

# Інформаційна технологія відбору відеоматеріалів. Програмні засоби на основі колаборативної фільтрації

Підготував:  
студент групи 2КН-16м  
Щербань Вадим Сергійович

Науковий керівник:  
д.т.н., проф. Яровий А.А.

# Мета, задачі, об'єкт та предмет дослідження

- ▶ **Мета дослідження** – підвищення ефективності процесу надання рекомендацій щодо відбору відеоматеріалів.
- ▶ **Задачі дослідження:**
  - проаналізувати предметну область рекомендаційних систем;
  - класифікувати та проаналізувати методи надання рекомендацій у рекомендаційних системах;
  - обґрунтувати вибір програмного інструментарію для реалізації інформаційної технології відбору відеоматеріалів, зокрема програмних засобів на основі колаборативної фільтрації;
  - розробити схему алгоритму роботи програмного продукту;
  - здійснити програмну реалізацію та тестування програмного засобу відбору відеоматеріалів.
- ▶ **Об'єкт дослідження** – процес надання рекомендацій щодо відбору відеоматеріалів.
- ▶ **Предмет дослідження** – програмні засоби надання рекомендацій на основі колаборативної фільтрації.

# Актуальність

- **Потреби користувачів** у споживанні конкретної інформації та їх неможливість обробляти всю інформацію привели до стрімкого розвитку технологій пошуку інформації.
- З ростом обсягів даних, що зберігаються у мережі і пропонуються користувачу, зростає актуальність **проблеми випередження запиту** користувача шляхом пропонування йому потенційно цікавої інформації.
- **Рекомендаційні системи** — зручна альтернатива пошуковим алгоритмам, оскільки дозволяють виявити об'єкти, які не можуть бути знайдені останніми.

# Приклади рекомендаційних систем для відбору відеоматеріалів

- ▶ IMDb
- ▶ КиноПоиск
- ▶ Movielens
- ▶ Taste
- ▶ Rotten Tomatoes





# Збір даних про користувачів

## ▶ Приклади явного збору даних:

- користувач оцінює запропонований об'єкт за диференційованою шкалою;
- користувач ранжує групу об'єктів від найкращого до найгіршого;
- користувач вибирає кращий з двох запропонованих об'єктів;
- користувачу пропонують створити список його улюблених об'єктів.

## ▶ Приклади неявного збору даних:

- спостереження за тим, що користувач оглядає в інтернет-магазинах або базах даних іншого типу;
- ведення записів про поведінку користувача онлайн;
- відстеження вмісту комп'ютера користувача.

# Стратегії створення рекомендаційних систем

Дві основні стратегії створення рекомендаційних систем:

## 1. Контентна фільтрація

При фільтрації вмісту створюються профілі користувачів і об'єктів.

Профілі користувачів можуть містити демографічну інформацію або відповіді на певний набір питань.

Профілі об'єктів можуть містити назви жанрів, імена акторів, імена виконавців тощо.

## 2. Колаборативна фільтрація

Використовується інформація про поведінку користувачів у минулому — наприклад, інформація про придбання або оцінки.

