

ДОДАТОК А

ЦІЛІ ТА ЗАДАЧІ ДОСЛІДЖЕНЬ

Актуальність	Сучасні підходи до рекомендацій не враховують швидкість обробки і розраховані на рекомендації в режимі онлайн.
Об'єкт дослідження	Процес побудови рекомендацій в соціальних мережах в режимі онлайн.
Предмет дослідження	Методи рекомендацій які використовуються у рекомендаційних системах в режимі онлайн
Мета роботи	Підвищення ефективності методу колаборативної фільтрації для рекомендацій які використовуються у соціальних мережах в режимі онлайн.
Задачі дослідження	<ol style="list-style-type: none">1. Дослідження характеристик рекомендаційних систем.2. Дослідження рекомендаційних систем які призначені для соціальних мереж.3. Порівняльний аналіз методів побудови рекомендаційних систем для соціальних мереж.4. Створення проекту з розробки рекомендаційної системи5. Удосконалення методу колаборативної фільтрації для рекомендацій в режимі онлайн.
Наукова новизна	Удосконалений метод колаборативної фільтрації, який врахує демографічні дані користувачів соціальної мережі в режимі онлайн.
Практичні результати	Підвищення швидкості видачі рекомендацій для користувача в рекомендаційній підсистемі соціальної мережі зі збереженням релевантності.

РІВНІ ПЕРСОНАЛІЗАЦІЙ В РЕКОМЕНДАЦІЙНИХ СИСТЕМАХ

Рекомендаційні системи можна класифікувати за рівнем персоналізації на неперсоніфіковані та персоніфіковані методи. Перші методи не враховують характеристики і переваги клієнтів, в той час як другі тісно залежать від профілю користувача.



КАТЕГОРІЇ РЕКОМЕНДАЦІЙНИХ СИСТЕМ

Виділяються три основні категорії РС, де всі вони є персоналізованими, а саме: колаборативна фільтрація, фільтрація на основі контенту і гібридні методи. Ці три категорії є найбільш популярними і значущими методами рекомендацій.



ТЕХНІКИ ПОБУДОВИ РЕКОМЕНДАЦІЙ

РС можуть бути розділені на різні групи за різних критеріїв. залежно від типу бази та вхідних даних, а також алгоритму, який використовується для створення пропозицій. Вхідні дані надаються користувачами і безпосередньо пов'язані з користувачем, для якого генеруються рекомендації. Вихідні дані є основою, що дозволяє виділити наступні методи рекомендації:



ПОСТАНОВКА ЗАДАЧІ

Існуючі підходи рекомендацій у соціальних мережах найчастіше розраховані на оффлайн рекомендацію і ці методи працюють повільно.

Проблема: удосконалення методу рекомендацій в соціальних мережах в режимі онлайн. Підвищити швидкість обробки не знижуючи релевантність рекомендацій.

ТИПИ СОЦІАЛЬНИХ МЕРЕЖ

Соціальні мережі можна класифікувати в залежності від типу відносин, які пов'язують двох людей. У цьому випадку можна виділити ділові та соціальні зв'язки. Перші містять соціальні мережі, що складаються з людей, які пов'язані один з одним через те, що вони роблять разом, але одночасно вони не розділяють своє приватне життя. Їх можна назвати професійними мережами.



ПІДХОДИ ДО РЕКОМЕНДАЦІЙ У СОЦІАЛЬНИХ МЕРЕЖАХ

Методи рекомендації діляться на інтелектуальну фільтрацію і спрощений підхід. Інтелектуальна фільтрація складається з чотирьох основних категорій: демографічна, контент-орієнтована, колаборативна фільтрація і гібридний метод. До спрощених підходів відносяться методи статистичного аналізу і нерозрахункового підходу. Обидва вони не вимагають складних обчислень і є неперсоніфікованими методами рекомендації.



