer and information science</li> <li>14) International World Wide Web Conference</li> <li>15) International Conference on Engineering Management, Information Technology and Intelligence</li> <li>16) International Conference on Emerging Trends in Information and Communication Security</li> <li>17) Proceedings of the 2008 ACM Conference on Recommender Systems</li> <li>18) WWW '10: Proceedings of the 19th international conference on World wide web</li> </ol>	Технічні блоги та есеї	<ol style="list-style-type: none"> <li>1) Netflix Technology Blog</li> <li>2) Pinterest Engineering Blog</li> <li>3) Non-Brand Data</li> <li>4) Medium</li> <li>5) Числа стратегій State під ризиком</li> <li>6) Байєсівська логіка політикою</li> <li>7) Оптимізація аудіорекомендацій на довідковому перспективі</li> <li>8) Підприємство перспектива наванчання</li> <li>9) Spotify Research</li> </ol>
Публікації в журналах	<ol style="list-style-type: none"> <li>1) Machine Learning with Applications</li> <li>2) Expert Systems with Applications</li> <li>3) IEEE Transactions on Knowledge and Data Engineering</li> <li>4) iBusiness</li> <li>5) Modeling and User-Adapted Interaction</li> <li>6) Science</li> <li>7) User Modeling and User-Adapted Interaction</li> <li>8) Knowledge and Information Systems</li> <li>9) Egyptian Informatics Journal</li> <li>10) AI magazine</li> <li>11) ACM Transactions on Information Systems</li> <li>12) User Modeling and User-Adapted Interaction</li> <li>13) International Journal of Information Technology</li> <li>14) Artificial Intelligence Review</li> <li>15) ACM Transactions on Internet Technology (TOIT)</li> <li>17) Expert Systems with Applications</li> <li>18) Journal of Management Information Systems</li> <li>19) Journal of Big Data</li> <li>20) Computational Biology and Chemistry</li> <li>21) Adv. Artificial Intelligence</li> <li>22) Physica A: Statistical Mechanics and its Applications</li> </ol>	Заяти та дослідження	<ol style="list-style-type: none"> <li>1) Електронний архів Харківського національного університету радіоелектроніки «ENR KNURE»</li> <li>2) Physics Reports</li> </ol>



Академічні джерела, журнали, конференції джерел ACM, IEEE, Google Scholar про опис рекомендаційних систем, проблеми, сучасні дослідження Tiktok, Pinterest

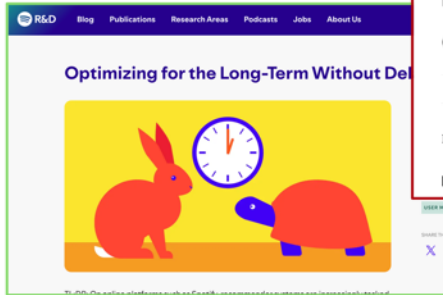
# Огляд літератури

## Огляд предметної області

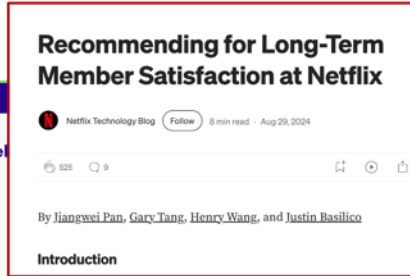


# Огляд літератури

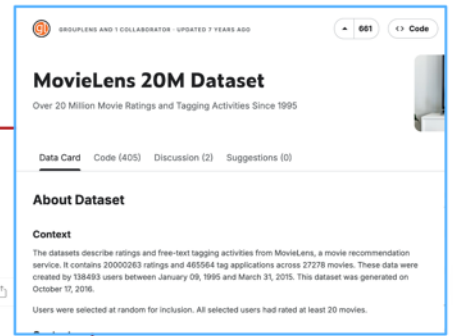
Перелік основних джерел та теорій у галузі



