

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
(повна назва)

Кафедра Інформаційних управляючих систем
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА Пояснювальна записка

рівень вищої освіти другий (магістерський)

Дослідження методів побудови рекомендацій в інформаційній системі міжнародної організації AISEC для волонтерських стажувань молоді
(тема)

Виконав:
здобувач 2 року навчання,
групи ІУСТМ-24-1

Софія КОСЯЧУК
(Власне ім'я, ПРІЗВИЩЕ)

Спеціальність 122 Комп'ютерні науки
(код і повна назва спеціальності)

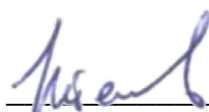
Тип програми освітньо-професійна
(освітньо-професійна або освітньо-наукова)

Освітня програма Інформаційні управляючі системи та технології
(повна назва освітньої програми)

Керівник доц. каф. ІУС Аліна МІХНОВА
(посада, власне ім'я, ПРІЗВИЩЕ)

Допускається до захисту

Зав. кафедри


(підпис)

Костянтин ПЕТРОВ
(Власне ім'я, ПРІЗВИЩЕ)

2025 р.

Харківський національний університет радіоелектроніки

Факультет _____ Комп'ютерних наук _____
Кафедра _____ Інформаційних управляючих систем _____
Рівень вищої освіти _____ другий (магістерський) _____
Спеціальність _____ 122 Комп'ютерні науки _____
(код і повна назва)
Тип програми _____ освітньо-професійна _____
(освітньо-професійна або освітньо-наукова)
Освітня програма _____ Інформаційні управляючі системи та технології _____
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

« 24 » листопада 20 25 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві _____ Косячук Софії Леонідівні _____
(прізвище, ім'я, по батькові)

1. Тема роботи Дослідження методів побудови рекомендацій в інформаційній системі міжнародної організації AIESEC для волонтерських стажувань молоді затверджена наказом університету від «24» листопада 2025 р. № 1055Ст
2. Термін подання здобувачем роботи до екзаменаційної комісії «17» грудня 2025 р.
3. Вихідні дані до роботи наукова література з тематики кваліфікаційної роботи, матеріали передатестаційної практики, матеріали про методи рекомендації, джерела інформації про рекомендаційні системи та волонтерські платформи
4. Перелік питань, що потрібно опрацювати в роботі аналіз предметної області, опис діяльності волонтерських організацій, аналіз науково-технічної літератури у сфері побудови персональних рекомендаційних систем, огляд існуючих інформаційних систем-аналогів для підбору волонтерських стажувань, опис існуючих методів побудови інформаційної системи для вирішення задачі персоналізованого підбору волонтерських проєктів для молоді, постановка задачі дослідження, дослідження методів для формування рекомендацій проєктів AIESEC, інформаційна технологія побудови рекомендації в інформаційній системі, практична реалізація гібридного методу, експериментальна оцінка гібридного методу

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Аналіз предметної області	24.11.2025 – 26.11.2025	Виконано
2	Опис існуючих методів побудови ІС для підбору волонтерських проєктів для молоді	27.11.2025 – 29.11.2025	Виконано
3	Постановка задачі дослідження	30.11.2025 – 01.12.2025	Виконано
4	Дослідження методів для формування рекомендацій	01.12.2025 – 02.12.2025	Виконано
5	Інформаційна технологія побудови рекомендацій в ІС	02.12.2025 – 03.12.2025	Виконано
6	Практична реалізація гібридного методу	03.12.2025 – 06.12.2025	Виконано
7	Експериментальна оцінка гібридного методу	06.12.2025 – 08.12.2025	Виконано
8	Оформлення пояснювальної записки	08.12.2025 – 09.12.2025	Виконано
9	Оформлення графічних матеріалів	09.12.2025 – 10.12.2025	Виконано
10	Захист кваліфікаційної роботи	17.12.2025	Виконано

Дата видачі завдання «24» листопада 2025 р.

Здобувач _____


(підпис)

Керівник роботи _____


(підпис)

доц. каф. ІУС Аліна МІХНОВА
(посада, власне ім'я, ПРІЗВИЩЕ)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 97 с., 17 рис., 2 табл., 2 дод., 31 джерел.

АІЕSEC, ГІБРИДНІ МЕТОДИ, ІНФОРМАЦІЙНА СИСТЕМА, КОЛАБОРАТИВНА ФІЛЬТРАЦІЯ, КОНТЕНТНО-ОРІЄНТОВАНИЙ МЕТОД, РЕКОМЕНДАЦІЙНІ МОДЕЛІ

Об'єктом дослідження є процес побудови рекомендацій у інформаційній системі волонтерських стажувань АІЕSEC.

Предметом дослідження є методи формування рекомендацій та їх поєднання для якісного підбору волонтерських проєктів.

Метою роботи є аналіз сучасних підходів до рекомендаційних систем, оцінка їх ефективності у контексті волонтерських програм та розробка гібридного методу рекомендацій для системи АІЕSEC.

У роботі проведено аналіз предметної області та діяльності АІЕSEC, розглянуто платформи-аналоги та визначено їхні технічні обмеження. Досліджено основні методи рекомендацій: контентно-орієнтовані моделі, колаборативну фільтрацію на основі взаємодій та атрибутивну колаборативну фільтрацію.

Розроблено інформаційну технологію рекомендацій для АІЕSEC, що охоплює модулі збору даних, формування профілів користувачів і проєктів, вибору рекомендацій та ранжування. Запропоновано гібридний метод, який поєднує контентний аналіз з двома видами колаборативної фільтрації, що підвищує точність добору проєктів для різних категорій користувачів.

Проведена експериментальна оцінка підтвердила вищу точність і релевантність гібридного підходу. Розроблений метод може бути використаний як основа для впровадження рекомендаційної підсистеми в ІС АІЕSEC.

ABSTRACT

Explanatory note of the qualification work: 97 p., 17 fig., 2 t., 2 appendix, 27 sources.

AIESEC, HYBRID METHODS, INFORMATION SYSTEM, COLLABORATIVE FILTERING, CONTENT-ORIENTED APPROACH, RECOMMENDATION MODELS

The object of the study is the process of building recommendations in the AIESEC volunteer internship information system.

The subject of the study is the methods of forming recommendations and their combination for the high-quality selection of volunteer projects.

The aim of the work is to analyze modern approaches to recommendation systems, assess their effectiveness in the context of volunteer programs and develop a hybrid recommendation method for the AIESEC system.

The work analyzes the subject area and activities of AIESEC, considers similar platforms and identifies their technical limitations. The main recommendation methods are studied: content-oriented models, collaborative filtering based on interactions and attributive collaborative filtering.

An information technology of recommendations for AIESEC has been developed, which includes modules for data collection, user and project profiling, recommendation selection and ranking. A hybrid method is proposed that combines content analysis with two types of collaborative filtering, which increases the accuracy of project selection for different categories of users.

The experimental evaluation confirmed the higher accuracy and relevance of the hybrid approach. The developed method can be used as a basis for implementing a recommender subsystem in the AIESEC.

ЗМІСТ

Скорочення та умовні позначки	9
Вступ.....	10
1 Аналіз предметної області.....	12
1.1 Опис діяльності волонтерських організацій, що є предметною областю та інформаційних систем, що використовуються	12
1.1.1 Коротка характеристика AIESEC	12
1.1.2 Модель діяльності та продуктові напрямки.....	13
1.1.3 Інформаційні системи AIESEC.....	13
1.2 Аналіз науково–технічної літератури у сфері побудови персональних рекомендаційних систем	15
1.2.1 Опис рекомендаційних систем	15
1.2.2 Функціональні особливості рекомендаційних систем	17
1.2.3 Приклади реалізації популярних рекомендаційних систем	20
1.3 Огляд існуючих інформаційних систем–аналогів для підбору волонтерських стажувань. Аналіз переваг та недоліків.....	23
1.3.1 Особливості ринку міжнародних молодіжних програм	23
1.3.2 Огляд ключових систем-аналогів, подібних до AIESEC	24
1.3.3 Порівняльний аналіз інформаційної системи AIESEC та платформ–аналогів.....	27
1.4 Опис існуючих методів побудови інформаційної системи для вирішення задачі персоналізованого підбору волонтерських проєктів для молоді	29
1.4.1 Контентно-орієнтовані моделі.....	29
1.4.2 Моделі колаборативної фільтрації	30
1.4.3 Атрибутивна колаборативна фільтрація.....	30
1.4.4 Гібридні моделі	32
1.4.5 Контекстно–залежні моделі	32

1.5	Постановка задачі дослідження.....	33
2	Дослідження методів для формування рекомендацій проєктів AIESEC	36
2.1	Основні проблеми при побудові рекомендаційних систем	36
2.2	Дослідження колаборативної фільтрації на основі взаємодії користувачів.....	37
2.3	Дослідження атрибутивної колаборативної фільтрації	39
3	Інформаційна технологія побудови рекомендації в інформаційній системі волонтерських стажувань AIESEC	46
3.1	Опис контекстної діаграми	46
3.2	Опис діаграми декомпозиції	47
4	Практична реалізація гібридного методу	49
4.1	Програмна реалізація методів рекомендаційної системи	49
4.1.1	Обґрунтування вибору технологічних засобів	49
4.1.2	Реалізація контентної фільтрації	50
4.1.3	Реалізація колаборативної фільтрації. Вирішення проблеми «холодного старту»	51
4.1.4	Реалізація гібридного методу	53
4.1.5	Демонстрація рекомендаційної системи із застосуванням реалізованого гібридного методу	55
4.2	Експериментальна оцінка гібридного методу.....	65
4.2.1	Опис метрик для експериментальної оцінки	65
4.2.2	Результати експерименту	68
	Висновки	72
	Перелік джерел посилання	74
	Додаток А Фрагменти коду	78
A.1	Реалізація контентно-орієнтованої моделі з використанням TD-IDF ..	78
A.2	Реалізація колаборативної фільтрації на основі атрибутів	79
A.3	Реалізація колаборативної фільтрації на основі взаємодії користувачів.....	80
A.4	Генерація кандидатів	81

А.5 Реалізація гібридного методу.....	82
Додаток Б Графічний матеріал кваліфікаційної роботи	85

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ

ІС – інформаційна система

ІУС – інформаційні управляючі системи

КФ – колаборативна фільтрація

ООН – Організація Об'єднаних Націй

СУБД – система управління базами даних

ХНУРЕ – Харківський національний університет радіоелектроніки

AIESEC – Association internationale des étudiants en sciences économiques et commerciales

AP – average precision

DCG – discounted cumulative gain

ESC – European Solidarity Corps

EXPA – Experience AIESEC

IAESTE – International Association for the Exchange of Students for Technical Experience

IDCG – ideal discounted cumulative gain

IVHQ – International Volunteer HQ

MAP – mean average precision

NDCG – normalized discounted cumulative gain

NPS – net promoter score

TF-IDF – term frequency – inverse document frequency

ВСТУП

На сьогоднішній день, молодь часто отримує багато пропозицій щодо навчання, кар'єри чи допомоги через волонтерство. Організація AIESEC надає численні місця для волонтерства за кордоном – але великий вибір ускладнює якісне прийняття рішення. Тож важливо створити систему, яка рекомендувала б підходящі варіанти на основі особистих переваг, минулого досвіду й мети. Такий інструмент може полегшити пошук актуального проєкту.

