

Розробка WEB- орієнтованої рекомендаційної системи

ВИКОНАВ РАКІТЕНКО ДМИТРІЙ АНДРІЙОВИЧ, СТУДЕНТ БП КН-4

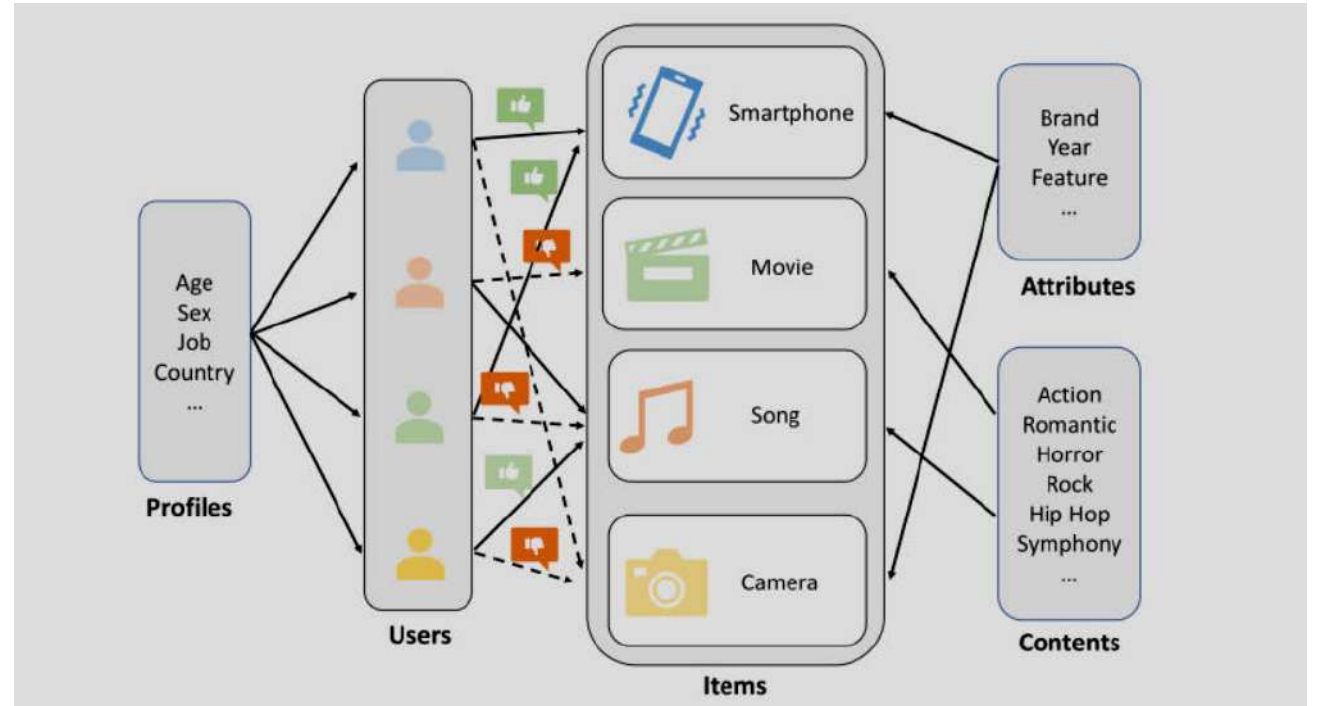
НАУКОВИЙ КЕРІВНИК ГОРБОРУКОВ В.В., КАНДИДАТ ТЕХНІЧНИХ НАУК,
ДОЦЕНТ

Постановка задачі

- 1) Здійснити аналіз наявних прикладів рекомендаційних систем, особливостей розробки рекомендаційної системи.
- 2) Дослідити популярні типи та методи реалізації рекомендаційних систем, описати популярні алгоритми та методи оцінки ефективності роботи рекомендаційної системи.
- 3) Розробити веб-сайт для перегляду інформації про анімаційні серіали та надання їм оцінок.
- 4) Розробити для створеного веб-сайту рекомендаційну систему з використанням фільтрації за вмістом та колаборативної фільтрації.
- 5) У розробленій системі імплементувати гібридний метод надання рекомендацій.