**Spotify Research: нетерплячі бандити та навчання з підкріпленням**



**Netflix Technology Blog: контекстуальні бандити та проксі-нагороди**

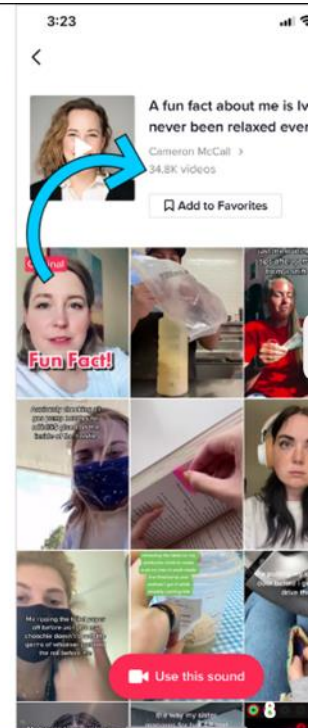
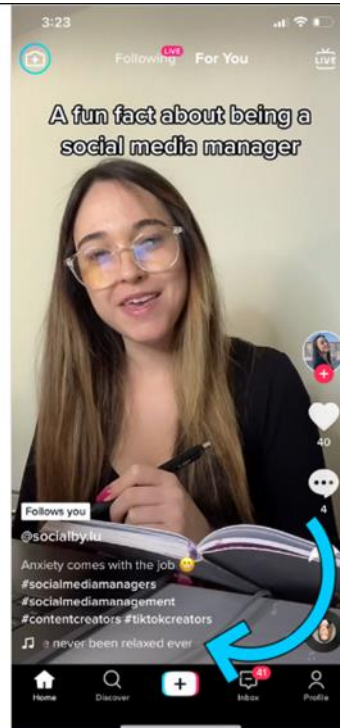


**Дослідження MovieLens та інших датасетів**

# Огляд літератури

Зазначення прогалин у наявних дослідженнях

Рекомендаційні системах часто використовуються короткострокові метрики, такі як кліки чи перегляди, які легко виміряти



## Огляд літератури

Зазначення прогалин у наявних дослідженнях



Однак, для багатьох сервісів головна мета – **довгострокове задоволення користувача**. Якщо модель навчання надто оптимізується під ці короткострокові метрики, вона може втратити зв'язок із довгостроковими цілями



## Огляд літератури

Зазначення прогалин у наявних дослідженнях

Подивилася за одну ніч та поставила відгук 5 зірок



Ганна




Почала дивитись, але працює над дипломною роботою, тому нема часу закінчити

Додала в подивитись пізніше



## Постановка задачі

### Формулювання проблеми

 Як створити рекомендаційну систему, що враховує довгострокові потреби користувачів?

 Як використати проксі-нагороди для оптимізації задоволення?


### Опис очікуваних результатів


 **Гібридна** рекомендаційна система з **проксі-нагородами**


 **Демонстрація ефективності** в порівнянні з традиційними підходами


## Методологія

### Опис використаних методів дослідження

 Аналіз існуючих алгоритмів рекомендацій. Теоретичне дослідження забезпечення довгострокового задоволення користувача

 Розробка гібридних підходів та їх тестування на основі реальних даних MovieLens (жанрова на основі скалярного добутку, колаборативна фільтрація (SVD))

 Використання проксі-нагород та алгоритму LinUCB для контекстуальних бандитів та методи оптимізації довгострокової задоволеності користувача

 Експериментальна оцінка через порівняння рекомендацій

## Методологія

Інструментарій та технології, використані в роботі



Набір даних: MovieLens, що надає достатньо інформації для моделювання рекомендаційної системи



Python як основна мова програмування, бібліотеки для машинного навчання (Surprise), бібліотеки для аналізу даних (Pandas, NumPy, Pickle), інструменти для візуалізації результатів (Matplotlib)