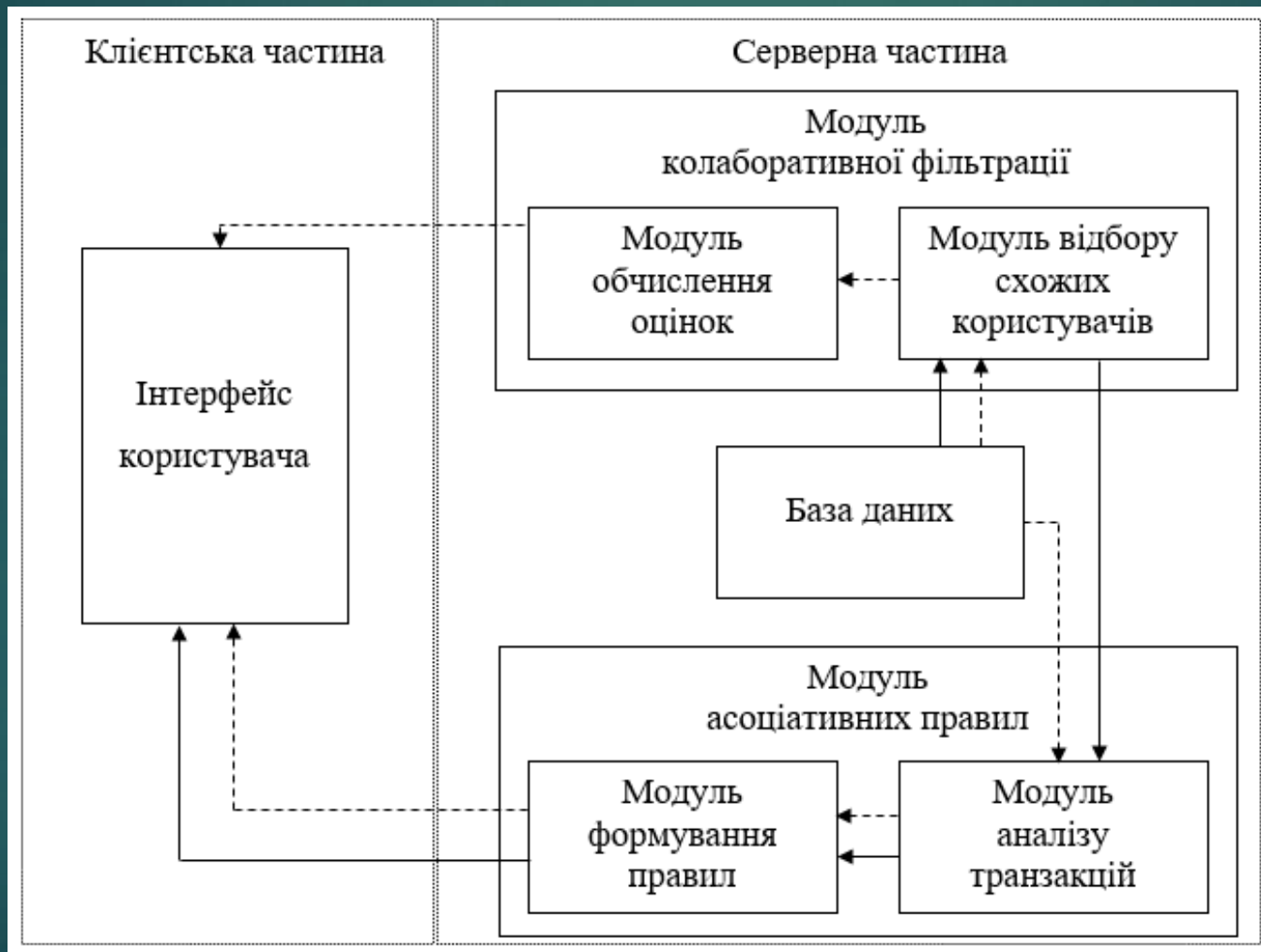
# Колаборативна фільтрація

**Підхід, заснований на сусідстві**, в якому для активного користувача підбирається підгрупа користувачів схожих з ним. Комбінація вагів і оцінок підгрупи використовується для прогнозу оцінок активного користувача.