Актуальність проблеми

Наразі продовжується активне вдосконалення наявних методів рекомендаційних систем, а потреба в таких системах надзвичайно висока. Вони використовуються у таких популярних веб-платформах, як “Netflix”, “Youtube”, “Rozetka” тощо. Перед сучасними рекомендаційними системами постає задача генерації якісних рекомендацій для різних груп користувачів та вирішення пов’язаних з цим проблем.



Зміст роботи

- 1) Розділ 1 - дослідження теорії рекомендаційних систем та використовуваних алгоритмів
- 2) Розділ 2 - розробка веб-сайту для перегляду інформації про серіали
- 3) Розділ 3 - розробка рекомендаційної системи
- 4) Розділ 3 - опис імплементації у рекомендаційній системі вдосконалених методів рекомендацій

Метою даної роботи є реалізація веб-сайту для перегляду даних про анімаційні серіали та наданням їм оцінок і гібридної персоналізованої рекомендаційної системи для нього.

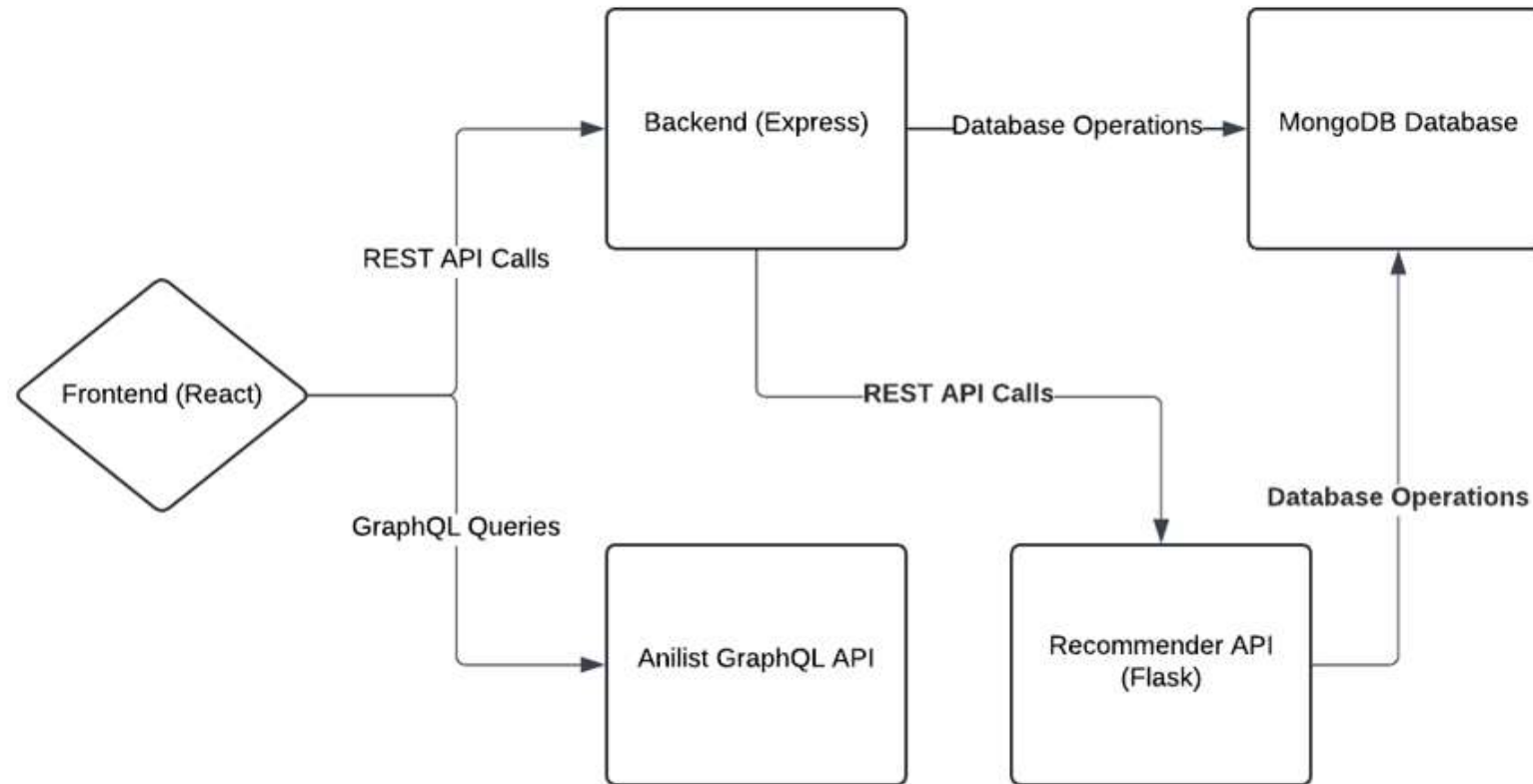
Об'єктом дослідження є методи сучасних веб-орієнтованих рекомендаційних систем, зокрема для генерації рекомендацій для веб-сайтів перегляду фільмів та читання книжок.