Фреймворк для створення REST API. Фронтенд HTML/CSS/JavaScript



Реляційна база даних для збереження інформації про користувачів

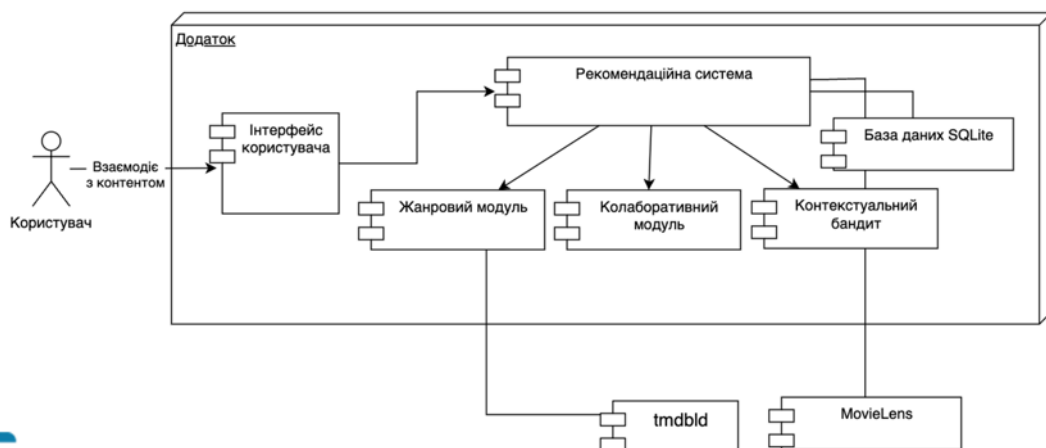


+ невелика інтеграція з **TMDB**

13

## Архітектура система для проведення експериментального дослідження

Архітектура розробленої системи



14

## Архітектура система для проведення експериментального дослідження

### Ключові компоненти – Жанрова фільтрація

1. Для кожного твору формується вектор жанрів, де значення 1 означає належність до жанру, а 0 – відсутність. Усі вектори творів розташовуються в n-вимірному просторі, де n – це кількість всіх жанрів.
2. Формується вектор користувача
3. Обчислюється подібність

$$DotSim(\vec{D}, \vec{Q}_i) = \sum_{j=1}^n d_j q_j$$

де  $d$  та  $q$  – два вектори.



## Архітектура система для проведення експериментального дослідження

### Ключові компоненти – Колаборативна фільтрація

1. Створюємо матрицю користувачів x фільмів із рейтингами;
2. Використовуємо SVD для представлення кожного користувача та фільм у вигляді векторів у спільному латентному просторі;
3. Визначаємо схожість користувачів. Для кожного користувача обчислюються ймовірні оцінки тих фільмів, які він ще не переглядав. Прогноз ґрунтується на схожості між векторами користувача та фільму;
4. Виконуємо сортування, фільтрацію, видаляємо з рекомендацій фільми, які користувач вже дивився, пріоритизуємо нові релізи, щоб користувачі мали актуальні пропозиції та обираємо N кількість для показу.



## Архітектура система для проведення експериментального дослідження

Ключові компоненти – Гібридна фільтрація



**Жанрова:**

100% співпадіння



**Колаборативна:**

RMSE = 0.9674 для колаборативної моделі

70% колаборативна + 30% жанрова



17

## Архітектура система для проведення експериментального дослідження

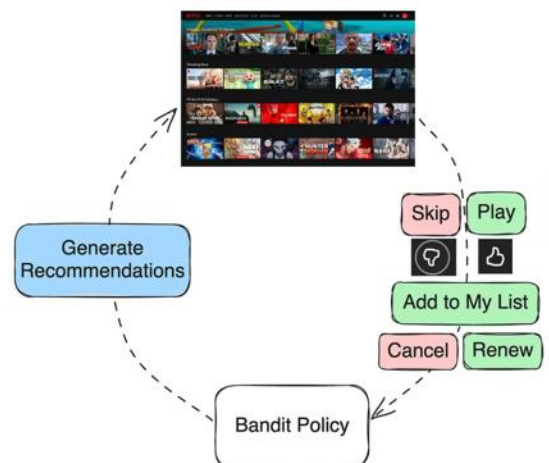
Ключові компоненти – Проксі-нагороди

**Явні:**

- 1) додати у подивитись пізніше: чи користувач додав фільм у збережений список -  $0.3$ ;
- 2) рейтинг, як безпосередній показник задоволення від перегляду фільму -  $-1.0 - 1.0$ ;

**Неявні:**

- 1) чи почав користувач дивитися фільм -  $0.2$ ;
- 2) чи завершив дивитися -  $0.7$ .



18

## Архітектура система для проведення експериментального дослідження

Ключові компоненти – Контекстуальні бандити



$$\text{Сумарна винагорода} = \sum_{t=1}^T r_t$$

де  $r_t = R_t \alpha_t$  – винагорода, отримана в результаті виконання дії  $\alpha$ .