Рекомендаційні системи – важливий елемент сучасних платформ, оскільки допомагають налаштувати взаємодію з людьми й прискорюють пошук необхідного продукту. Натомість у багатьох випадках використовують методи аналізу вмісту чи взаємодії, а також їх комбінації, особливо в онлайн-торгівлі, навчанні та соцмережах. Однак для волонтерства ці підходи мають обмеження: проєкти часто розрізняються між собою, критерії входження можуть швидко змінюватися, і починаючи з нуля, система працює не якісно. Крім того, доводиться враховувати внутрішні особливості кожної конкретної людини.

Об'єктом дослідження є процес формування персоналізованих рекомендацій у інформаційній системі волонтерських стажувань.

Предметом дослідження є методи побудови рекомендацій та їхнє комбінування в інформаційній системі AIESEC для підвищення якості добору волонтерських проєктів.

Актуальність дослідження зумовлена необхідністю покращити процес пошуку можливостей для молодих людей і поліпшити досвід роботи з AIESEC. Створення комбінованого способу, який об'єднує аналіз матеріалів і спільні оцінки учасників, допоможе якісно пропонувати більш персоналізовані варіанти.

Мета кваліфікаційної роботи – вивчення існуючих методів побудови рекомендацій, оцінка їх переваг та недоліків сторони, створення комбінованого способу персоналізації рекомендацій волонтерських проєктів AIESEC.

Для досягнення мети буде виконано аналіз предметної області, проведено

огляд інформаційних систем AIESEC, розглянуто сучасні науково–технічні підходи до рекомендаційних технологій, визначено недоліки існуючих моделей, розроблено інформаційну технологію рекомендацій, реалізовано гібридний метод та проведено експериментальну оцінку його ефективності.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Опис діяльності волонтерських організацій, що є предметною областю та інформаційних систем, що використовуються

1.1.1 Коротка характеристика AIESEC

AIESEC – молодіжний лідерський рух і платформа розвитку, створена та керована молоддю для молоді. Базова місія формується як досягнення миру та реалізації людського потенціалу шляхом надання молодим людям практичних лідерських досвідів і середовища для дій відповідно до цінностей організації.

Організація виникла після Другої світової війни з ініціативи європейських студентів; політичний і соціальний контекст повоєнної Європи сформував акцент AIESEC на міжкультурному порозумінні, толерантності та запобіганні конфліктам через відмінності (культурні, релігійні тощо). У сучасному трактуванні «мир» розглядається не лише як відсутність війни, а як стан гармонійної взаємодії між людьми та всередині особистості.

Організація декларує унікальну модель розвитку лідерства:

- цінності задають орієнтири для щоденних рішень і формують культуру;
- кожен лідерський досвід (участь у програмах, командна робота) має бути носієм цих цінностей і спрямованим на довгостроковий позитивний вплив;
- очікуваним результатом повинна стати позиційність молодої людини щодо важливих для неї питань та послідовне «зростання до кращої версії себе» з одночасним створенням можливостей для інших.

AIESEC пропонує спектр можливостей – обміни/стажування, волонтерство та інші активності – які навмисне виводять учасника за межі звичного середовища і розвивають лідерські компетентності через практику.

Партнерами вважаються фізичні особи, корпорації, інституції та організації, які поділяють цінності AIESEC, дотримуються її етичних принципів і підсилюють вплив – зокрема через надання можливостей, ресурсів або експертизи [1].

1.1.2 Модель діяльності та продуктові напрямки

Ядром діяльності є надання міжкультурних лідерських досвідів через програми стажувань і членства. Ключовий волонтерський продукт – Global Volunteer. Це короткострокові проекти, що прив'язані до Цілей сталого розвитку ООН і спрямовані на суспільну користь.

Організація описує стандартний шлях кандидата:

- реєстрація;
- пошук проекту;
- відбір/співбесіди;
- підготовка до виїзду та підтримка від AIESEC.

Всеукраїнська молодіжна організація «AIESEC в Україні» деталізує умови програм (житло, гнучкі дати, підтримка менеджерів) та типові напрями впливу [2].

Другий напрям – Global Talent (професійні стажування тривалістю 2–18 місяців у сферах маркетингу, фінансів, ІТ тощо). Всеукраїнська молодіжна організація «AIESEC в Україні» фіксує етапи подачі (реєстрація, пошук, співбесіда, підготовка, стажування) й надає вимоги до кандидатів та інформацію про організаційний внесок [3].

Окремо в Україні представлена програма Local Volunteer – участь у локальних соціальних проектах разом з іноземними волонтерами (2–6 тижнів) і ключові метрики локального впливу (соціальні проекти, охоплення бенефіціарів тощо).

1.1.3 Інформаційні системи AIESEC

AIESEC включає користувацькі платформи для кандидатів/партнерів та

внутрішні системи для членів організації.

AIIESEC Opportunity Portal – це єдиний глобальний каталог можливостей з пошуком, фільтрами та подачею заявок на волонтерські програми. Кандидати реєструються через центр автентифікації і далі працюють зі своїм профілем та заявками. Для організацій–партнерів передбачено розміщення можливостей і процес відбору кандидатів на платформі.

Функціональні можливості інформаційної системи для користувачів:

- аутентифікація та обліковий запис – централізований логін/реєстрація, керування профілем, перегляд «my applications», є офіційні інструкції логіну та видалення акаунта;

- пошук і відбір можливостей – глобальний каталог із фільтрами сортування і картки можливостей; сторінки продуктів описують покроковий шлях кандидата;

- комунікації та трекінг – після подачі кандидат виходить на відбір (інструкції/зв'язок через менеджерів можливостей), прогрес відслідковується статусами заявки.

Національний сайт AIIESEC в Україні – інформаційна система українського національного комітету із деталями продуктів, вимог, вартісних умов і локальної специфіки. Заявки на консультацію/перший контакт здебільшого збираються через інтегровані форми Turiform. Наявна політика конфіденційності, що регулює умови використання сайту та обробку даних [4].

EXPA – внутрішня операційно–CRM платформа AIIESEC для членів організації, побудована навколо стандартного потоку клієнтів та портфельних продуктів. Її використовують члени AIIESEC для управління різними аспектами своєї діяльності, включаючи відстеження можливостей стажерів, а також їхніх особистих профілів та профілів комітетів. Доступ отримується за допомогою створення облікового запису AIIESEC та отримання запрошення.

Ключові функції даної системи:

- ведення профілів стажерів, їхніх заявок, членського статусу, призначення менеджерів; у тому числі, перегляд профільної сторінки;

- управління можливостями локального комітету;
- моніторинг і оновлення статусів заявок, нотатки/теги, таймлайни руху по етапах, NPS/стандарти;
- облік і сегментація партнерських організацій;
- довідник структур AIESEC, у тому числі, налаштування параметрів порталу можливостей для молоді на рівні сутності (програмні внески, LC-assignment, кампанії, візи/страхування, транзакційні листи).
- аналітика за метриками діяльності та ресурсний центр знань для AIESECсерів [5].

1.2 Аналіз науково–технічної літератури у сфері побудови персональних рекомендаційних систем

1.2.1 Опис рекомендаційних систем

У сучасному цифровому середовищі кількість інформації, з якою стикається користувач, зростає в геометричній прогресії. Людина фізично не може переглянути всі доступні товари, статті, вакансії чи освітні можливості, тому виникає потреба у механізмах, що допомагають швидко знаходити найбільш релевантні результати. Саме цю функцію виконують рекомендаційні системи – інтелектуальні підсистеми, які аналізують дані про користувачів і об'єкти та формують персональні пропозиції.

Рекомендаційна система – це сукупність програмних і аналітичних засобів, призначених для автоматизованого відбору об'єктів, що найімовірніше відповідають інтересам конкретного користувача. За визначенням Resnick і Varian, рекомендаційні системи є механізмами, які допомагають приймати рішення в умовах надлишку інформації, створюючи персоналізований інформаційний простір для кожного користувача [6]. Рекомендаційні системи перетворюють великі масиви даних на індивідуальні рекомендації, скорочуючи

час пошуку й підвищуючи задоволеність користувача результатом.

В основі будь-якої рекомендаційної системи лежить збір, зберігання та аналіз даних про попередні дії користувачів. Цими діями можуть бути перегляди сторінок, оцінки, покупки, подані заявки або навіть час, проведений на певному ресурсі. На основі таких даних система формує профіль користувача, який описує його вподобання, стиль поведінки та ймовірні потреби. Паралельно формується профіль об'єкта, який містить його описові характеристики.