Методами дослідження є вивчення та аналіз наявних здобутків в галузі рекомендаційних систем на основі джерел, розробка моделі власного методу рекомендацій, порівняння методів рекомендаційних систем.

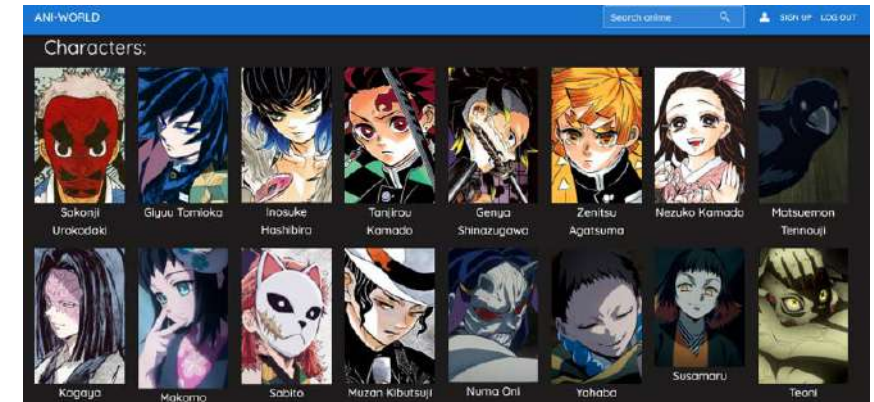
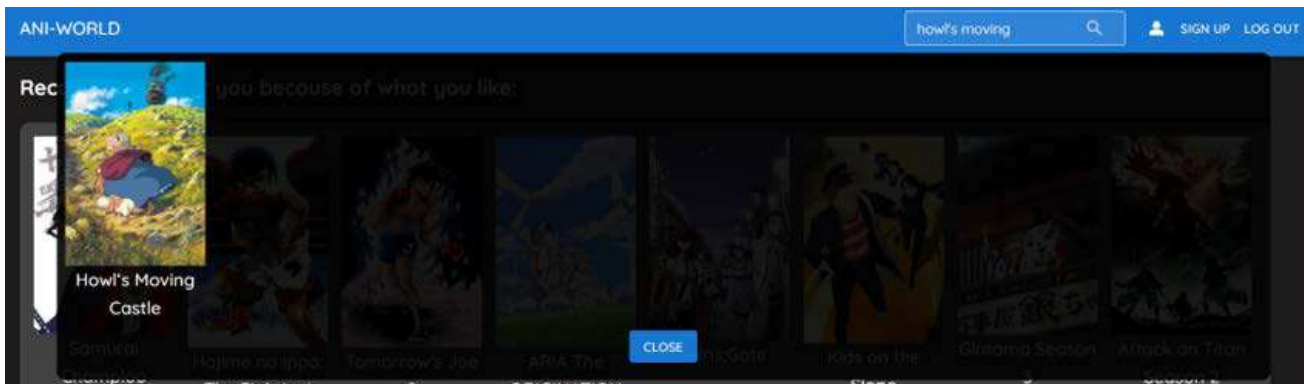
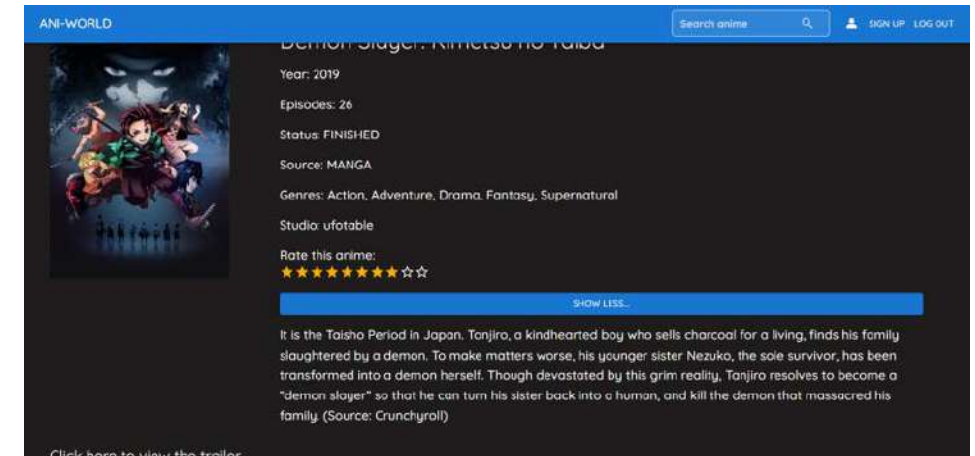
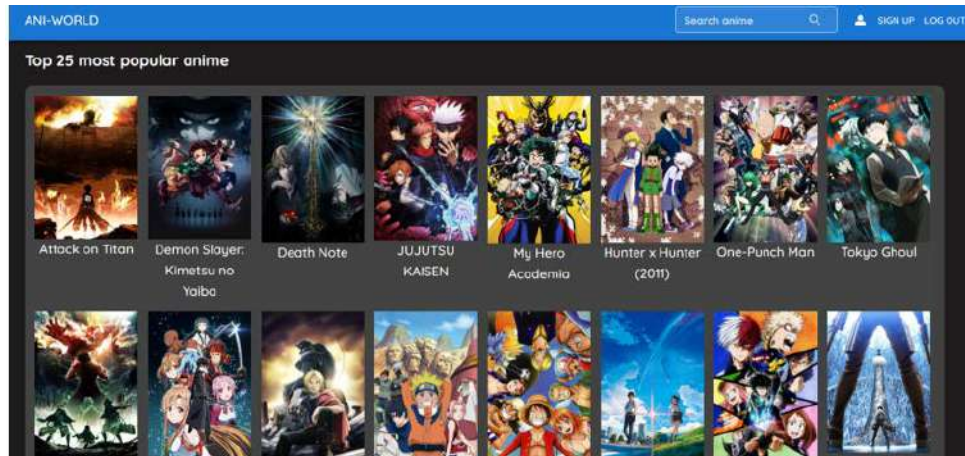
Теоретичний огляд