## Архітектура система для проведення експериментального дослідження

Перенавчання

### Колаборативна

Кожен раз, коли користувач нові  
N ставить відгуків

### Проксі-нагород

Після кожної дії користувача

## Зміст проведеного експерименту

### Вхідні дані:

- 100,000 рейтингів з MovieLens 20M
- Профілі користувачів з улюбленими жанрами
- Проксі-нагороди за дії користувачів

### Критерії:

- Середнє задоволення користувача
- Адаптація до змін вподобань
- Позитивний досвід взаємодії



21

## Зміст проведеного експерименту

### Методи

#### 1. Симуляція користувача з динамічними вподобаннями



#### 2. Порівняння контекстуального бандиту vs колаборативної фільтрації

Гаусівський шум ( $\sigma = 0.05$ ) для реалістичності

100 незалежних експериментів для статистичної значущості



22

## Зміст проведеного експерименту

### Послідовність

1. Ініціалізація профілю користувача з улюбленими жанрами
2. Генерація сесій з поступовою зміною вподобань
3. Отримання рекомендацій від обох систем
4. Симуляція взаємодії користувача з фільмами
5. Оновлення моделей на основі проксі-нагород
6. Повторення циклу для 100 експериментів



23

## Результати експерименту

### Порівняння базових моделей з моделлю з проксі-нагородами



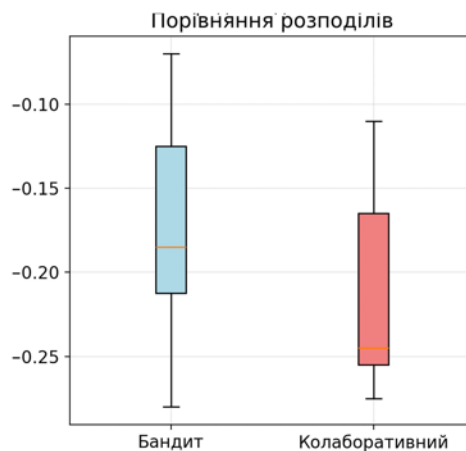
Порівняння середнього задоволення користувачів



24

## Результати експерименту

Порівняння базових моделей з моделлю з проксі-нагородами



Статистичний розподіл середнього задоволення користувачів

25

## Аналіз отриманих результатів

- 👍 Контекстуальний бандит виграв у 70% за середнім задоволенням
  - 👍 60% перемог за адаптацією та позитивним досвідом
  - 👍 Загалом 63.3% перемог по всіх метриках (190 з 300)
- ↓
- 👍 Краща стабільність результатів бандиту
  - 👍 Ефективна адаптація до змін вподобань

26

## Аналіз отриманих результатів

- ✓ Створено гібридну систему з проксі-нагородами
- ✓ Доведено ефективність контекстуальних бандитів
- ✓ Продемонстровано кращу адаптацію до користувачів

Контекстуальні бандити систематично перевершують традиційну колаборативну фільтрацію. Проксі-нагороди ефективно відображають довгострокове задоволення. Гібридний підхід забезпечує баланс між точністю та різноманітністю.



### Публікація результатів

**SCI-CONF.COM.UA**  
**SCIENCE AND TECHNOLOGY:  
 CHALLENGES, PROSPECTS  
 AND INNOVATIONS**

PROCEEDINGS OF IX  
 SCIENTIFIC AND PRACTICAL CONFERENCE  
 APRIL 24-26  
**OSAKA  
 2025**

**CERTIFICATE**  
 is awarded to  
**Ogu Stephanie**  
 for being an active participant in  
 XI International Scientific and Practical Conference  
**"SCIENCE AND TECHNOLOGY: CHALLENGES,  
 PROSPECTS AND INNOVATIONS"**  
*24 Hours of Participation  
 (0.8 ECTS credits)*

**OSAKA**  
 19-21 June 2025

sci-conf.com.ua

UDC 004.832.23  
**RESEARCHING RECOMMENDATIONS FOR LONG-TERM USER  
 SATISFACTION USING A MODERN CONTENT PERSONALISATION  
 PLATFORM AS AN EXAMPLE**

**Ogu Stephanie Hoana,**  
 Student at the Department of Software Engineering  
**Valenda Natalia Anatoliivna**  
 Associate Professor of the Department of Software Engineering,  
 PhD in Engineering Sciences,  
 Associate Professor  
 Kharkiv National University of Radio Electronics  
 Kharkiv, Ukraine

**Abstract.** The object of the research is recommender systems that take into account long-term user satisfaction in modern content personalisation platforms. The research and design methods are the use of collaborative filtering, matrix factorisation, the use of contextual bandits and machine learning. An approach for using proxy rewards has also been developed, which allows taking into account both short-term and long-term user satisfaction.