Після цього система обчислює ступінь відповідності між профілем користувача та профілем об'єкта, використовуючи різноманітні алгоритмічні підходи. На практиці це означає, що рекомендаційна система постійно оцінює, наскільки певний об'єкт цікавий для конкретного користувача, і формує список рекомендацій. Важливо підкреслити, що сучасні системи працюють не лише з явними оцінками, але й з неявними сигналами – такими як час перегляду, кількість натискань або частота повернення до певного розділу сайту. Це дозволяє моделі навчатися навіть тоді, коли користувач не залишає прямих відгуків.

Наукові праці Adomavicius і Tuzhilin виділяють кілька ключових напрямів розвитку рекомендаційних систем – від класичних контентних і колаборативних методів до гібридних та контекстно-залежних систем. В останніх враховується не лише історія взаємодій, а й поточний контекст – місце перебування, час доби, пристрій, погода, соціальне оточення тощо. Це дає змогу будувати динамічні моделі персоналізації, які реагують на зміну поведінки користувача в реальному часі [7].

З точки зору інформаційних технологій, рекомендаційна система є комплексом програмних модулів, інтегрованих у більшу інформаційну систему. Її структура зазвичай містить такі компоненти: модуль збору та попередньої обробки даних, модуль зберігання профілів користувачів та об'єктів, аналітичний модуль для розрахунку рекомендацій та інтерфейс користувача для відображення результатів. У промислових реалізаціях ці компоненти функціонують у середовищі великих даних та машинного навчання, що дозволяє

обробляти мільйони запитів у реальному часі.

Окрім чисто технічних переваг, рекомендаційні системи мають і значний соціальний ефект. Вони впливають на те, яку інформацію бачить користувач, формують його цифрові звички та можуть навіть спрямовувати професійний чи освітній вибір.

Таким чином, рекомендаційна система є не лише технічним засобом покращення зручності користувача, а й важливим елементом сучасних інформаційних екосистем.

1.2.2 Функціональні особливості рекомендаційних систем

Сучасні рекомендаційні системи є складними багаторівневими інформаційними комплексами, у яких поєднуються технології збирання, зберігання та аналітичної обробки великих обсягів даних. Вони інтегрують у собі як елементи штучного інтелекту, так і класичні принципи проектування інформаційних систем – масштабованість, надійність, швидкодію та модульність. Основна функціональна мета таких систем полягає у забезпеченні швидкого, гнучкого та точного формування персоналізованого контенту для кожного користувача [7].

У типовій архітектурі рекомендаційна система складається з кількох логічних рівнів:

- збір даних;
- збереження та обробка інформації;
- аналітичний рівень;
- рівень представлення результатів.

Ці рівні взаємопов'язані між собою через єдину інфраструктуру даних, що підтримує обмін інформацією у режимі реального часу.

На рівні збору даних система отримує інформацію про взаємодію

користувача з платформою. Джерелами можуть бути логи вебсайту, мобільні додатки, форми зворотного зв'язку або інтегровані зовнішні сервіси. У великих проєктах, таких як YouTube чи LinkedIn, для цього використовуються системи потокової обробки подій, які дозволяють обробляти мільйони дій користувачів щосекунди [8]. Кожна подія – перегляд, натискання, перехід, додавання у список – перетворюється на структуровану сутність, що зберігається у базі даних.

Наступним етапом у роботі рекомендаційної системи є збереження та підготовка інформації. Після збору дані проходять низку процедур, спрямованих на приведення їх до уніфікованої форми. Відбувається очищення від шумів, усунення дублювань, нормалізація значень, заповнення пропусків і категоризація показників. Така обробка є необхідною, оскільки дані, зібрані з різних джерел, можуть мати різну структуру, точність і формат подання. На цьому етапі формується набір характеристик – ознаки, які в подальшому використовуються для аналітичного аналізу та побудови профілів користувачів і об'єктів.

Для забезпечення узгодженості між етапом навчання моделі та її реальним функціонуванням застосовуються спеціалізовані сховища ознак. Такі сховища відіграють ключову роль у сучасній архітектурі рекомендаційних систем, оскільки дають змогу уникнути розбіжностей між даними, що використовувалися під час тренування алгоритмів, і тими, які подаються до моделі під час роботи в реальному часі. Подібні рішення детально описані в інженерних дослідженнях Uber AI, які стали еталоном побудови інфраструктури для обробки ознак у промислових умовах [9].

Після підготовки даних формується аналітичний рівень, що є ядром рекомендаційної системи. На цьому етапі відбувається пошук релевантних об'єктів, їхня фільтрація, ранжування та формування фінального списку. Для прискорення процесу в архітектурі зазвичай передбачено два основні контури:

- контур попереднього відбору;
- контур формування кінцевих результатів.

Така структура дозволяє скоротити час відповіді системи до кількох

десятків мілісекунд навіть за великої кількості користувачів [9].

Важливою частиною є постобробка результатів, де рекомендації проходять додаткову перевірку на відповідність певним обмеженням і політикам платформи. Наприклад, для волонтерських або соціальних систем враховується географічна доступність програми, терміни проведення, рівень підготовки кандидата, а також вимоги партнерських організацій. Подібна логіка дозволяє адаптувати систему не лише до інтересів користувача, а й до особливостей контенту, який вона обробляє.

Рівень представлення результатів забезпечує передачу рекомендацій користувачу у зручній формі. Це може бути інтеграція у головну сторінку сайту, персональний розділ користувача або модуль “схожих пропозицій”. На цьому етапі важливо дотримуватись балансу між релевантністю рекомендацій і зручністю інтерфейсу. Дослідження Netflix показують, що навіть дрібні зміни у способі відображення рекомендацій можуть суттєво вплинути на взаємодію користувача з системою [10].

Не менш важливими є підсистеми моніторингу та оновлення даних. Вони відстежують стабільність роботи системи, коректність джерел даних і швидкість оновлення. Також вони забезпечують повторне тренування моделей і адаптацію до змін у поведінці користувачів. Для великих платформ цей процес є безперервним: дані постійно надходять, агрегуються, перевіряються й перетворюються на нові версії рекомендацій [10].

Технічні вимоги до таких систем передбачають високу масштабованість, відмовостійкість та низьку затримку. Наприклад, рекомендаційна система LinkedIn формує результати менше ніж за 100 мілісекунд, забезпечуючи одночасну роботу для мільйонів користувачів [11].

Отже, функціональні та технічні особливості рекомендаційних систем визначають їхню здатність працювати стабільно й ефективно у динамічному інформаційному середовищі. Успішна система не лише формує релевантні рекомендації, а й забезпечує постійну адаптацію до змін користувацької поведінки, високу швидкодію та безпеку даних. Саме такі принципи мають бути

покладені в основу побудови рекомендаційної системи для інформаційної платформи волонтерських стажувань AIESEC.

1.2.3 Приклади реалізації популярних рекомендаційних систем

Рекомендаційні системи сьогодні є фундаментальним елементом більшості великих цифрових платформ. Їхня ефективність визначає не лише зручність користувача, а й економічні показники компаній, оскільки персоналізовані рекомендації безпосередньо впливають на залучення, утримання та активність аудиторії.

Компанія Netflix була однією з перших, хто перетворив персоналізовані рекомендації на центральний елемент свого продукту. Її система побудована на глибокому аналізі даних про перегляди, рейтинги, тривалість сесій та поведінкові патерни користувачів. Кожна взаємодія – перегляд серії, зупинка відео, вибір мови чи навіть швидкість прокручування – стає сигналом для алгоритмів, які моделюють індивідуальні вподобання глядача.

За даними Netflix Research, система складається з кількох послідовних модулів: формування базового списку релевантних фільмів, їх ранжування за прогнозованим рівнем зацікавленості користувача та динамічного оновлення результатів з урахуванням нових дій. Усе це відбувається в масштабах понад 250 млн користувачів, а рекомендаційна система генерує до 80 % переглядів контенту на платформі. Ключовою рисою Netflix є здатність адаптуватися до короткострокових змін поведінки користувачів: якщо глядач протягом кількох днів переглядає документальні фільми, система тимчасово змінює баланс категорій, щоб запропонувати більше подібного контенту.

Платформа YouTube використовує рекомендаційну систему, описану в роботі Covington, Adams і Sargin, яка обробляє сотні мільярдів сигналів щодня. Її архітектура ґрунтується на двоетапному підході: спочатку виконується

генерація кандидатів, коли з усіх відео обирається відносно невелика підмножина потенційно цікавих, а потім – ранжування, у межах якого ці кандидати впорядковуються за індивідуальною релевантністю.

Особливістю системи є здатність враховувати як короткострокові інтереси (останні перегляди, поточні тренди), так і довгострокові переваги користувача. Наприклад, якщо глядач переглянув декілька музичних кліпів певного жанру, алгоритм пропонуватиме схожі відео, але водночас намагатиметься урізноманітнити вибір. YouTube активно застосовує модулі контролю різноманітності контенту, а також системи обмеження рекомендацій для відео з потенційно спірним змістом, що демонструє поєднання технічної та етичної відповідальності у розробці рекомендаційних механізмів.

Інтернет-платформа Amazon стала піонером у використанні рекомендацій у сфері електронної комерції. Ще на початку 2000-х років компанія впровадила власну систему колаборативної фільтрації, яка порівнює не користувачів між собою, а товари між собою. Такий підхід виявився значно ефективнішим у масштабних умовах, адже дозволяє обробляти інформацію незалежно від кількості користувачів і динамічно оновлювати результати.

Рекомендаційна система Amazon аналізує історію покупок, переглядів, додавання товарів до кошика та навіть поведінку користувача під час пошуку. На основі цих даних формується мережа взаємозв'язків між товарами, що дозволяє прогнозувати, які з них найімовірніше зацікавлять покупця. За оцінками самої компанії, понад 30 % продажів генерується саме завдяки блоку «Рекомендовані для вас». Така ефективність зумовила поширення подібних моделей у більшості сучасних маркетплейсів [12].