- 1) Приклади рекомендаційних систем
- 2) Збір даних для рекомендаційної системи
- 3) Історична довідка
- 4) Класифікація рекомендаційних систем та проблема “холодного старту”
- 5) Базові методи та алгоритми: оцінки подібності, кластеризація, SVD, TF-IDF, LDA, оцінки ефективності роботи системи
- 6) Характеристика колаборативного та контентного підходів надання рекомендацій

Архітектура веб-застосунку



Інтерфейс розробленого веб-сайту



Профіль користувача

Register on Ani-World

Enter email...

Enter nickname...

MM.YY.PPPP

Enter password...

Confirm password...

Gender *

SIGN UP

ALREADY HAVE AN ACCOUNT? CLICK HERE TO LOG IN

Welcome to Ani-World!

Enter email...

Enter password...

LOG IN

DON'T HAVE AN ACCOUNT? CLICK HERE TO REGISTER

ANI-WORLD

Search anime

SIGN UP LOG OUT

 **Dima**

Email: test@mail.com
Date of birth: 05.10.2002
Gender: male

Description: Robots in space and cowboys in spaceships
Genres you like: Mystery, Psychological

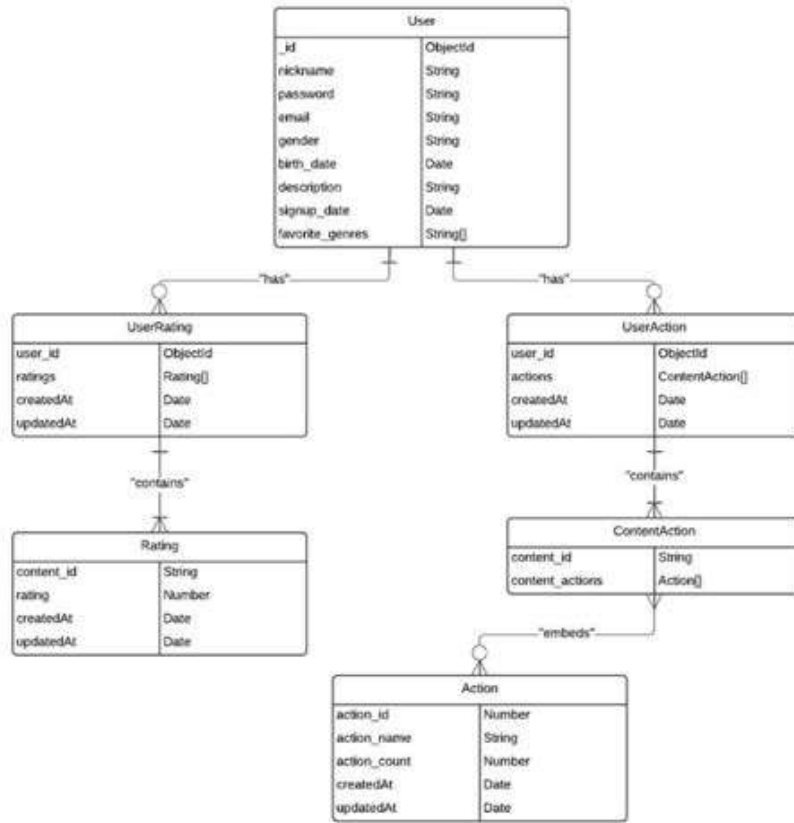
UPDATE PREFERENCES

What you might enjoy based on genres you like:

Monster Steins;Gate Monogatari Owarimonogatari Mob Psycho 100 Attack on Titan

База даних користувачів

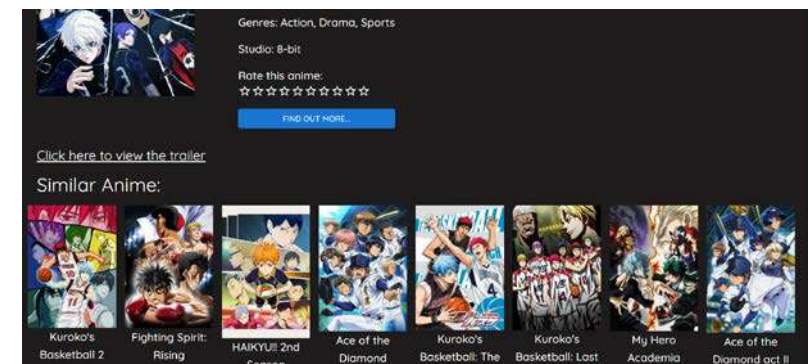
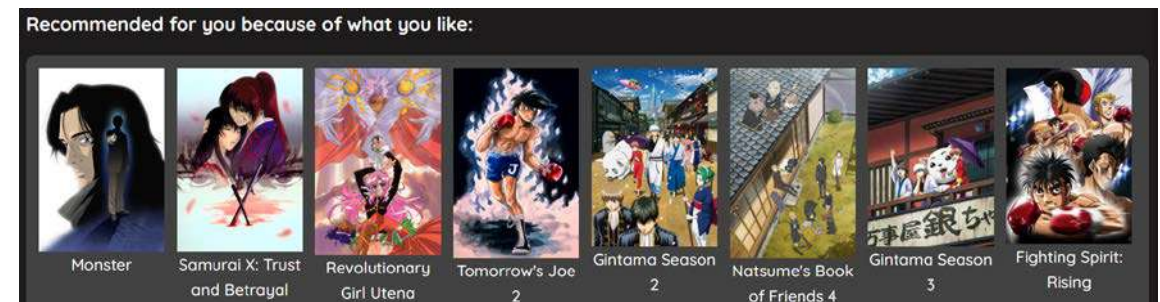


```
_id: ObjectId('6633b3ca868c3e212e636b03')
nickname: "Dima"
password: "$2a$10$keeUYjvp20UXIL3yD6uKS0yhBzU/cs1ApqeuRQNrYqLqCAXd.uNQW"
email: "test@mail.com"
gender: "male"
birth_date: 2002-10-05T00:00:00.000+00:00
signup_date: 2024-05-02T15:39:54.623+00:00
favorite_genres: Array (2)
__v: 0
description: "Robots in space and cowboys in spaceships"
```