**Keywords:** Collaborative filtering, content-based filtering, contextual bandit, matrix factorisation, Netflix, proxy-rewards, recommendation systems.

Recommender systems play a crucial role in modern digital platforms, providing personalised user experiences. Their key feature is the ability to process large amounts of data and generate recommendations that match user preferences. However, with the development of information systems, user demand and industry requirements pose new challenges to recommender systems.

Many popular services, such as social networks TikTok, Instagram, etc., work with short-term user interactions and prioritise their short-term retention. At the same time, the problem of some large platforms is that long-term user retention is more difficult to achieve. This reflects one of the specifics of recommendation platforms.

---

## Підсумки

### Реалістичність та корисність отриманих результатів

1. Система протестована на реальних даних MovieLens
2. Практичне застосування для Netflix, Spotify подібних платформ
3. Покращення утримання користувачів через довгострокове задоволення



29

---

## Підсумки

### Обмеження

1. Масштабованість (тільки 100K записів)
2. Відсутність реального тестування з користувачами
3. Емпірично встановлені параметри потребують A/B тестування



30

## Підсумки

### Ключові напрямки алгоритмічних досліджень

- Ефективна інтеграція різних моделей рекомендацій в одну систему
- Дослідження значення проксі-нагород на якість рекомендацій
- Використання ШІ для передбачення на основі проксі-нагород
- Оптимізація та тестування різних видів моделей бандитів
- Поглиблена інтеграція з TMDb та IMDb
- Пошук релевантних залежностей між фільмами на основі метрик
- Додавання збагачених даних через зовнішні API
- Врахування демографічних характеристик користувачів
- Тестування системи на одночасному використанні різної кількості користувачів

## Підсумки

### Глобальні напрямки алгоритмічних досліджень

- Дослідження проблеми відкладеного зворотного зв'язку (за досвідом Spotify)
- Поєднання вивчення користувачів з боку бізнесу та соціальних наук
- Адаптація здобутків у алгоритми покращення взаємодії з користувачем

ДОДАТОК В  
АПРОБАЦІЯ РЕЗУЛЬТАТІВ РОБОТИ



Рисунок В.1 – Сертифікат участі у ХІ міжнародній науково-практичній конференції «Science And Technology: Challenges, Prospects And Innovations», 19-21.06.2025, Осака, Японія, ISBN 978-4-9783419-4-5

**UDC 004.832.23**

**RESEARCHING RECOMMENDATIONS FOR LONG-TERM USER  
SATISFACTION USING A MODERN CONTENT PERSONALISATION  
PLATFORM AS AN EXAMPLE**

**Ogu Stephanie Ifeoma,**

Student at the Department of Software Engineering

**Valenda Natalia Anatoliivna**

Associate Professor of the Department of Software Engineering,

PhD in Engineering Sciences,

Associate Professor

Kharkiv National University of Radio Electronics

**Abstract.** The object of the research is recommender systems that take into account long-term user satisfaction in modern content personalisation platforms. The research and design methods are the use of collaborative filtering, matrix factorisation, the use of contextual bandits and machine learning. An approach for using proxy rewards has also been developed, which allows taking into account both short-term and long-term user satisfaction.

**Keywords:** Collaborative filtering, content-based filtering, contextual bandit, matrix factorisation, Netflix, proxi-rewards, recommendation systems.

Recommender systems play a crucial role in modern digital platforms, providing personalised user experiences. Their key feature is the ability to process large amounts of data and generate recommendations that match user preferences. However, with the development of information systems, user demand and industry requirements pose new challenges to recommender systems.