Соціальна мережа LinkedIn, орієнтована на професійні зв'язки, використовує рекомендаційні системи у кількох напрямках: добір вакансій, пропозиції контактів, курсів і тематичного контенту. Система працює у режимі майже реального часу: будь-яка дія користувача – перегляд профілю, подання заявки, зміна статусу чи геолокації – миттєво враховується у формуванні нових рекомендацій.

Платформа має власну інфраструктуру для обробки ознак та швидкого оновлення моделей. Як зазначено в LinkedIn Engineering, система постійно навчається на потокових даних і здатна реагувати на зміни в поведінці користувачів без повного перенавчання. Це забезпечує баланс між точністю рекомендацій і швидкодією сервісу. Крім того, LinkedIn активно впроваджує політики справедливості та прозорості, щоб запобігти надмірному домінуванню великих компаній у видачі вакансій .

Прикладом соціально-спрямованої реалізації рекомендаційних систем є платформа VolunteerMatch, що поєднує волонтерів з неприбутковими організаціями. Її основна мета полягає не лише у підборі релевантних проєктів, а й у забезпеченні справедливого розподілу можливостей між різними програмами. Розробники системи відзначають, що класичні моделі ранжування часто призводять до ефекту «популярності», коли найвідоміші ініціативи отримують надмірну увагу, тоді як інші залишаються непоміченими.

Щоб уникнути цього, у VolunteerMatch застосовується механізм повторного ранжування, який оптимізує баланс між релевантністю та видимістю різних пропозицій. Це дозволяє підвищити різноманітність рекомендацій, не знижуючи їх точності. Такий підхід особливо цінний для волонтерських організацій, де поряд із технологічними факторами важливу роль відіграють етичні та соціальні аспекти [13].

Попри різницю в цілях і масштабах, усі розглянуті системи мають спільні риси: вони базуються на безперервному зборі даних, підтримують оновлення у реальному часі та забезпечують індивідуалізований досвід для кожного користувача. У комерційних системах, таких як Amazon або Netflix, основний акцент робиться на підвищенні залучення й продажів, тоді як соціальні платформи на кшталт VolunteerMatch або LinkedIn доповнюють технічні рішення етичними принципами – прозорістю, різноманітністю та рівністю доступу.

Ці приклади демонструють, що рекомендаційні системи можуть ефективно застосовуватись не лише у бізнесі, а й у соціально значущих галузях, зокрема у

сфері волонтерських стажувань.

1.3 Огляд існуючих інформаційних систем–аналогів для підбору волонтерських стажувань. Аналіз переваг та недоліків

1.3.1 Особливості ринку міжнародних молодіжних програм

Сегмент інформаційних систем для організації міжнародних волонтерських стажувань для молоді характеризується наявністю низки платформ, які виконують функції посередника між молодими людьми та організаціями–партнерами в різних країнах. До таких систем належать, зокрема, портали програм Європейського Союзу, міжнародні студентські асоціації та приватні організації, що спеціалізуються на волонтерських і стажувальних поїздках за кордон.

Спільними рисами для цих систем є: орієнтація на молодь, наявність централізованої онлайн-платформи з базою можливостей, механізми реєстрації та створення користувачького профілю, система подання заявок, а також різний ступінь автоматизації підбору програм. Деякі платформи фокусуються переважно на пошуку волонтерських можливостей, інші поєднують волонтерство з професійними стажуваннями.

Для детального аналізу, варто зосередитись на тих інформаційних системах, чия модель найближча до AIESEC. Це повинні бути міжнародні платформи з чітко структурованою базою програм, наявністю партнерських організацій у різних країнах і наданням молоді можливості поєднати розвиток компетенцій, міжкультурний обмін та соціальний вплив.

1.3.2 Огляд ключових систем-аналогів, подібних до AIESEC

У межах аналізу інформаційних систем, що функціонують на ринку міжнародних волонтерських та стажувальних можливостей, важливо оцінити не просто їх організаційну модель або масштаби діяльності, а те, наскільки кожна з них інтегрує принципи персоналізації у власну цифрову інфраструктуру. Саме наявність або відсутність рекомендаційних функцій визначає, наскільки така система відповідає сучасним тенденціям розвитку інформаційних сервісів та наскільки вона може слугувати прикладом для побудови вдосконаленої моделі в межах AIESEC.

Під час дослідження виявлено, що більшість платформ-аналогів перебувають на етапі переходу від класичних каталогів, де користувач самостійно фільтрує пропозиції, до більш інтерактивних моделей, які повинні були б аналізувати поведінку користувача та адаптуватися до неї. Однак повноцінних рекомендаційних систем, здатних автоматично прогнозувати ймовірний інтерес волонтера на основі комплексного профілю, серед конкурентів практично немає. Це створює унікальну можливість для AIESEC зайняти позицію технологічного лідера.

Європейський корпус солідарності (European Solidarity Corps) є найбільшою європейською організацією для волонтерських можливостей, проте його інформаційна система побудована переважно як структурований каталог проєктів. Користувач отримує доступ до великої кількості можливостей, однак добір цих можливостей здійснюється майже виключно за допомогою популярних фільтрів:

- країна виконання;
- тривалість участі;
- тематика діяльності;
- вікові обмеження.

З погляду рекомендаційної системи ESC має низку суттєвих обмежень.

По–перше, платформа не буде повноцінного цифрового профілю волонтера. Дані, які зберігаються у профілі, носять адміністративний характер і не використовуються для прогнозування інтересів користувача. По-друге, система не відстежує поведінкові патерни. Через це ESC не може адаптувати пошук під конкретні переваги користувача.

Таким чином, ESC є якісним прикладом інформаційної системи з потужною базою можливостей, але без механізмів персоналізованого добору. З позиції побудови рекомендаційної підсистеми ESC скоріше демонструє те, чого слід уникати, а саме, жодного аналізу профілю, жодної адаптації рекомендацій, лише традиційні фільтри [14].

Платформа IAESTE (International Association for the Exchange of Students for Technical Experience) від самого початку орієнтована на студентів технічних спеціальностей, тому містить досить деталізовані описи програм, вимоги до кваліфікації та специфічні критерії добору. Проте це не означає, що система має розвинений рекомендаційний механізм. Навпаки, логіка підбору суттєво спрощена – користувач вводить фільтри після чого переглядає доступні варіанти у вигляді списку.

На відміну від ESC, профіль користувача тут дещо змістовне, адже кандидат у стажування зобов'язаний описати свій академічний рівень, технічні навички, володіння мовами. Теоретично ці дані могли б слугувати основою для персоналізованих рекомендацій. Однак на практиці IAESTE використовує профіль лише для формальної відповідності вимогам програми, а не для оцінювання схожості між користувачем і потенційними можливостями.

Важливою характеристикою IAESTE є те, що система передбачає активну участь локальних комітетів у відборі. Саме вони вирішують, кому рекомендувати ту чи іншу позицію, а не автоматизована цифрова система. Це означає, що рекомендаційні функції тут делеговано людським експертам, а не алгоритмам. З одного боку, це забезпечує якість оцінювання; з іншого – робить платформу неспроможною до масштабованої персоналізації [15].

Платформа IVHQ (International Volunteer HQ) демонструє певні ознаки

руху в напрямі персоналізації. На сайті можна помітити модулі «рекомендовані програми» чи «проекти, які могли би Вам сподобатись». Проте характер цих рекомендацій є поверхневим. Система пропонує варіанти, що відповідають категоріям, які користувач переглядав раніше.

Головне обмеження полягає в тому, що IVHQ не використовує багатовимірний профіль користувача. У системі відсутні такі дані, як:

- рівень компетенцій волонтера;
- попередній досвід участі;
- часові обмеження;
- мотивації або сильні сторони;
- поведінкові патерни (глибина перегляду, повторні відвідування тощо).

Через це IVHQ не може формувати персоналізований рейтинг програм у термінах рекомендаційних систем. Вона лише “підтягує” подібні варіанти на основі ключових слів. Це ближче до пошукового автодоповнення, ніж до алгоритмічного аналізу вподобань [16].

Projects Abroad має дещо іншу логіку роботи. Хоча на платформі існує цифровий каталог із фільтрами, фактична рекомендаційна функція делегується консультантам, які допомагають людині обрати програму на основі телефонної або електронної комунікації. Це означає, що система використовує «людську персоналізацію», а не алгоритмічну.

Такий підхід дозволяє врахувати індивідуальні потреби, проте з точки зору інформаційних технологій він є немасштабованим і не відповідає концепції рекомендаційних систем, описаних у науковій літературі. Платформа Projects Abroad майже не аналізує поведінку користувачів, не використовує машинне навчання, не моделює інтереси та не здійснює ранжування можливостей автоматично. Функціонально вона ближча до цифрового представництва туристичного агента, ніж до інтелектуальної системи добору [17].

1.3.3 Порівняльний аналіз інформаційної системи AIESEC та платформ-аналогів

Аналіз проведених у попередньому підрозділі систем показав, що більшість конкурентів AIESEC перебувають на початкових етапах використання персоналізації: вони застосовують статичні фільтри або поверхневі підбірки «подібних» програм, проте не використовують машинне навчання, поведінкові сигнали чи багатовимірні профілі користувача. У цьому контексті AIESEC має унікальну перевагу – наявність розширених профілів користувачів і повних даних про їхню взаємодію з платформою, що створює потенціал для побудови повноцінної рекомендаційної системи, хоча фактична реалізація поки що перебуває на базовому рівні.

Для наочності наведено таблицю 1.1, що узагальнює ключові параметри рекомендаційності кожної з аналізованих платформ.