ПРОФІЛЬ КОРИСТУВАЧА



ДЕМОГРАФІЧНІ ЕЛЕМЕНТИ СОЦІАЛЬНИХ МЕРЕЖ

Interests & Personality	Background & Lifestyle	Schools	Companies	Networking	Name	Basic Info	Profile Songs
<p>Display Name: Kaska</p> <p>MySpace Name: kate_had_to</p> <p>MySpace Url: http://www.myspace.com/kate_had_to</p> <p>First Name: kaska</p> <p>Last Name: musial</p>							

Interests & Personality	Background & Lifestyle	Schools	Companies	Networking	Name	Basic Info	Profile Songs
<p>Gender: <input type="radio"/> Male <input checked="" type="radio"/> Female</p> <p>Date Of Birth: Oct / 28 / 1982</p> <p>Occupation: Student</p> <p>City: Wroclaw</p> <p>State/Region: N/A (US State) [] (Non-US region) []</p> <p>Country: Poland</p> <p>Zip/Postal Code: []</p> <p>Ethnicity: White / Caucasian</p> <p>Body Type: No Answer</p> <p>Height: 0 Feet 0 Inches <input checked="" type="checkbox"/> No Answer</p> <p>I would like to make space for:</p> <ul style="list-style-type: none"><input type="checkbox"/> Dating<input type="checkbox"/> Serious Relationships<input checked="" type="checkbox"/> Friends<input type="checkbox"/> Networking							

ДЕМОГРАФІЧНІ ЕЛЕМЕНТИ СОЦІАЛЬНИХ МЕРЕЖ

Interests & Personality	Background & Lifestyle	Schools	Companies	Networking	Name	Basic Info	Profile Songs
<p>Marital Status: <input type="radio"/> Swinger <input checked="" type="radio"/> In a Relationship <input type="radio"/> Single <input type="radio"/> Divorced <input type="radio"/> Married</p> <p>Sexual Orientation: <input type="radio"/> Bi <input type="radio"/> Lesbian <input checked="" type="radio"/> Straight <input type="radio"/> Not Sure <input type="radio"/> No Answer</p> <p>Hometown: <input type="text" value="Radomsko"/></p> <p>Religion: <input type="text" value="Catholic"/> ▼</p> <p>Smoker: <input type="radio"/> Yes <input checked="" type="radio"/> No <input type="radio"/> No Answer</p> <p>Drinker: <input checked="" type="radio"/> Yes <input type="radio"/> No <input type="radio"/> No Answer</p> <p>Children: <input type="radio"/> I don't want kids <input checked="" type="radio"/> Someday <input type="radio"/> Undecided <input type="radio"/> Love kids, but not for me <input type="radio"/> Proud parent <input type="radio"/> No Answer</p> <p>Education: <input checked="" type="radio"/> High school <input type="radio"/> Some college <input type="radio"/> In college <input type="radio"/> College graduate <input type="radio"/> Grad / professional school <input type="radio"/> Post grad <input type="radio"/> No Answer</p> <p>Income: <input type="text" value="Less than \$30,000"/> ▼</p>							

УДОСКОНАЛЕНИЙ МЕТОД РЕКОМЕНДАЦІЇ В СОЦІАЛЬНИХ МЕРЕЖАХ В РЕЖИМІ ОНЛАЙН

На основі зібраної інформації може бути побудована рекомендаційна структура, що дозволяє створювати пропозиції для соціальної мережі. Основною метою системи є надання найбільш актуальних рекомендацій користувачам. Крім того, аналізуючи кілька різних джерел даних, метод дозволяє новому користувачеві приєднатися до мережі і задовольнити його очікування.

СТАП 1: МОНІТОРИНГ АКТИВНОСТІ ТА СБІР ЗНАНЬ ПРО КОРИСТУВАЧЕЙ

СТАП 2: ПОПЕРЕДНЯ ФІЛЬТРАЦІЯ КОРИСТУВАЧІВ ЗА ДЕМОГРАФІЧНИМИ ОЗНАКАМИ

**СТАП 3: РОЗРАХУВАННЯ ФІНАЛЬНОЇ СХІДНОСТІ МІЖ УСЬОМА ПАРАМЕТРАМИ
КОРИСТУВАЧІВ ВИКОРИСТОВУЮЧИ КОЛАБОРАТИВНУ ФІЛЬТРАЦІЮ**

СТАП 4: ФІЛЬТРАЦІЯ ПО СОЦІАЛЬНИМ ПАРАМЕТРАМ

ЕТАП 2: ПОПЕРЕДНЯ ФІЛЬТРАЦІЯ КОРИСТУВАЧІВ ЗА ДЕМОГРАФІЧНИМИ ОЗНАКАМИ

Вхідні дані:

Множина користувачів $U = \{U_k\}; U_k = \{c, l, a, s\}$,

де: c – країна користувача;

l – мова користувача;

s – соціальний статус користувача;

a – вік користувача.