Many popular services, such as social networks Tiktok, Instagram, etc., work with short-term user interactions and prioritise their short-term retention. At the same time, the problem of some large platforms is that long-term user retention is more difficult to achieve. This reflects one of the specifics of recommendation platforms, which is the need to take into account individual user preferences not only in real time, but also in the long term.

This study examines two globally popular systems: the streaming service Netflix and the music service Spotify, which are examples of successful and well-known recommendation systems.

Netflix accounts for 80% of streaming time through its recommendation system. They also claim to improve the user experience to retain customers [1].

And Spotify and its Made for You recommendations, according to the recently released Made to be Found report, have led to the discovery of over a third of all new artists. Both platforms aim to entertain the world, consistently deliver a positive user

experience, and, most importantly for our purposes in this study, create long-term enjoyment [2] [3].

As Spotify's own scientists, developers, and researchers point out, traditionally most recommendation systems have been trained to optimise for short-term feedback (clicks, likes, session duration, etc.) [3]. And Netflix research has shown that optimising recommendation systems for long-term outcomes is a more challenging task than optimising for short-term outcomes [4].

Short-term optimisation works with feedback that can be obtained almost immediately when a user directly interacts with a recommended item.

Data for long-term optimisation is more difficult to track because it:

- is typically less dense: there is much less data on long-term user actions (e.g., whether a user returned a month after interacting with a recommended item) than there is on short-term actions;

- has a higher level of noise: may contain more random factors that are not directly related to the recommendation itself. For example, user satisfaction may depend not only on the quality of the item, but also on their mood, circumstances, or external factors that are difficult or even impossible to take into account;

- has weaker connections to specific recommendations: long-term results (e.g., satisfaction with a watched movie weeks or months later) are more difficult to directly associate with specific recommendations that were provided in the past.

However, the customer must be satisfied, because participant satisfaction is closely related to the likelihood that the user will stay on the platform, renew their monthly subscription, which directly affects the company's revenue. Therefore, optimising the recommendation system to maximise customer satisfaction and increase customer retention is a critical issue [5].

Based on the analysis of Netflix and Spotify solutions, it was determined that contextual bandit methods can be applied, where each interaction with the user is treated as a separate context, and recommendations are adapted according to the feedback received. This communication can be immediate (skipping, playing, rating "dislike" or adding items to a playlist) or delayed (ending a show or renewing a subscription).

During this research, the obtained knowledge was practically applied to create a recommendation system that uses proxy rewards to evaluate short-term interactions and simulate long-term satisfaction, testing how the chosen approach affects the quality of recommendations in the long term.

The project will be implemented in several stages: development of initial recommendations (implementation of a content-based recommendation algorithm, when the user indicates only initial preferences in film genres), development of a recommendation system (implementation of a recommendation algorithm based on collaborative filtering to create an initial model), integration of modern methods for long-term recommendations (integration of proxy rewards to optimise long-term user satisfaction using contextual bandits). For the data the MovieLens dataset was used [6].

The first part of the practical implementation of the study consists of implementing initial recommendations based on the dataset. The problem that is present at this stage is the presence of minimal information about the user, i.e. the system faces the problem of a cold start. To implement recommendations in this context, it was chosen to use content-based filtering. Elements (films) are converted into vectors using metadata descriptions or internal characteristics as features [7]. In the case of this study, each genre is considered as a separate binary feature. For each movie, a genre vector is formed, where the value 1 means belonging to the genre, and 0 means absence. All vectors of works are located in an  $n$ -dimensional space, where  $n$  is the number of all genres. Proximity is determined using the scalar product [7].

Further, the user profile is formed as a set of vectors of those elements that he was interested in. Based on these vectors, an aggregated preference vector (e.g., average or weighted sum) is constructed. The algorithm compares the user profile vector with vectors of new (unwatched) items. Those movies whose vectors are closest to the user profile vector according to the selected metric are recommended to the user as the most relevant. The algorithm also chose to use the dot product.

The main recommendation model implements collaborative filtering based on the hypothesis that users with similar past preferences will have similar future preferences. This approach requires forming an interaction matrix  $R$  of dimensions  $m \times n$  (users  $\times$

movies) using rating data, where the system predicts unknown ratings for user-movie pairs not yet connected in the database.

