

РЕКОМЕНДАЦІЙНА СИСТЕМА ДЛЯ СТРИМІНГОВОГО СЕРВІСУ ПЕРЕГЛЯДУ ФІЛЬМІВ

Вінницький національний технічний університет

Анотація

У даній роботі досліджено основні підходи щодо розв'язання задачі формування рекомендацій вибору фільмів у стримінговому сервісі. Наведено принцип дії рекомендаційної системи підбору фільмів на основі фільтрації за контентом.

Ключові слова: надання рекомендацій, стримінговий сервіс, потокове відео, перегляд фільмів.

Abstract

This paper explores the main approaches to solving the problem of generating recommendations for selecting films in a streaming service. The principle of operation of a recommendation system for selecting films based on content filtering is presented.

Keywords: providing recommendations, streaming service, streaming video, watching movies.

Вступ

Стримінгові сервіси, які надають доступ до великої кількості відео-, аудіо- і текстового контенту через інтернет, стали важливою частиною сучасного життя. Вони дозволяють глядачам і слухачам переглядати фільми, серіали, веб-шоу, слухати музику, аудіокниги, та навіть читати книги і газети на своїх смартфонах, планшетах та комп'ютерах у будь-який зручний для них час. Це стало можливим завдяки швидкому розвитку інтернет-з'єднань і високоякісному відео- та аудіокодуванню [1].

Метою роботи є розширення функціональних можливостей програмного забезпечення стримінгового сервісу для перегляду фільмів.

Результати дослідження

Потокове відтворення відео – це спосіб передачі медіафайлів, таких як відео, без потреби завантажувати їх із мережі Інтернет. Такі сервіси, як Netflix, Hulu, Megogo та YouTube використовують цей метод для надання можливості переглядати відео на різних типах пристроїв [1].

Потокове відтворення відео стало важливою частиною сучасного споживчого досвіду. Воно дозволяє глядачам зручно переглядати вміст в режимі реального часу, без необхідності очікування завантаження великих файлів на пристрій. Це особливо корисно в епоху швидкісного інтернет-каналу та розширеної мобільності.

На даний момент існує достатня кількість стримінгових сервісів і кожна людина обирає той, який їй подобається більше всього. Кожен з них намагається привернути увагу глядача всім, чим може, будь то унікальність, зручність, доступність, яскравий і приємний оку інтерфейс чи найбільше різноманіття контенту [2].

Одним з аспектів роботи стримінгових сервісів є пропонування до перегляду відеоконтенту, який є близьким до переглянутих відеофільмів чи відеороликів [3].

Існує цілий ряд способів створити більш точну систему генерації рекомендацій для перегляду відеоконтенту користувачем (рис. 1):

- алгоритми фільтрації на основі контенту, що базуються на інформації про жанри, акторів, країни тощо. Ці алгоритми аналізують, які шоу користувач переглянув і поставив лайк, а потім знаходять подібний вміст;

- алгоритми спільної фільтрації, які покладаються на інформацію про минулу поведінку користувачів. Вони аналізують смаки кожного користувача, а потім дають рекомендації на основі вмісту, який подобається користувачам зі схожими смаками;
- алгоритми машинного навчання, які можуть підвищити точність прогнозів, але потребують часу для навчання нейронних мереж.

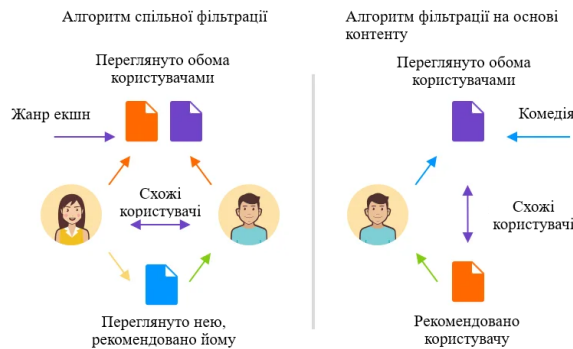


Рис. 1. Схема функціонування рекомендаційних систем

Фільтрація на основі контенту перш за все опирається на інформацію про медіа-контент в системі, а не про користувачів [3]. Тобто рекомендації базуються на знаходженні схожих елементів, до тих, що користувач вже оцінив в минулому. Для цього необхідно створити профіль користувача та профіль елемента. Після цього, на основі параметрів елементів системи, можна зробити висновок щодо відповідності конкретного елемента конкретному користувачу. Для опису елементів системи та створення їх профілю рекомендаційні системи ставлять у відповідність кожному елементу певний набір ключових слів. Формально роботу рекомендаційної системи на основі фільтрації контенту можна зобразити так, як показано на рисунку 2.