Таблиця 1.1 – Порівняння інформаційної системи AIESEC з аналогами з точки зору реалізації рекомендацій

Платформа	Тип цифрового добору	Профіль користувача	Використання поведінкових даних	Рівень персоналізації	Тип рекомендацій	Технічні переваги	Технічні недоліки
AIESEC	Каталог з фільтрами; часткове ручне коригування локальними комітетами	Розширений (навички, інтереси, досвід)	Частково фіксуються, але не використовуються алгоритмічно	Середній (потенційно високий)	Відсутні повноцінні алгоритмічні рекомендації	Наявність багатовимірних даних та історії взаємодій; можливість створення навчальних вибірок	Персоналізація не реалізована; профіль не інтегрований у моделі добору

Кінець таблиці 1.1

Платформа	Тип цифрового добору	Профіль користувача	Використання поведінкових даних	Рівень персоналізації	Тип рекомендацій	Технічні переваги	Технічні недоліки
ESC	Статичний каталог; фільтрація за основними параметрами	Базовий профіль	Не використовується	Низький	Відсутні	Висока структурованість каталогу	Відсутній аналіз профілю; відсутність алгоритмів персоналізації
IAESTE	Каталог; добір виконується локальними комітетами	Профіль студента з технічними навичками	Не використовується	Низький	Відсутні	Наявність формалізованих даних про навички	Добір не автоматизований; система не виконує обчислювального аналізу відповідності
IVHQ	Каталог; поверхневі рекомендації	Профіль обмеженої деталізації	Не використовується	Середній (лише категорійний)	Схожість за країною або тематикою	Мінімальна автоматизація добору; наявність блоку «similar opportunities»	Рекомендації неглибокі; відсутність моделей машинного навчання
Projects Abroad	Каталог; ручна персоналізація консультантами	Формальний профіль	Не використовується	Низький	Людська персоналізація	Індивідуальний підбір на основі інтерв'ю	Немає алгоритмічної персоналізації; система не масштабується

Проведений аналіз не лише дав змогу виявити сильні й слабкі сторони існуючих платформ, а й чітко визначити напрям подальшої роботи – необхідність створення інтелектуального рекомендаційного модуля, який використовує дані про профіль користувача, його поведінку та характеристики програм для формування персоналізованих рекомендацій.

1.4 Опис існуючих методів побудови інформаційної системи для вирішення задачі персоналізованого підбору волонтерських проєктів для молоді

1.4.1 Контентно-орієнтовані моделі

Контентно-орієнтовані моделі – одні з найдавніших способів створювати персоналізовані поради. Головний принцип – система вивчає особливості предметів, потім зрівнює їх із даними про користувача. На платформах для волонтерства будь-який проєкт можна описати через конкретні параметри – напрям, локація, термін, тип участі, мовні очікування, ключові компетенції, група призначення тощо. Анкета учасника, що шукає активність, також формується як набір рис – схильності, минулі практики, знання мов, доступні час і бюджет.

На підставі цих описів алгоритм визначає ступінь схожості між особистими даними й окремим проєктом, зазвичай через такі показники, як кутова близькість. Замість роботи з текстом безпосередньо – дані перетворюють у числовий формат; наприклад, за допомогою TF-IDF або складніших нейромережевих методів типу Word2Vec чи BERT. Один із ключових плюсів такого способу – можливість простежити логіку системи – видно конкретні параметри, що спрацювали при ранжируванні [18].

У той самий час методи, спрямовані на вміст, мають певні межі. Через те, що система найчастіше пропонує схожі елементи до тих, які людина вподобала раніше, коливання у варіантах стають меншими, а нові напрями з'являються рідко. Також такі моделі ігнорують дані про вподобання інших користувачів, бо й не можуть опиратися на групові тенденції чи приховані шаблони поведінки.

1.4.2 Моделі колаборативної фільтрації

Моделі колаборативної фільтрації (КФ) аналізують дії багатьох людей стосовно численних елементів. Натомість – вони не враховують внутрішній зміст цих матеріалів, а спираються на те, що робили інші користувачі. Порівнюючи активність однаково орієнтованих осіб, система передбачує смаки конкретного читача. Такий підхід ґрунтується на подібностях у вподобаннях.

Взаємодії записують як матрицю «людина – елемент», на базі якої створюють різні схеми. Нинішні способи спільного фільтрування нерідко ґрунтуються на розкладанні даних – завдяки цьому кожному людину й об'єкт можна описати набором прихованих ознак. Ці системи помічають тонкі залежності, недоступні для прямої перевірки через видимі параметри. Порівнюючи такі набори за допомогою добутку, отримують міру потенційного інтересу особи до певної речі.

Колаборативна фільтрація добре працює там, де багато минулих даних. Проте вона слабка перед новими – якщо учасники не мають історії взаємодій, система просто не зможе щось порекомендувати. У волонтерських системах додатковою складністю є нерівномірність взаємодій, коли частина користувачів має дуже обмежену активність, а інша – надзвичайно високу, що впливає на стабільність навчання моделей [18].

1.4.3 Атрибутивна колаборативна фільтрація

Атрибутивна колаборативна фільтрація – це один із методів рекомендацій, де поряд з даними про взаємодії беруть до уваги й ознаки користувачів чи товарів. У той час як класична версія аналізує лише рейтинги чи кліки, новий підхід враховує контекстну інформацію, наприклад вік або жанр. Це особливо

допомагає, коли даних замало. Завдяки такому поєднанню система краще справляється з ситуаціями, коли немає достатньо активності нового користувача. Крім того, пропозиції стають більш адаптованими під окремого користувача.

Додавання ознак до рекомендаційних моделей знижує чутливість системи до розміру історії взаємодій – при цьому пропозиції можуть ґрунтуватися на особистих параметрах користувачів і специфіці товарів. У нинішніх цифрових платформах, де регулярно зареєстровуються нові люди, подібний метод сприяє кращій стабільності алгоритму.

У рамках атрибутивної колаборативної фільтрації використовують різноманітні дані – наприклад, демографічні параметри на кшталт віку чи освіти; географічне положення, досвід роботи тощо. Також враховуються професійні уміння й особисті зацікавлення людей. Завдяки таким даним система може порівнювати схожість між людьми або предметами, навіть якщо немає багато інформації про поведінку.

У порівнянні з традиційними методами колаборативної фільтрації, атрибутивний підхід пропонує кілька переваг. Він частково усуває труднощі «холодного старту», адже враховує характеристики елементів замість масового аналізу минулої поведінки. Також такий спосіб краще себе показує за умов малої густини даних – поширена проблема в великих платформах. Крім того, результати легше пояснити через видимі ознаки предметів; це робить модель зрозумілішою для користувача.

Хоча атрибутивний підхід має чимало переваг, він також стикається з окремими межами. Серед ключових труднощів – потрібно правильно і повно визначити атрибути; помилки чи застарілі дані тут можуть погіршити роботу рекомендацій.

Отже, атрибутивна колаборативна фільтрація допомагає значно покращити роботу рекомендаційної системи завдяки формуванню пропозицій не тільки по взаємодіях, але і по атрибутах інших користувачів. Це дає змогу точніше адаптувати пропозиції, менше опиратись на минулі взаємодії й легше пояснювати варіанти. Такий підхід використовують у гібридних моделях [19].

1.4.4 Гібридні моделі

Гібридні моделі для рекомендаційних систем комбінують декілька методів у одній структурі. Дослідження показують, що існує низка шляхів такої інтеграції: зважене об'єднання результатів, каскадні процедури, об'єднання на рівні ознак, мета – рекомендації тощо. Основна вигода даних моделей полягає в тому, що слабкі місця одного підходу можуть бути покритими перевагами іншого.

У практичних реалізаціях часто застосовується двошарова архітектура. Спочатку контентно-орієнтований модуль відбирає з каталогу підмножину «кандидатів», які за змістом відповідають профілю користувача, а потім колаборативний або нейронний ранжувальник упорядковує їх з урахуванням поведінкових патернів.

Це дозволяє зменшити простір пошуку, уникнути суворих обмежень холодного старту й одночасно використати колективний досвід.

У волонтерських платформах краще працюють гібридні підходи. Інформація про минулу активність людей, те, що їм подобається, і як вони діють – це покращує передбачення; аналіз матеріалів дає чітке співвіднесення за змістом і сенсом. Приклад великого досвіду в різних сферах показує: поєднані моделі все частіше стають нормою. У разі міжнародних волонтерських програм такий комбінований механізм рекомендацій слугує найпрацевдатнішим базисом для адаптації під особу [18].

1.4.5 Контекстно-залежні моделі

Контекстно-залежні системи покращують базову модель «користувач – елемент», впроваджуючи фактори оточення, пов'язані з моментом вибору. У

різних дослідженнях контекст розглядається як додатковий вимір даних: час, місце, настрій, тип пристрою, соціальне оточення тощо. Це перетворює двовимірну матрицю взаємодій на багатовимірну структуру.

У волонтерстві важливий контекст. Молодь залучається до волонтерських проєктів залежно від навчання, канікул, коштів на дорогу, віз, знання мови чи безпеки. Такі моделі працюють по-різному. Простіші підходи ґрунтуються на фільтрації об'єктів, які не відповідають жорстким контекстним обмеженням, після чого застосовується базовий алгоритм рекомендацій. Більш складні техніки інтегрують контекст безпосередньо у модель через модифіковані факторизаційні методи або нейронні архітектури. Для AIESEC значення контексту має велике значення. Пропозиції без урахування часу, віз чи грошей іноді технічно правильні – проте на практиці не працюють. Системи, що залежать від обставин, дають поради, адаптовані до справжнього стану справ [18].

1.5 Постановка задачі дослідження

У сучасних умовах цифрової трансформації молодіжних та волонтерських організацій виникає потреба в створенні інформаційних систем, здатних не лише надавати доступ до каталогу програм, а й забезпечувати персоналізований добір волонтерських можливостей для молоді. Виходячи з аналізу предметної області, можна дійти до висновку, що існуючі платформи міжнародних волонтерських організацій здебільшого обмежуються статичними каталогами, простими фільтрами та мінімальною автоматизацією добору, що ускладнює орієнтацію користувачів у великій кількості програм та знижує ефективність взаємодії з системою. Інформаційна система AIESEC містить значні обсяги даних про профілі користувачів, характеристики програм і взаємодію волонтерів з платформою, однак ці дані практично не використовуються для побудови повноцінного рекомендаційного механізму.