**Переваги:** легкий та «дешевий» в реалізації.

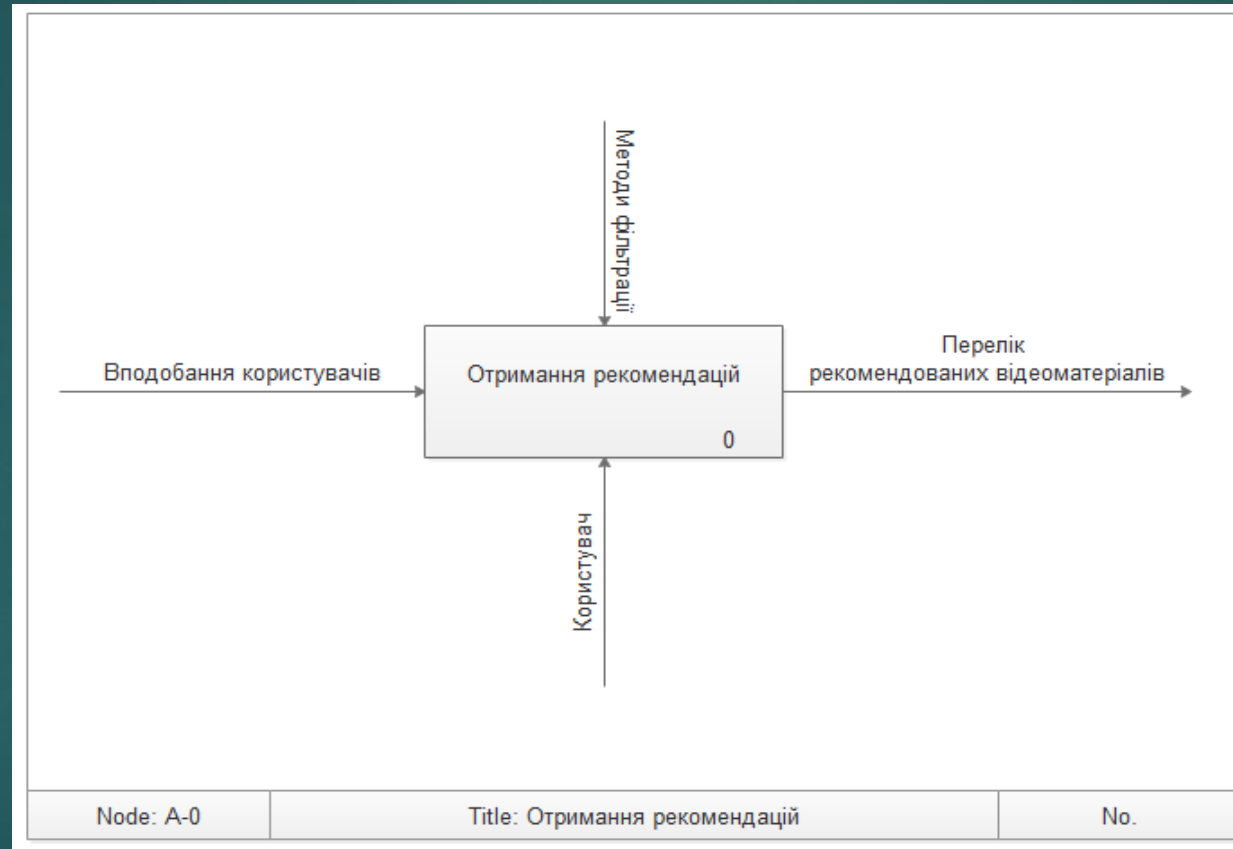
**Недоліки:** погано обробляє розріджені матриці