Обмеження за демографічною ознакою: $R = \{[a_1, a_2], C, L, S\}$

де C – множина країн;

L – множина допустимих мов;

S – множина статусів користувача в соц мережі;

$[a_1, a_2]$ – допустимий діапазон віку користувача.

Відбір користувачів виконується за наступних умов:

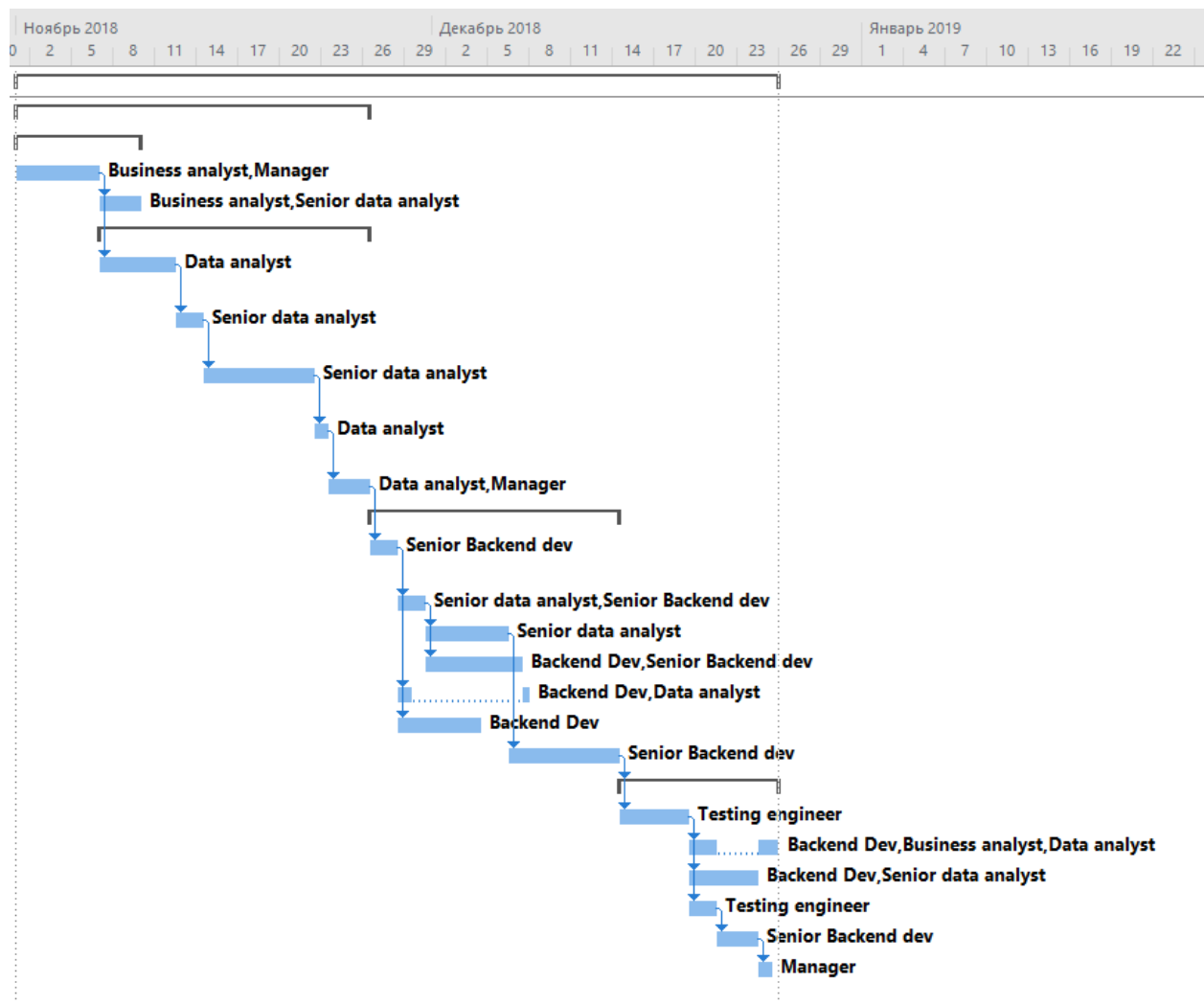
$U^* = \{U_k: \forall U_k c \in C, l \in L, s \in S, a_1 \leq a \leq a_2\}$.