За таких умов постає науково-практична задача – дослідити існуючі методи побудови рекомендаційних та суміжних інформаційних систем, визначити їх придатність для вирішення проблеми персоналізованого підбору волонтерських програм, а також обґрунтувати архітектурні рішення, які мають бути покладені в основу рекомендаційного модуля для платформи AIESEC. Особливість предметної області полягає в тому, що підбір волонтерських можливостей залежить не лише від інтересів користувача, а й від низки контекстних, поведінкових, соціальних і логістичних факторів, що потребує комплексного поєднання методів аналізу даних.

Об'єктом дослідження є процес персоналізованого добору волонтерських проєктів у масштабованих інформаційних системах молодіжних організацій.

Предметом дослідження є методи та моделі побудови інформаційної системи, орієнтованої на автоматизований підбір волонтерських стажувань з урахуванням профілю, поведінки й контексту користувача.

Метою дослідження є аналіз існуючих моделей побудови рекомендаційних та суміжних інформаційних систем, визначення їх можливостей та обмежень у контексті волонтерських програм, а також обґрунтування оптимального підходу до створення рекомендаційної системи для інформаційної платформи AIESEC.

Для досягнення поставленої мети необхідно розв'язати такі основні завдання:

- здійснити огляд світових підходів до побудови рекомендаційних систем і проаналізувати їх застосовність до специфіки волонтерських програм;
- визначити ключові вимоги до інформаційної системи персоналізованого підбору проєктів для молоді та специфічні фактори, що впливають на вибір волонтера;
- дослідити контентно–орієнтовані, колаборативні, гібридні, контекстно–залежні в аспекті їх використання у платформах на кшталт AIESEC;
- встановити переваги та недоліки кожного з методів у контексті предметної області та визначити доцільність їх поєднання;
- сформулювати концептуальну постановку архітектурного рішення щодо

побудови рекомендаційної підсистеми для AIESEC, яка забезпечить персоналізований і контекстуально релевантний добір програм;

– визначити напрями подальшої розробки та майбутнього впровадження отриманих результатів у практичну частину інформаційної системи.

Таким чином, постановка задачі визначає необхідність комплексного дослідження методів побудови рекомендаційних систем з метою створення інтелектуального модуля персоналізації, здатного підвищити якість добору волонтерських програм та ефективність взаємодії користувачів з інформаційною системою AIESEC.

2 ДОСЛІДЖЕННЯ МЕТОДІВ ДЛЯ ФОРМУВАННЯ РЕКОМЕНДАЦІЙ ПРОЄКТІВ AIESEC

2.1 Основні проблеми при побудові рекомендаційних систем

Рекомендаційні системи – це інструменти, що відбирають потрібний контент для окремого користувача через надлишок даних. Якісні результати вони відображають у випадках, де існує велика кількість даних. Також активно застосовуються в медіапросторі чи проєктах громадської участі.

Один із головних недоліків таких систем – надто розріджена матриця «користувач-елемент», де реальних взаємодій набагато менше, ніж потенційних пар. Через це важко визначити схожих користувачів або точно передбачити оцінки за допомогою колаборативної фільтрації. З цим часто постає й проблема нового елемента: для свіжих профілів чи предметів просто немає минулих даних, тож класичні алгоритми не дають надійних результатів без додаткової інформації.

Другим недоліком є слабка персоналізація – платформа ігнорує чи погано використовує дані про освіту, навички, мови, переважну країну чи фінансові межі. Через це людину сприймають як безликого користувача, орієнтуючись лише на дії у системі. У результаті, пропозиції формально схожі на основі минулих кліків, проте рідко стосуються справжніх потреб або можливостей особи.

Останній фактор – поширена упередженість на користь популярних варіантів. Тобто частіше показують те, що вже має високий попит, хоча менш помітні, але більш влучні за характеристиками пропозиції просто не доходять до користувача.

Для вирішення цих проблем при побудові рекомендаційних систем, необхідно використовувати поєднання декількох методів.

2.2 Дослідження колаборативної фільтрації на основі взаємодії користувачів

Колаборативна фільтрація – один з найпоширеніших способів створення рекомендацій. Принцип роботи даного методу полягає в порівнянні користувачів, які мали однакові вподобання в минулому, і найбільш вірогідно будуть мати однакові у майбутньому. Зазвичай дані для таких систем організовують як розріджену таблицю типу «людина–елемент». У ній кожне число позначає рейтинг чи показник активності людини щодо конкретного елемента.

Колаборативна фільтрація ґрунтується на певному алгоритмі дій:

- для потрібного користувача знаходять кілька найближчих за характеристиками сусідів;
- на підставі даних про сусідів кожному потенційному об'єкту присвоюється передбачуване значення рейтингу;
- елементи з кращими передбаченими показниками пропонуються конкретному користувачеві.

Головне – це поррахувати схожість між уподобаннями людей, ґрунтуючись на їхніх оцінках. Однак іноді замість оцінок беруть дані про взаємодії. Наприклад, якщо хтось часто клікає чи переглядає контент. Такий підхід допомагає краще зрозуміти переваги.

У колаборативних методах часто використовують коефіцієнт Пірсона – це показник лінійного зв'язку між двома векторами оцінок. Коли два учасники, x та y , мають загальну групу елементів E , схожість між ними розраховують за наступною формулою:

$$s(x, y) = \frac{\sum_{e \in E^*} (R_{x,e} - \bar{R}_x)(R_{y,e} - \bar{R}_y)}{\sqrt{\sum_{e \in E^*} (R_{x,e} - \bar{R}_x)^2} \sqrt{\sum_{e \in E^*} (R_{y,e} - \bar{R}_y)^2}}, \quad (2.1)$$

де $R_{x,e}$ та $R_{y,e}$ – рейтинги (або числові значення взаємодій) користувачів x та y для об'єкта e ;

\bar{R}_x та \bar{R}_y – середні значення оцінок відповідних користувачів по всіх оцінених ними об'єктах;

E – множина об'єктів, оцінених обома користувачами.

Для використання у якості подібності в рамках рекомендаційних алгоритмів часто застосовують лінійну нормалізацію:

$$sim(x, y) = \frac{s(x,y)+1}{2}, \quad (2.2)$$

де $sim(x, y)$ – нормалізована міра подібності між користувачами x та y , що належить інтервалу $[0; 1]$;

$s(x, y)$ – значення коефіцієнта кореляції Пірсона за формулою (2.1).

На основі подібності з обраною підмножиною сусідів $N(x)$ прогнозований рейтинг елемента e для користувача x у класичній може визначатися як зважене відхилення від середнього рейтингу:

$$\hat{R}_{x,e} = \bar{R}_x + \frac{\sum_{y \in N(x)} sim(x,y) (R_{y,e} - \bar{R}_y)}{\sum_{y \in N(x)} |sim(x,y)|}, \quad (2.3)$$

де $\hat{R}_{x,e}$ – прогнозований рейтинг (або агреговане значення взаємодій) користувача x для об'єкта e ;

\bar{R}_x – середній рейтинг користувача x ;

$R_{y,e}$ – рейтинг (або значення взаємодій) користувача y для об'єкта e ;

\bar{R}_y – середній рейтинг користувача y ;

$N(x)$ – множина «сусідів» користувача x , тобто найбільш подібних користувачів;

$sim(x, y)$ – нормалізована подібність між користувачами x та y за формулою (2.2).

У якості взаємодій, планується використовувати перегляди, вподобання, зберігання та подання заявки на проєкт.

Отже, колаборативна фільтрація на основі користувачів формує рекомендації через справжню поведінку людей – завдяки цьому система помічає приховані зв'язки між діями й підказує матеріали, які не видно за словами. Не дивлячись на це, спосіб потребує багато минулих взаємодій; без них він просто не працює, особливо для нових користувачів або проєктів. Розрідженість даних та чутливість до шумових або випадкових взаємодій також обмежують точність моделі.

Підбиваючи підсумок, можна сказати, що колаборативна фільтрація є цінною та якісною для рекомендаційних систем, але в ізольованому вигляді не є активною [20].

2.3 Дослідження атрибутивної колаборативної фільтрації

Звичайна КФ працює якісно лише тоді, коли є багато взаємодій. У разі появи нового користувача, тобто «холодного старту», КФ не працює коректно. Для вирішення цієї проблеми, варто використовувати атрибутивну колаборативну фільтрацію, яка вирішує дану проблему методом порівняння користувачі на основі їхніх описових характеристик.

У випадку волонтерської платформи AIESEC використовуються наступні популярні атрибути:

- освітній фон;
- навички;
- мови;
- пріоритетні країни;
- рівень освіти;
- бюджет.

Для множинних категоріальних атрибутів, як міра подібності використовується індекс Жаккара. Міра Жаккара для двох множин A та B визначається як:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}, \quad (2.4)$$

де A, B – порівнювані множини категоріальних значень;

$|A \cap B|$ – кількість спільних елементів у множинах A та B ;

$|A \cup B|$ – кількість усіх унікальних елементів у їх об'єднанні;

$J(A, B)$ – значення індексу Жаккара, що належить інтервалу $[0; 1]$.

Загальна подібність між двома користувачами u_1 та u_2 визначається як зважена сума часткових подібностей, де ваги встановлено відповідно до значущості атрибутів:

$$\begin{aligned} sim_{attr}(u_1, u_2) &= 0.25 \cdot J(backgrounds) \\ &\quad + 0.30 \cdot J(skills) \\ &\quad + 0.20 \cdot J(languages) \\ &\quad + 0.20 \cdot J(preferred_countries) \\ &\quad + 0.025 \cdot J(study_level) \\ &\quad + 0.025 \cdot J(budget), \end{aligned} \quad (2.5)$$

де $J(\cdot)$ – індекс Жаккара для відповідного множинного атрибуту.

Таким чином, подібність $sim_{attr}(u_1, u_2)$ належить інтервалу $[0; 1]$ та відображає ступінь відповідності профільних характеристик користувачів.