# Загальна структурна схема





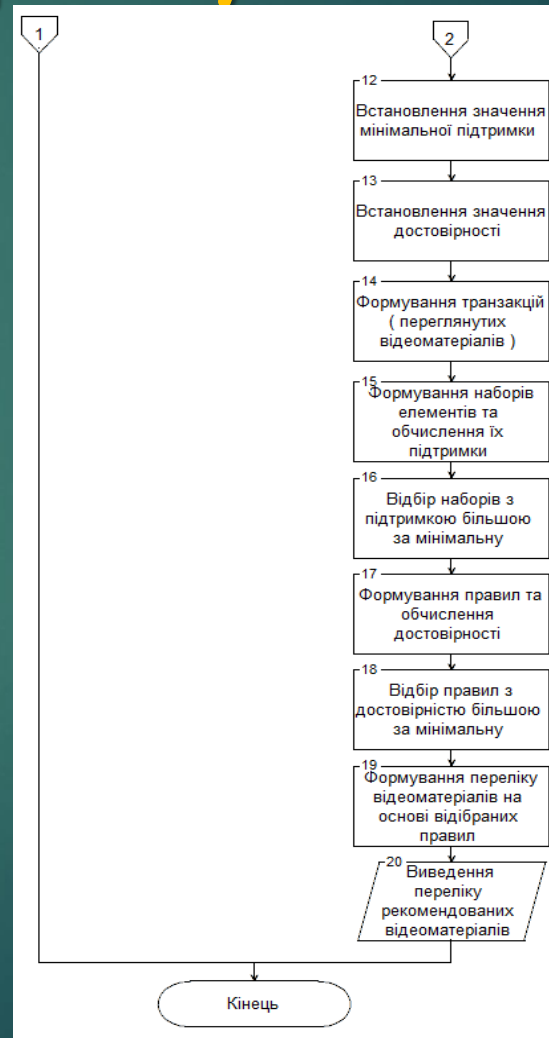
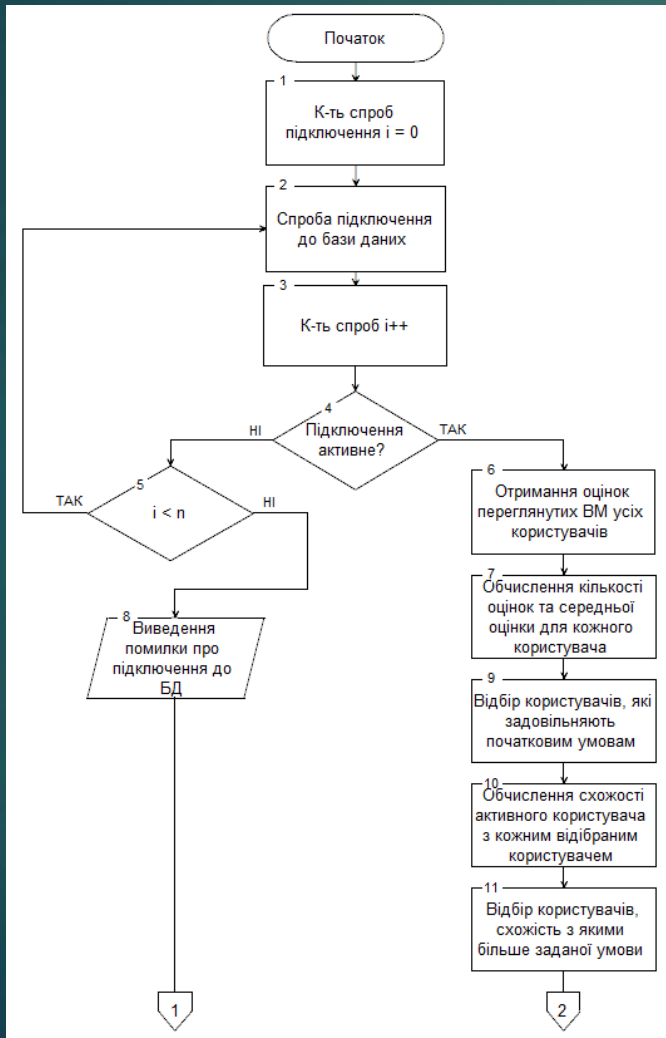
# Контекстна діаграма IDEFO



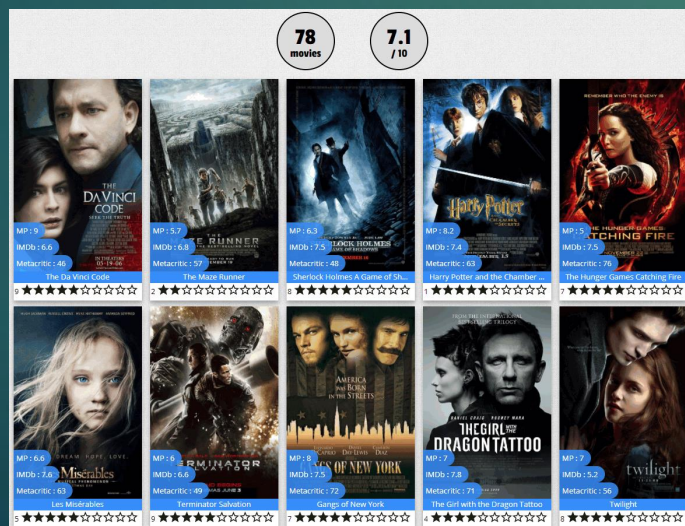
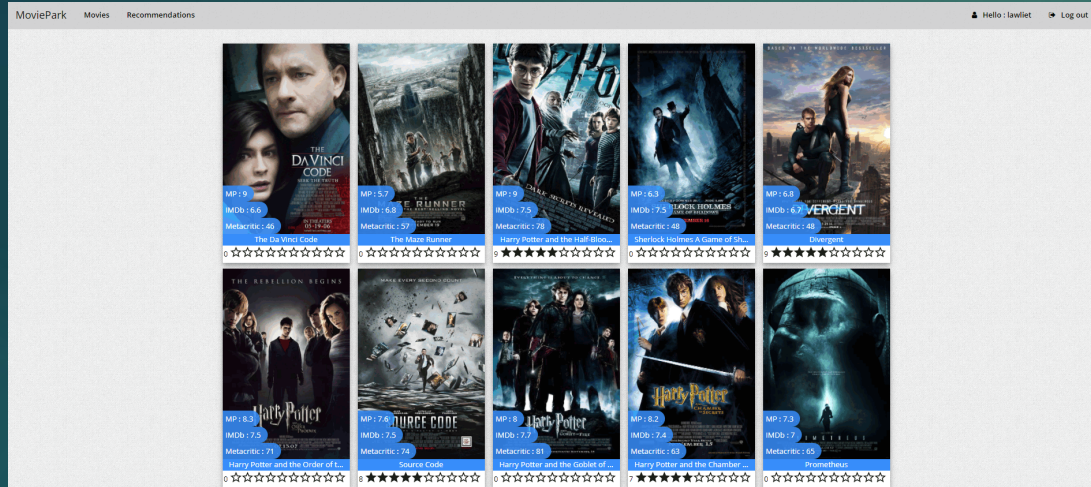
# UML-діаграма варіантів використання