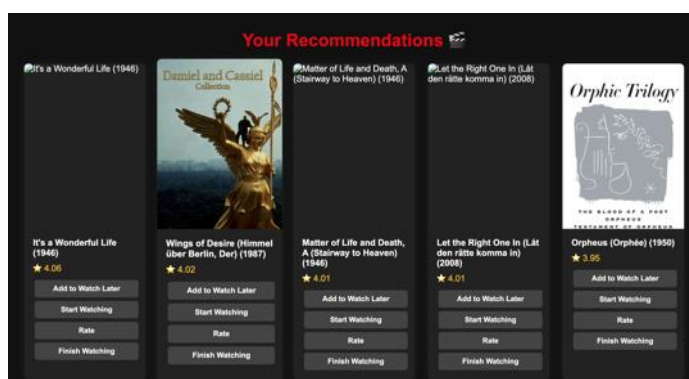
A critical characteristic of this data is its extreme sparsity – in real systems like MovieLens 20M, less than 1% of all possible user-movie combinations are filled, with density  $\rho$  typically below 0.01. This sparsity creates significant challenges, including insufficient data for similarity calculations, cold start problems for new users or items, and computational inefficiency when working directly with sparse matrices. To address the sparsity problem, the system employs matrix factorisation through Singular Value Decomposition (SVD), which approximates the sparse rating matrix  $R$  as a product of three smaller matrices.

This approach represents each user and movie as vectors in a latent space where hidden factors capture underlying preferences and characteristics that influence ratings. Rather than working with the large sparse matrix directly, the algorithm learns compact "profiles" for users and movies based on existing ratings, automatically discovering latent factors that might represent genre preferences or other implicit characteristics. The analysis process involves creating the user-movie rating matrix, applying SVD to represent users and movies as vectors in shared latent space, calculating user similarity using cosine similarity between latent factor vectors, and predicting probable ratings for unseen movies. Final recommendations are generated by sorting predictions, filtering out already-watched movies, prioritising recent releases for relevance, and selecting the top  $N$  items to display to users.

Analysis of Netflix and Spotify experiences demonstrates the power of proxy rewards and contextual bandits for addressing long-term user satisfaction. However, the current system adapts only to explicit ratings, creating a significant limitation where users who don't finish watching movies or don't provide feedback cannot receive meaningful recommendations. The solution involves implementing proxy rewards by logging user actions through interface buttons representing both explicit feedback (adding movies to "watch later" lists and direct ratings) and implicit signals (whether users started watching and whether they completed viewing). These proxy rewards will serve as the foundation for a contextual bandit model aligned with reinforcement learning principles. The

contextual bandit approach addresses scalability challenges by using a single universal model operating on contextual vectors rather than maintaining separate "arms" for each movie. Each film is represented through a contextual vector incorporating genre classification, rating characteristics, temporal features, personalized user preference metrics, and popularity/quality indicators. The algorithm learns to predict rewards based on this context rather than movie identifiers, making the system computationally feasible while maintaining effectiveness.

The system was implemented as a web app, which tracks user interaction through buttons that simulate core user behaviours, as shown in Figure 1.



**Figure 1. Recommender system interface**

Each action is assigned specific proxy reward values based on implicit feedback research in recommendation systems. For example, completing a movie viewing receives higher reward values as it strongly indicates user satisfaction, while low ratings (one or two stars) receive negative values to reflect user dissatisfaction. The contextual bandit system is initialised with two key data structures: a  $20 \times 20$  covariance matrix  $A$  that tracks relationships between movie characteristics (initially an identity matrix indicating no prior knowledge), and a  $20 \times 1$  vector  $b$  that accumulates weighted rewards for each contextual feature (initially zeros). The system uses the LinUCB (Linear Upper Confidence Bound) algorithm, where each movie's value combines expected reward (the system's prediction of user preference) with confidence interval (uncertainty measure that gives unknown films a chance for recommendation).