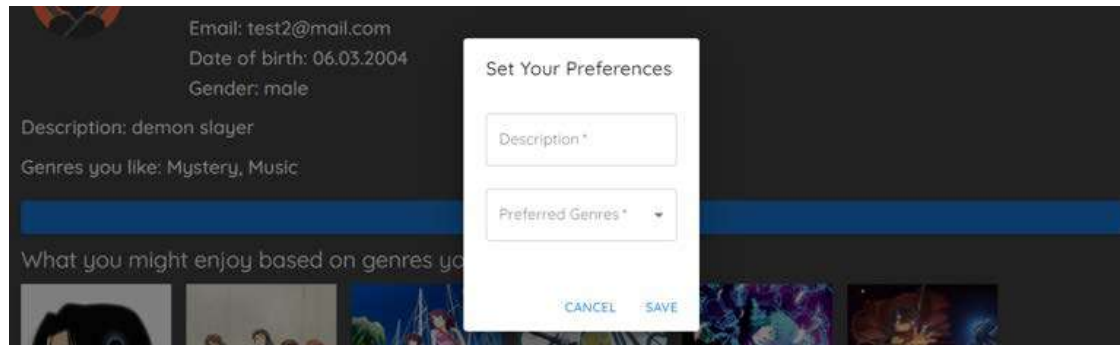
```
_id: ObjectId('6634ecb08bb1a91105cb6851')
user_id: ObjectId('6633b3ca868c3e212e636b03')
ratings: Array (11)
  0: Object
    content_id: "113415"
    rating: 7
    _id: ObjectId('6634ecb08bb1a91105cb6852')
    createdAt: 2024-05-03T14:28:35.585+00:00
    updatedAt: 2024-05-03T14:28:35.585+00:00
  1: Object
  2: Object
  3: Object
```

Рекомендаційна система

- 1) Персоналізовані рекомендації на головній сторінці з використанням гібридного методу (поєднання колаборативного та контентного методів)
- 2) Подібні серіали (TF-IDF, натренована LDA модель)
- 3) Подолання проблеми “холодного старту” рекомендаціями у профілі користувача на основі опису та улюблених жанрів (TF-IDF)



Подолання проблеми “холодного старту”



за жанрами “Кохання”, “Спорт”

$$TF = \frac{\text{кількість входжень слова в документ}}{\text{кількість слів в документі}}$$

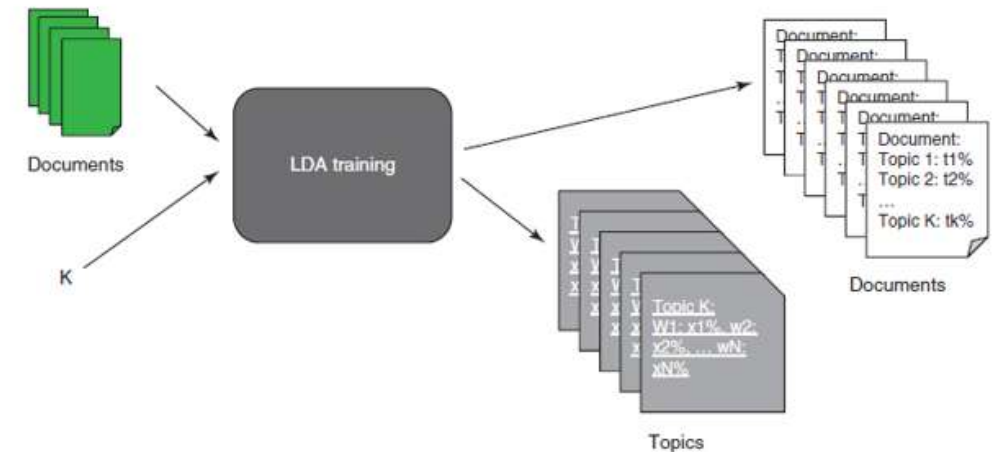
$$IDF = \log\left(\frac{\text{загальна кількість документів}}{\text{кількість документів, що містять слово}}\right)$$



за описом “I enjoy space battles and huge spaceships!”