# Схема алгоритму роботи програмного засобу відбору відеоматеріалів на основі комбінованого алгоритму



# Приклад роботи програмного засобу відбору відеоматеріалів



### Les Misérables

Premiere date: 07 February 2013  
Slogan: Fight. Dream. Hope. Love.  
Length: 158 min  
Genres: Drama, History, Romance, Musical, War  
Countries: USA, UK  
Budget: 61000000  
Age limit: PG-13  
Description: In 19th-century France, Jean Valjean, who for decades has been hunted by the ruthless policeman Javert after breaking parole, agrees to care for a factory worker's daughter. The decision changes their lives forever.

Les Misérables - Official Movie Trailer 2012 (HD)

7.6 IMDb 63 Metacritic 6.6 MP



# Тестування роботи програмного засобу відбору відеоматеріалів

- ▶ Для тестування було обрано вибірку з 10 людей, включаючи 2 розробників. На момент тестування в базі даних містилося 300 відеоматеріалів та 10 облікових записів. В ролі активного користувача обрано обліковий запис *Milky\_Way*. Активний користувач оцінив 78 відеоматеріалів. Оцінки користувача є суб'єктивними.
- ▶ Алгоритм колаборативної фільтрації порекомендував 32 відеоматеріали, алгоритм Apriori - 7 відеоматеріалів, а комбінований алгоритм надав 8 рекомендацій.
- ▶ Отримавши рекомендації за допомогою кожного алгоритму, користувач *Milky\_Way* переглянув по 5 випадкових фільмів з кожного переліку та виставив їм оцінки.
- ▶ **Якість** наданих рекомендацій визначається середнім значенням оцінок, які користувач виставить рекомендованим відеоматеріалам після їх перегляду. Чим вище це значення, тим більша якість надання рекомендацій.



# Порівняльна характеристика якості наданих рекомендацій для користувача Milky\_Way

	К-сть рекомендацій	Кількість переглянутих рекомендацій	Сума оцінок переглянутих рекомендацій	Середнє значення $E_{сер}$	Якість Q, %
<u>Колабора-тивна фільтрація</u>	32	5	35	7	70,0
<u>Алгоритм Аргіогі</u>	7	5	38	7,6	76,0
<u>Комбінований алгоритм</u>	8	5	39	7,8	78,0

# Комерційний потенціал розробки інформаційної технології відбору відеоматеріалів

- ▶ Рівень комерційного потенціалу розробки вище середнього.
- ▶ Розраховано витрати на заробітну плату, амортизаційні відрахування, витрати на комплектуючі та на силову електроенергію складають 203741,5 грн.
- ▶ Очікується, що основний прибуток буде отримуватись протягом трьох років. Чистий прибуток підприємства від впровадження розробки за три роки становить 533751,85 грн.
- ▶ Здійснено розрахунок ефективності вкладених інвестицій та періоду їх окупності. Життєвий цикл наукової розробки становить 3 роки. Абсолютна ефективність вкладених інвестицій дорівнює 330010,35 грн. Відносна щорічна ефективність вкладених коштів становить 38%, при мінімальному порозі в 19%. А термін окупності інвестицій складає 2,63 роки. Тому фінансування розробки інформаційної технології відбору відеоматеріалів є доцільним для інвесторів і може принести їм потенційний прибуток.