ЗАДАЧІ ПРИ РОЗРОБЦІ ПРОЕКТУ З РЕКОМЕНДАЦІЙ В СОЦІАЛЬНИХ МЕРЕЖАХ

Розробка програмного продукту буде складатися з наступних задач.

➤	➤ Розробка проекту	39 дней	1/11/18	25/12/18		
➤	➤ Планування	17.5 дней	1/11/18	26/11/18		
➤	➤ Визначення потреб	7 дней	1/11/18	9/11/18		
➤	Визначення бізнес цілей	4 дней	1/11/18	6/11/18		Business analyst, Manager
➤	Написання потреб до продукту	3 дней	7/11/18	9/11/18	4	Business analyst, Senior data analyst
➤	➤ Проектування Структури рекомендацій	13.5 дней	7/11/18	26/11/18		
➤	Аналіз існуючих алгоритмів рекомендацій	3.5 дней	7/11/18	12/11/18	4	Data analyst
➤	Вибір та перевірка існуючого методу рекомендацій	2 дней	12/11/18	14/11/18	7	Senior data analyst
➤	Удосконалення структури та алгоритму рекомендації	6 дней	14/11/18	22/11/18	8	Senior data analyst
➤	Симуляція рекомендацій з використанням алгоритму	1 день	22/11/18	23/11/18	9	Data analyst
➤	Презентація алгоритму	1 день	23/11/18	26/11/18	10	Data analyst, Manager
➤	➤ Розробка сервісу рекомендацій	14 дней	26/11/18	14/12/18		
➤	Розробка схеми зберігання даних для рекомендацій	2 дней	26/11/18	28/11/18	11	Senior Backend dev
➤	Імплементация алгоритму на платформі	2 дней	28/11/18	30/11/18	13	Senior data analyst, Senior Backend dev
➤	Розробка конекторів до сервісу	4 дней	30/11/18	6/12/18	14	Senior data analyst
➤	Інтеграція з існуючою системою	5 дней	30/11/18	7/12/18	14	Backend Dev, Senior Backend dev
➤	Тестування стабільності системи	1.5 дней	28/11/18	7/12/18	13	Backend Dev, Data analyst
➤	Інтеграційне тестування на тестовій вибірці	4 дней	28/11/18	4/12/18	13	Backend Dev
➤	Коригування алгоритму	6 дней	6/12/18	14/12/18	15	Senior Backend dev
➤	➤ Завершение	7.5 дней	14/12/18	25/12/18	19	
➤	Тестування на навчальній вибірці	3 дней	14/12/18	19/12/18	19	Testing engineer
➤	Опис документації	3.5 дней	19/12/18	25/12/18	21	Backend Dev, Business analyst, Data analyst
➤	Відладка алгоритму на великій вибірці	3 дней	19/12/18	24/12/18	21	Backend Dev, Senior data analyst
➤	Перевірка та виправлення помилок	2 дней	19/12/18	21/12/18	21	Testing engineer
➤	Аудит	1 день	21/12/18	24/12/18	24	Senior Backend dev
➤	Запуск системи в експлуатацію	1 день	24/12/18	25/12/18	25	Manager

ДИАГРАМА ГАНТА



ВХІДНІ ДАНІ ДЛЯ ОБРОБКИ РЕКОМЕНДАЦІЙ

Розглянемо базову ситуацію в соціальній мережі, коли користувач заходить на свою сторінку та відкриває лист своїх контактів. У багатьох соціальних мережах існує такий блок, як Рекомендовані контакти.

Для цього експерименту, користувач матиме такі поля.

Поле	Значення
ПІБ	Іван Іванов Миколайович
Стать	Чоловік
Дата народження	12.01.1990
Місто	Харків
Мова	Українська,англійська
Активність авторизацій	Мапа (дата:кількість авторизацій)
Інтереси	Мапа(тип інтересу:перелік інтересів)
Соціальний статус	Вільний