Подібні серіали. Тренування LDA моделі

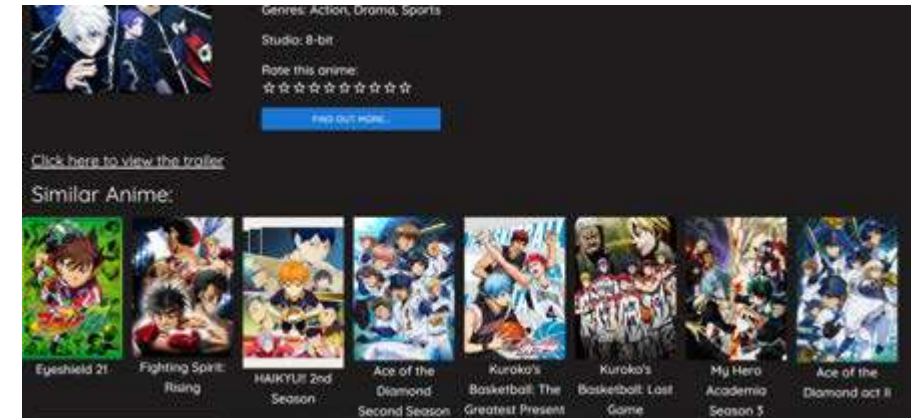
- 1) Токенізація та лематизація.
- 2) Збереження обробленого списку тегів.
- 3) Створення векторів частоти входження слів у списки.
- 4) Налаштування параметрів LDA моделі.
- 5) Тренування LDA моделі з використанням отриманих векторів.
- 6) Збереження отриманих векторів та моделі для подальшого використання.
- 7) Перевикористання отриманої моделі для підвищення швидкодії системи.



Подібні серіали. Приклади

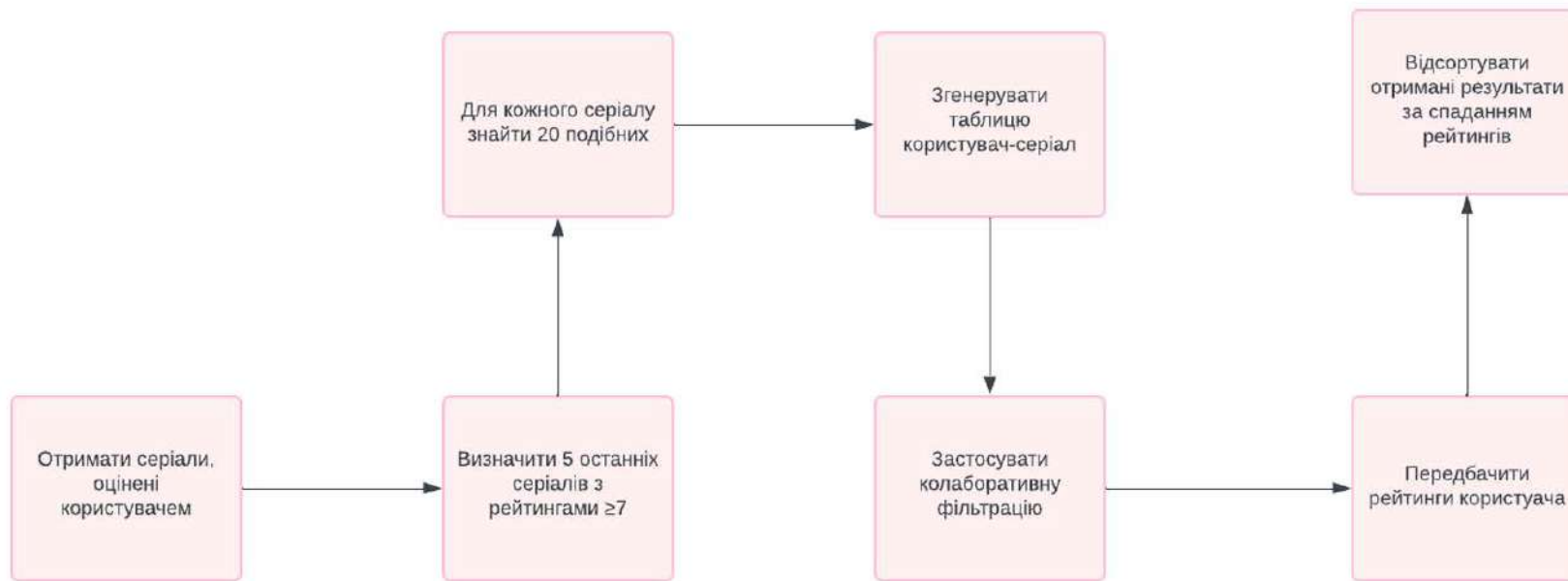


Рекомендації за тематикою до серіалу “Neon Genesis Evangelion” (sci-fi)



Подібні за метаданими серіали до серіалу “Blue Lock” (спортивна тематика)