This work attempted to address the long-term user satisfaction problem in recommendation systems, though the solution has inherent limitations. The system's scalability is constrained by resource limitations, utilising only 100,000 records from

MovieLens 20M compared to the billions of interactions processed by platforms like Netflix. Additionally, the system lacks comprehensive testing metrics for full system evaluation. While the LinUCB algorithm's conservative bias toward older films with more historical data was identified and addressed through recency bonuses and increased exploration parameters, other hidden issues may exist due to the limited research timeframe. The research successfully developed a hybrid recommendation system that integrates three complementary approaches: genre-based content filtering for initial recommendations, collaborative filtering using SVD matrix factorisation (achieving 0.9674 RMSE), and contextual bandits with proxy rewards for long-term satisfaction optimisation. The system demonstrates effective handling of data sparsity through matrix factorisation techniques and implements user interaction tracking through explicit feedback (ratings, watch-later lists) and implicit signals (viewing start/completion). The contextual bandit approach using LinUCB algorithm balances exploration of new content with exploitation of learned preferences, adapting to user behaviour through continuous learning. Future development opportunities include optimising proxy reward values for different user actions, integrating advanced AI/ML techniques for prediction enhancement, testing various bandit model configurations, conducting scalability testing with multiple concurrent users, and expanding external data integration with TMDb and IMDb APIs. The foundation established enables further research into demographic-based personalisation and delayed feedback mechanisms, positioning the system for evolution into a more comprehensive recommendation platform addressing both immediate user preferences and long-term engagement satisfaction.

## REFERENCES;

1. David Chong. Deep Dive into Netflix's Recommender System. Medium. URL: <https://towardsdatascience.com/deep-dive-into-netflixs-recommender-system-341806ae3b48>.
2. Deep learning for recommender systems: A Netflix case study / Harald Steck та ін. {AI} magazine. 2021. Т. 42, № 3. С. 7–18. URL: <https://onlinelibrary.wiley.com/doi/10.1609/aimag.v42i3.18140> (visited: 24.12.2024).

3. Deep learning for recommender systems: A Netflix case study / Harald Steck та ін. {AI} magazine. 2021. Т. 42, № 3. С. 7–18. URL: <https://onlinelibrary.wiley.com/doi/10.1609/aimag.v42i3.18140> (visited: 24.12.2024).
4. Optimizing for the Long-Term Without Delay / Thomas M. McDonald та ін. Spotify Research. URL: <https://research.atspotify.com/2023/07/optimizing-for-the-long-term-without-delay/> (visited: 27.12.2024).
5. Reward innovation for long-term member satisfaction / Gary Tang та ін. 17th {ACM} Conference on Recommender Systems. 2023. С. 396 – 399. URL: <https://doi.org/10.1145/3604915.3608873> (visited: 25.12.2024).
6. Find By ID. The Movie Database (TMDB). URL: <https://developer.themoviedb.org/reference/find-by-id> (дата звернення: 13.01.2025).
7. D J. M. P., Kavlakoglu E. What is content-based filtering? | IBM. IBM - United States. URL: <https://www.ibm.com/think/topics/content-based-filtering> (visited: 13.01.2025).

## ДОДАТОК Г

Експертний висновок результатів перевірки кваліфікаційної роботи на  
відповідність оформлення вимогам ДСТУ 3008: 2015

Експертний висновок результатів перевірки кваліфікаційної роботи

студент  
(посада)

програмної інженерії  
(кафедра)

ІПМ-23-1  
(група)

Огу Стефні Іфома

(прізвище, ім'я, по батькові)

Зауваження

Пункт ДСТУ 3008-2015	Зміст пункту	Сторінка кваліфікаційної роботи
1	2	3
	<b>7.1 Загальні положення</b>	
	<b>7.3 Нумерація сторінок звіту</b>	
	<b>7.4 Нумерація розділів, підрозділів, пунктів, підпунктів</b>	
	<b>7.5 Рисунки</b>	
	<b>7.6 Таблиці</b>	
	<b>7.7 Переліки</b>	
	<b>7.8 Примітки</b>	
	<b>7.9 Виноски</b>	
	<b>7.10 Формули та рівняння</b>	
	<b>7.11 Посилання</b>	
	<b>7.13 Список авторів</b>	
	<b>7.14 Скорочення та умовні позначки</b>	
	<b>7.15 Додатки</b>	

Експерт

зауважень немає

\_\_\_\_\_  
(підпис)

Олена ОЛІЙНИК

(прізвище, ініціали)

18.06.2025