Оскільки атрибути профілю користувача мають різний внесок у релевантність рекомендацій, у моделі повинні бути реалізовані вагові коефіцієнти, що задають важливість кожної ознаки. За допомогою цього, система може бути гнучкою та підлаштовуватися під специфіку предметної області [19].

2.4 Дослідження контентно-орієнтованої моделі на основі TD-IDF

Контентно-орієнтовані рекомендаційні моделі зіставляють опис об'єкта з профілем користувача, визначаючи їхню семантичну подібність. На відміну від колаборативних методів, які базуються на поведінці інших користувачів, контентний підхід використовує безпосередньо текстові характеристики об'єктів. Це робить його особливо корисним у випадках, коли інформація про взаємодії обмежена, але текстові описи проєктів є повними та структурованими.

Одним із найпоширеніших способів подання текстових документів є модель TF-IDF, що дозволяє формувати числовий вектор, у якому найбільшу вагу отримують терміни, характерні саме для певного документа, а не поширені в усій колекції.

Дані рекомендаційні моделі порівнюють опис об'єкта з профілем користувача, визначаючи семантичну подібність. На відміну від попередніх методів, які базуються на взаємодіях та поведінці користувачів, контентно-орієнтований метод використовує текстові характеристики. Це робить його також корисним, у випадку, коли відсутня або замало інформацію про взаємодії користувачів, але текстові характеристики присутні.

Одним із найпопулярніших способів є TD-IDF. Він дозволяє формувати числовий вектор з характеристик, у якому найбільшу вагу отримують ті терміни, які найхарактерніші саме для певного документа.

Нехай t – термін (слово), d – документ (опис окремого проєкту), D – колекція всіх документів. Частота терміна в документі визначається як:

$$TF(t, d) = \frac{f_{t,d}}{\sum_{t'} f_{t',d}}, \quad (2.6)$$

де $TF(t, d)$ – відносна частота терміна t у документі d ;

$f_{t,d}$ – кількість входжень терміна t у документі d ;

$\sum_{t'} f_{t',d}$ – загальна кількість термінів (слів) у документі d , тобто сума частот

усіх термінів t' .

Обернена частота документа (IDF) визначається формулою:

$$IDF(t, D) = \log \frac{|D|}{df_t}, \quad (2.7)$$

де $IDF(t, D)$ – обернена частота документа для терміна t у колекції D ;

$|D|$ – загальна кількість документів у колекції;

df_t – кількість документів, у яких зустрічається термін t .

Тоді вага терміна t у документі d задається як:

$$w(t, d) = TF(t, d) \cdot IDF(t, D), \quad (2.8)$$

де $w(t, d)$ – вага терміна t у документі d ;

$TF(t, d)$ – частота терміна t у документі d за формулою (2.6);

$IDF(t, D)$ – обернена частота документа для терміна t за формулою (2.7).

У результаті кожен документ представляється у вигляді вектора $v_d = (w_1, \dots, w_m)$ у просторі термінів розмірності m , де w_i – вага i -го терміна з словника.

Для того щоб порівняти розраховані вектори профілю користувача та проекту, слід використати косинусну подібність. Формула косинусної подібності виглядає наступним чином:

$$\cos(\theta) = \frac{v_u \cdot v_p}{\|v_u\| \cdot \|v_p\|}, \quad (2.9)$$

де v_u – TF-IDF вектор профілю користувача;

v_p – TF-IDF вектор опису проекту;

$v_u \cdot v_p$ – скалярний добуток векторів v_u та v_p ;

$\|v_u\|$, $\|v_p\|$ – норми відповідних векторів;

$\cos(\theta)$ – значення косинусної подібності.

Попри переваг даного методу, він суттєво залежить від повноти та структурованості текстових описів і має природну тенденцію рекомендувати лише ті проєкти, які схожі на попередні інтереси користувача. Він не враховує колективний досвід інших користувачів і тому може пропускати альтернативні можливості, які добре сприймаються спільнотою, але не збігаються з формальним контентом профілю.

У результаті контентні моделі є ефективним механізмом точного ранжування, проте найкраще працюють у поєднанні з іншими методами [21].

2.5 Опис створеного гібридного методу рекомендацій

Після проведення дослідження вищезазначених методів, прийнято рішення створити гібридний метод на їхній основі.

Запропонований гібридний метод поєднує колаборативну фільтрацію на основі користувачів або атрибутів поєднану з контентно-орієнтованим ранжуванням, утворюючи двоетапний процес формування рекомендацій.

Завдяки даній архітектурі відбирається найбільш підходящий список проєктів, а потім впорядковується відповідно до профілю користувача.

Першим етапом роботи гібридного методу є визначення стану користувача.

Якщо користувач не має історії взаємодії з проєктами, це означає, що він новий і система буде обчислювати рекомендації за атрибутною колаборативною фільтрацією, використовуючи міру Жаккара та зважену комбінацію часткових подібностей. Для цього використовуються формули (2.4 – 2.5).

Для активних користувачів використовується колаборативна фільтрація на основі користувачів, що базується на коефіцієнт Пірсона та його нормалізація. Для цього використовується формула (2.1 – 2.2).

Після знаходження найбільш подібних користувачів $N(u)$ система аналізує

всі проекти, з якими взаємодіяли ці користувачі. У колаборативній фільтрації передбачуваний рейтинг формується за формулою (2.3), де внесок кожного сусіда зважується пропорційно його подібності до цільового користувача.

Кожна взаємодія користувача з проектом стає числовою оцінкою – так званою «силою зацікавленості». Чим вищою є подібність між цільовим користувачем і сусідом, тим більше впливає його взаємодія на формування початкової оцінки проекту. Отже, система швидко складає перші рейтинги, спираючись на минулі дані людей із списку $N(u)$.

У результаті отримуємо список проектів на основі взаємодій або атрибутів інших користувачів.

Після формування списку проектів, кожен із них порівнюється із профілем користувача за допомогою контентно–орієнтованого методу на основі TD-IDF. Для цього використовуються формули (2.6 – 2.8). Для визначення ступеня семантичної близькості застосовується косинусна подібність (формула (2.9)). Вона відображає схожість напрямків проекту та користувача у багатовимірному просторі термінів. Отримане значення є фінальною оцінкою проекту-кандидата.

У результаті роботи контентно–орієнтованої моделі, користувач отримує кінцевий список найбільш підходящих проектів.

Гібридний спосіб об'єднує переваги трьох підходів. Через це система краще пристосовується до змін. Перш за все, такий метод допомагає подолати «проблему холодного старту».

Друга суттєва перевага даного методу – покращене ранжування. Хоча колаборативна фільтрація чітко вказує на ймовірно підходящі користувачу проекти, вона часто неточна на окремих прикладах. У той час як контентно–орієнтований аналіз добре розпізнає змістові ознаки елементів, однак обмежує різноманітність пропозицій. Об'єднання обох підходів формує стабільніший механізм, де один метод перекриває недоліки іншого.

Діаграма, яка відображає структуру новоствореного методу зображена на рисунку 2.1.

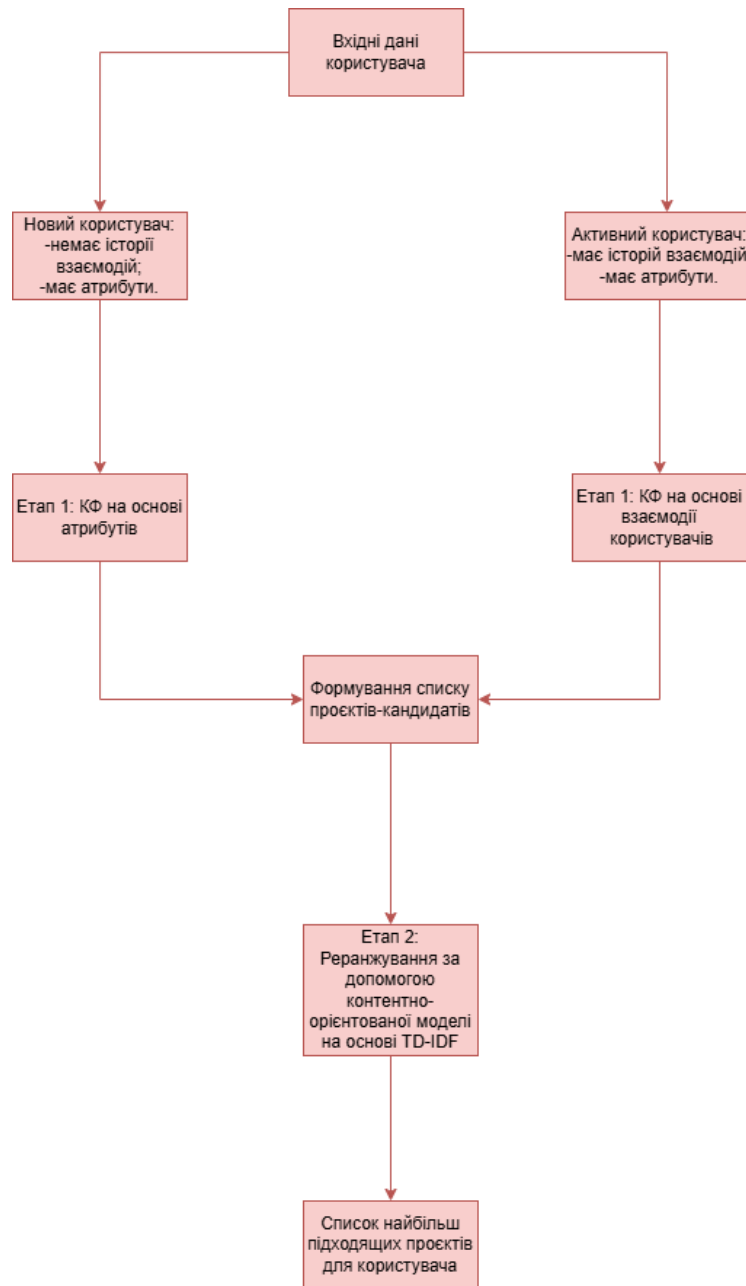


Рисунок 2.1 – Діаграма структури гібридного методу

3 ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ПОБУДОВИ РЕКОМЕНДАЦІЇ В ІНФОРМАЦІЙНІЙ СИСТЕМІ ВОЛОНТЕРСЬКИХ СТАЖУВАНЬ AIESEC

3.1 Опис контекстної діаграми

Побудовано технологію рекомендацій, яка реалізована у вигляді гібридного методу, який поєднує КФ на основі взаємодій або атрибутів та контентне-орієнтоване ранжування із використанням TF-IDF. У даному розділі представлено інформаційні моделі процесу у вигляді діаграм IDEF0, що дозволяють формально описати структуру, вхідні та вихідні дані, засоби та керуючі впливи.