Алгоритм гібридного методу



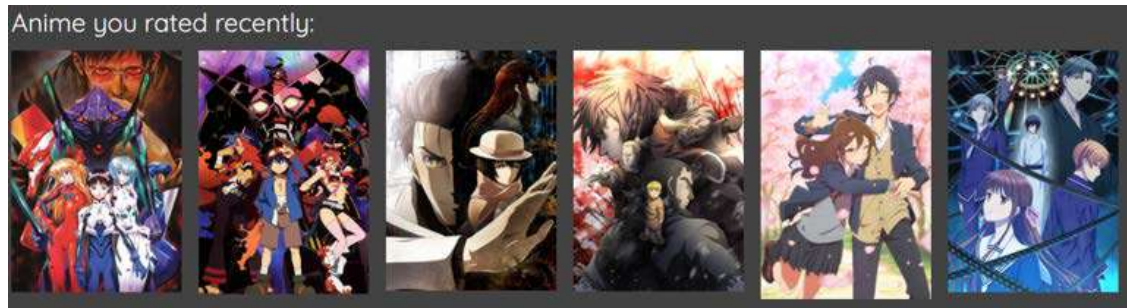
SVD RMSE: 1.3269800722559018
k-NN RMSE: 1.2489596160470882

Порівняння ефективності SVD
та k-NN за критерієм RMSE

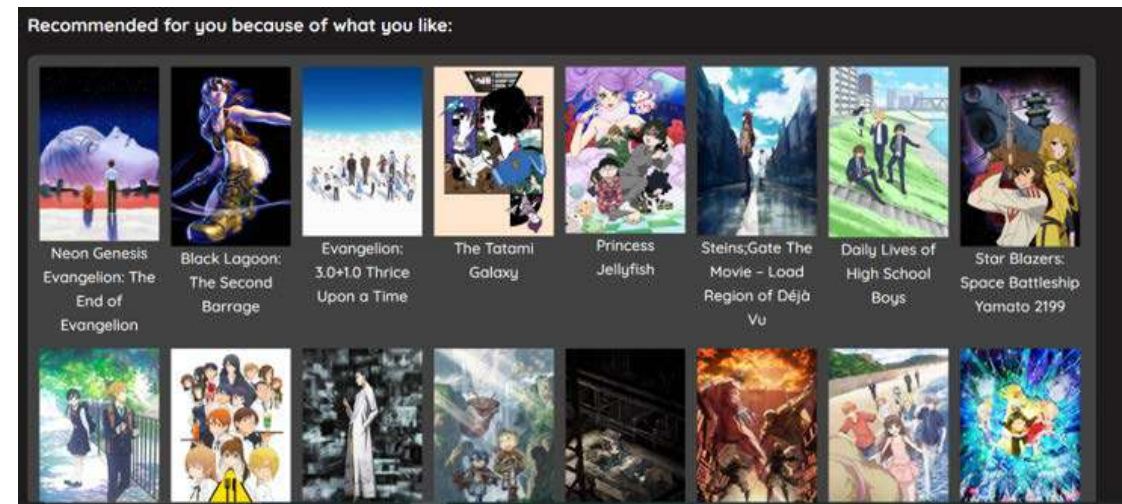
```
[{"user_id": "synthetic1", "anime_id": "6843", "rating": 2}, {"user_id": "synthetic2", "anime_id": "6843", "rating": 4},
```

Згенеровані рейтинги на основі середніх оцінок

Робота гібридного методу



Залишені оцінки



Отримані рекомендації на основі оцінок

Результати

- 1) Реалізовано веб-сайт з використанням сучасних технологій та архітектуру його взаємодії з рекомендаційною системою.
- 2) Натреновано LDA модель для забезпечення швидкодії системи.
- 3) Запропоновано гібридний метод, що забезпечує актуальність рекомендацій, вирішує проблему недостатньої користувацької бази та недостатню персоналізованість класичних колаборативних методів.
- 4) На основі розроблених методів реалізовано архітектуру рекомендаційної системи для веб-сайту.

Дякую за увагу!

Джерела:

- 1) Sun A. A Survey on Modern Recommendation System based on Big Data [Електронний ресурс] / A. Sun, Y. Peng – Режим доступу до ресурсу: <https://arxiv.org/html/2206.02631v2>.
- 2) Falk K. Practical Recommender Systems / Kim Falk. – Shelter Island: Manning, 2019. – 432 с. – (ISBN 9781617292705).