# Наукова новизна одержаних результатів

- ▶ Удосконалено спосіб відбору відеоматеріалів для надання рекомендацій, що відрізняється від відомих застосуванням комбінації алгоритмів колаборативної фільтрації та асоціативних правил, що забезпечило підвищення якості надання рекомендацій.
- ▶ Запропоновано інформаційну технологію відбору відеоматеріалів на основі колаборативної фільтрації та асоціативних правил, що забезпечило підвищення якості надання рекомендації.

# Практичне значення одержаних результатів

- ▶ Розроблено комбінований алгоритм колаборативної фільтрації та асоціативних правил для підвищення ефективності процесу надання рекомендацій щодо відбору відеоматеріалів;
- ▶ Розроблено програмний засіб надання рекомендацій щодо відбору відеоматеріалів, що дозволяє користувачу отримувати рекомендації, що відповідають його вподобанням;
- ▶ Здійснено практичну реалізацію програмного засобу з використанням сучасних технологій розробки веб-сайтів, програмних платформ та мов програмування: Loorback, Fireloop, Angular, Node.js, TypeScript, що забезпечило швидку програмну реалізацію та підвищення швидкодії обчислення рекомендацій.

# Апробація, публікації та впровадження одержаних результатів

- ▶ Результати досліджень апробовані на XLV та XLVI науково-технічних конференціях професорсько-викладацького складу, співробітників та студентів ВНТУ з участю працівників науково-дослідних організацій та інженерно-технічних працівників підприємств м. Вінниці та області, 2016-2017 рр.
- ▶ За результатами досліджень опубліковано двоє тез доповідей науково-технічних конференцій та подано заяву про реєстрацію авторського права на твір (комп'ютерну програму), реєстраційний номер АПС/2931-17 від 29.11.17 р.
- ▶ Результати даного дослідження протестовано та плануються для подальшого використання у ФОП Носенко С.І., м. Київ.




# Висновки

- ▶ Здійснено аналіз предметної області рекомендаційних систем, визначено основні задачі та вимоги до об'єкту проектування. В роботі проаналізовано існуючі аналоги, охарактеризовано прикладне значення та застосування рекомендаційних систем.
- ▶ Проаналізовано технології та системи надання рекомендацій. Здійснено класифікацію та аналіз методів надання рекомендацій у рекомендаційних системах, визначено основні їх переваги та недоліки.
- ▶ Розроблено загальну структуру програмного засобу, спроектовано IDEFO-, DFD-діаграми та UML-діаграму варіантів використання. Розроблено дві схеми алгоритму роботи програмного продукту: на основі алгоритму колаборативної фільтрації та на основі комбінованого алгоритму.

# Висновки

- ▶ Обґрунтовано вибір програмного інструментарію для реалізації інформаційної технології відбору відеоматеріалів, програмну реалізацію системи здійснено на платформі для створення веб-додатків NodeJS. Програмне забезпечення розроблено з використанням мов програмування високого рівня JavaScript та TypeScript та технології Loopback.
- ▶ Здійснено тестування програмного засобу надання рекомендацій щодо відеоматеріалів. Тестування показало, що на даному тестовому наборі якість рекомендацій, наданих комбінованим алгоритмом на 8% більша, ніж алгоритмом колаборативної фільтрації, та на 2% - в порівнянні з алгоритмом Apriori.
- ▶ Проведено оцінювання економічного потенціалу розробки, спрогнозовано витрати та комерційні ефекти від реалізації результатів розробки, розраховано ефективність вкладених інвестицій та періоду їх окупності.



ДЯКУЮ  
ЗА  
УВАГУ!