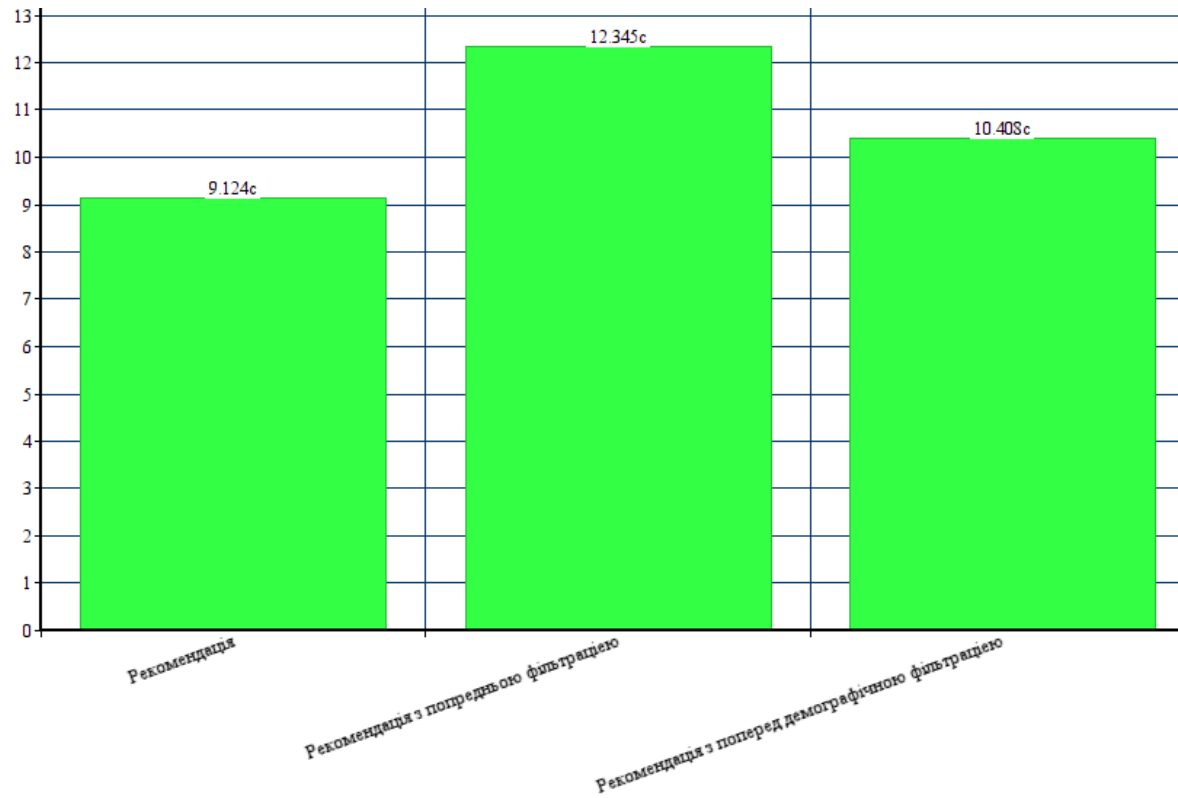
AUC ПОКАЗНИК ДЛЯ РІЗНИХ МЕТОДІВ РЕКОМЕНДАЦІЙ

Метою експерименту є визначення показника AUC – площі, яка обмежена ROC кривою і віссю частки помилкових позитивних класифікацій та швидкості обробки рекомендацій при збереженні показника AUC. Чим вище показник AUC, тим якісніше класифікатор, при цьому значення 0,5 демонструє непридатність обраного методу класифікації.

Метод	AUC
КФ без попередньої фільтрації	(0,862)
КФ з попередньою фільтрацією	(0,880)
КФ з попередньою демографічною фільтрацією	(0,871)

ШВИДКІСТЬ ОБРОБКИ РЕКОМЕНДАЦІЙ

Сгенеруємо вибірку розміром 10 тис. користувачів. Проведемо тіж самі операції але з вимірами по часу. Як результат отримаємо діаграму про затрачений час на обробку рекомендацій.



ВИСНОВКИ

В результаті виконання магістерської роботи був виконаний огляд методів рекомендацій в соціальних мережах. Проаналізовано існуючі методи фільтрацій для рекомендаційних систем.

Була розглянута проблема рекомендацій для соціальних мереж. А саме колаборативної фільтрації з використанням додаткових попередньої фільтрації по демографічному компоненту, який прискорює процес рекомендації і дозволяє збільшити швидкість рекомендацій користувачам соціальних мереж.

В практичному аспекті метод дозволяє покращити показник швидкості від 12,145с. до 10,408 с.