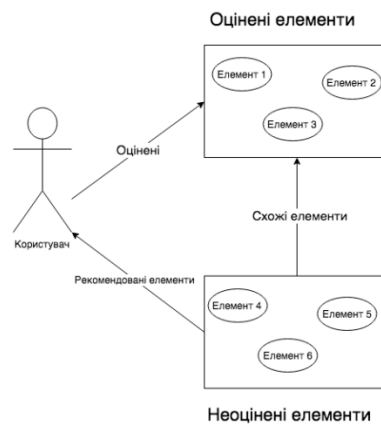


Рис. 2. Принцип дії системи рекомендацій, побудованої на основі аналізу контенту

У рекомендаційних системах з фільтрацією на основі контенту функція зіставлення $h(u, s)$ деякого користувача u з деяким елементом s використовує інформацію про інші елементи системи $s' \in S$, які користувач u вже оцінив в минулому та схожість цих елементів з даним елементом s [2].

Тобто, в загальному випадку, для того щоб рекомендувати користувачу нові елементи, система повинна проаналізувати елементи, які користувач високо оцінив в минулому, знайти ознаки, що об'єднують ці елементи, та на основі знайдених ознак знайти нові елементи, які їм відповідають, і які не мають оцінок від користувача, тобто є новими для нього.

Формально – якщо елемент A сподобався користувачу U , а елемент B схожий на елемент A , то можна зробити висновок, що елемент B також сподобається користувачу U [3].

Таким чином, користувацький профіль формується у вигляді параметрів, що характеризують кожен елемент $s' \in S'$, де S' – множина оцінених елементів. У якості таких параметрів найчастіше ви-

ступають ключові слова, а також вагові коефіцієнти ключових слів для кожного з елементів системи. Одним з найчастіше використовуваних способів знаходження даних вагових коефіцієнтів є TD-IDF міра [3].

Але при цьому враховується лише частоти входження ключового слова. Це може призвести до випадку, коли максимальну вагу будуть мати найбільш розповсюджені ключові слова, що в подальшому може призвести до некоректного прогнозування вподобань конкретного користувача. Для уникнення такого випадку використовується величина *IDF*, яка є зворотною до частоти входження ключового слова в елемент [3].

З множини всіх елементів в системі обираються лише ті, які є найбільш подібними до елементів, що вже були оцінені користувачем в минулому, таким чином будується гіпотеза, що нові елементи будуть також високо оцінені цим же користувачем. Даний набір елементів, що вже були оцінені користувачем, формують його профіль, а також вектор ваг ключових слів, при чому кожна вага визначає важливість певного ключового слова для поточного користувача.

Висновки

Рекомендаційні системи базуються не лише на використанні де-якої евристики та методів інформаційного пошуку, а досить часто використовують інші технології, такі як наївний байєсівський класифікатор, нейронні мережі, дерева рішень, кластеризацію та інші технології машинного навчання. Наприклад, рекомендаційна система, використовуючи оцінки елементів певним користувачем, за допомогою наївного байєсівського класифікатора може визначити елементи, які цей користувач ще не оцінив, та обчислити ймовірність належності певного елемента s_i до певного класу C_i (високо оцінені/низько оцінені елементи).

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Від Netflix до «Такфлікс»: великий гід стримінговими сервісами. The Village Україна – [Електронний ресурс]. – Режим доступу: <https://www.thevillage.com.ua/village/culture/tv/293007-tv-netflix-amazon-disney-plus-appletv-plus-hbo-max-peacock-streaming>.
2. Smith, J. «Streaming Technologies and Their Impact on the Film Industry» (International Journal of Film Studies), 2020, 635p.
3. Wang, L. «User Privacy and Data Protection in Streaming Services: Legal and Ethical Considerations» (Journal of Information Privacy), 2020, 1115p.

Озеранський Володимир Сергійович — канд. техн. наук, доцент кафедри комп'ютерних наук, Вінницький національний технічний університет, Вінниця

Сімончук Сергій Володимирович — асистент кафедри комп'ютерних наук, Вінницький національний технічний університет, Вінниця

Мельник Катерина Ігорівна — студентка групи 2КН-23м, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, Вінниця.

Ярова Олена Андріївна — студентка групи КОІС-21б, факультет інформаційних електронних систем, Вінницький національний технічний університет, Вінниця.

Ozeransky Volodymyr S. — Cand. Sc. (Eng), Assistant Professor of Computer Sciences Chair, Vinnytsia National Technical University, Vinnytsia.

Simonchuk Sergiy V. — Assistant of Computer Science Chair, Vinnytsia National Technical University, Vinnytsia.

Melnyk Kateryna I. — student of Faculty of Intellectual Information Technologies and Automation, Vinnytsia National Technical University, Vinnytsia.

Yarova Olena A. — student of Faculty of Information Electronic Systems, Vinnytsia National Technical University, Vinnytsia.