Діаграма A-0 показує систему в найзагальнішому вигляді. Основним елементом є один процес – «Побудова рекомендацій проєктів у інформаційній системі AIESEC». Вхідними даними слугують дані про користувача або історія його активності, які служать основою для подальшого аналізу. Основний напрям роботи системи задають три методи: КФ на основі взаємодії користувачів або атрибутів та контентно-орієнтований з використанням TD-IDF.

Для підтримки роботи системи використовуються системи управління базами даних – там містяться усі дані про користувачів, їхні взаємодії та проєкти. Окрім цього, задіяно мову Python для побудови алгоритмів порівняння й сортування.

У результаті, користувач отримує список рекомендацій, упорядкованих за ступенем підходящості. Контекстна діаграма відображає загальну структуру технології разом із зовнішніми взаємодіями, але не розглядає внутрішню логіку детально.

Контекстна діаграма зображена на рисунку 3.1

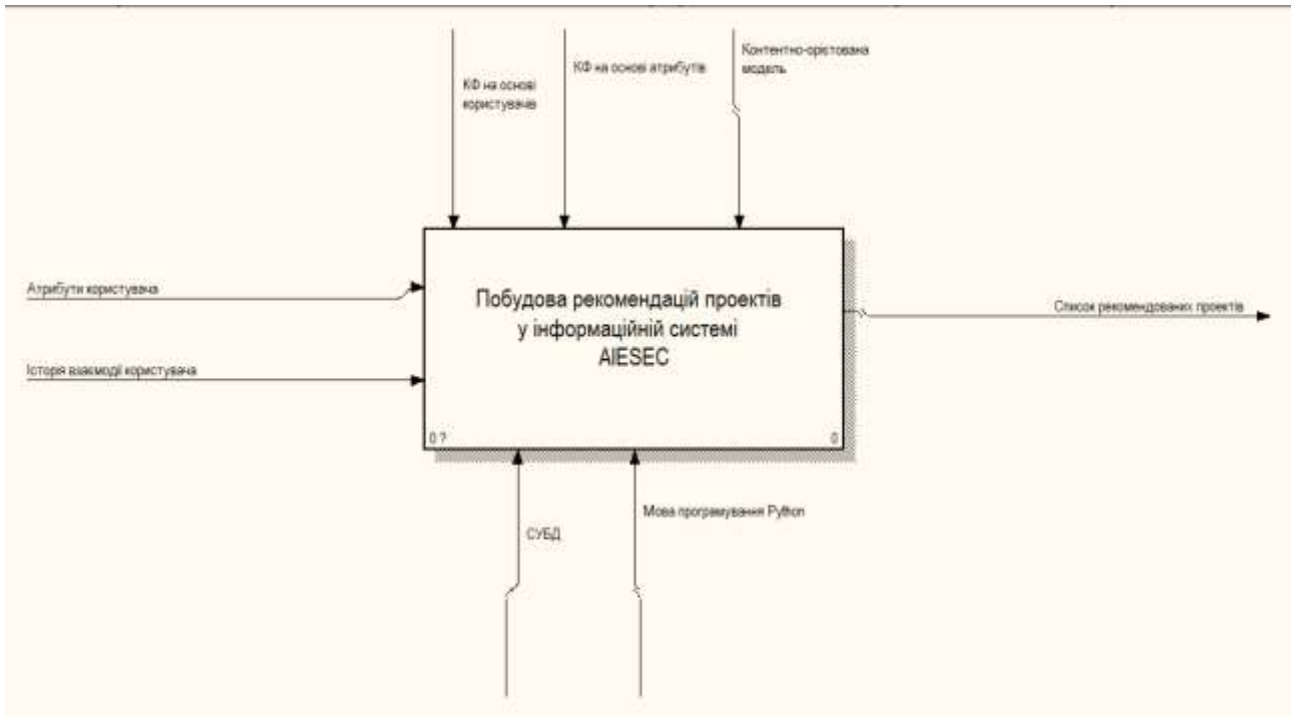


Рисунок 3.1 – Контекстна діаграма

3.2 Опис діаграми декомпозиції

На діаграмі декомпозиції показано, як працює система побудови рекомендацій – вона містить чотири логічно пов’язані процеси. У випадку, якщо користувач новий у системі, все починається із реєстрації разом із заповненням особистої інформації; при цьому дані перетворюються на організований профіль. Такий профіль потрібний для майбутнього аналізу, а також записується до бази даних. Особиста інформація користувача використовується для побудову початкового списку рекомендації за допомогою КФ на основі атрибутів. Для цього використовуються формули (2.4 – 2.5).

У випадку, якщо користувач зареєстрований у системі та має історію взаємодії, першим кроком є знаходження схожих користувачів через їхню активність та створення списку проєктів. Даний процес починається лише при наявності даних про минулі взаємодії із проєктами у системі. Тут використовують КФ, ґрунтуючись на поведінці користувача. Для цього

4 ПРАКТИЧНА РЕАЛІЗАЦІЯ ГІБРИДНОГО МЕТОДУ

4.1 Програмна реалізація методів рекомендаційної системи

4.1.1 Обґрунтування вибору технологічних засобів

Гібридна рекомендаційна система побудовано за допомогою мови програмування Python.

Python – це універсальна мова для програмування, де код виходить зрозумілим. Завдяки стислості синтаксису, розв’язок через Python часто значно компактніший, ніж версія на інших мовах.

Цей вибір обумовлений тим, що Python розвинутий у сфері науки про дані та машинного навчання, має у списку потужні бібліотеки для матричних обчислень та обробки природної мови [22].

Щоб посприяти швидкодії системи доцільним було використання спеціалізованих бібліотек для векторних операцій, що дозволяє уникнути повільних обчислень. Використані наступні бібліотеки:

- `numpy` – використовується як база для роботи з багатовимірними масивами. Відповідає за зберігання розріджених матриць взаємодій «користувач-проект» та виконання векторизованих арифметичних операцій. Це забезпечує ефективне використання пам'яті та високу швидкість обчислень при масштабуванні кількості користувачів [23];

- `scipy` – розширює можливості `numpy`, надаючи інструменти для наукових обчислень. Модуль `scipy.stats` використовується для точного розрахунку статистичних метрик, зокрема коефіцієнта кореляції Пірсона, який є основою алгоритму КФ на основі взаємодії користувачів [24].

Для реалізації контентно–орієнтованої моделі використовується бібліотека `scikit.learn`, яка є стандартом для машинного навчання [25].

4.1.2 Реалізація контентної фільтрації

У фрагменті коду A.1 з додатку A відображено роботу методу `build_project_text` із класу `ContentBasedRecommender` – він відтворює єдиний текстовий образ проєкту для аналізу через TF-IDF. Цей крок об'єднує важливі поля проєкту в один семантичний блок, забезпечуючи основу для порівнянь. Така трансформація переводить структуровані дані у простий текст, придатний до векторизації.

Спочатку алгоритм аналізує наявність ключові описових полів – наприклад, назву проєкту, посаду, опис діяльності. Потім ці елементи потрапляють до загального списку фрагментів. Наступним кроком обробляються додаткові параметри – `background`, `skills`, `languages`. Перед тим як їх приєднати, запускається функція `_normalize_attr`. Вона приводить дані до стандартного виду й ліквідує розбіжності у форматуванні.

Далі потрібно додати географічну інформацію – країну, місто і вид програми.

Після об'єднання всіх частин у цілісний текст метод виконує низку процедур з очищення даних:

- переведення всіх символів у нижній регістр для усунення чутливості до регістру;
- видалення розділових знаків і сторонніх символів через регулярні вирази;
- нормалізацію пробілів для отримання структурованого й однорідного тексту.

У результаті отримуємо текст, придатний для наступної обробки – він чітко й лаконічно передає суть проєкту. Завдяки такому підходу TF-IDF перетворює дані у числові вектори з високим рівнем інформативності, що покращує точність визначення подібності між проєктами та сприяє підвищенню якості контентно-орієнтованих рекомендацій.

4.1.3 Реалізація колаборативної фільтрації. Вирішення проблеми «холодного старту»

У системі реалізовано два підходи до підбору подібних користувачів, що дозволяє адаптувати процес побудови рекомендацій до різного рівня доступних даних. Якщо користувач має історію взаємодій із проєктами платформи, пріоритет надається класичним методам колаборативної фільтрації. Для нових користувачів, які не взаємодіяли з проєктами, такий варіант є неможливим через повну відсутність поведінкової інформації. В цих умовах система переходить до атрибутивного аналізу профілю, що дозволяє оцінити подібність між користувачами на основі їхніх характеристик.

У фрагменті блоку А.2 з додатку А, метод `calculate_user_attribute_similarity` виконує обчислення подібності між двома профілями користувачів за допомогою сукупності окремих атрибутів. Основою для порівняння множин атрибутів є коефіцієнт Жаккара, який дає змогу оцінити ступінь перетину між двома множинами. Такий підхід забезпечує коректність порівняння навіть у випадках неоднакового обсягу вихідних даних.

Спочатку метод здійснює нормалізацію полів профілю, перетворюючи атрибути на множини уніфікованих значень. Для кожної з категорій обчислюється окремий показник подібності. Кожний із цих показників робить внесок у підсумкову оцінку пропорційно своїй важливості, яка визначена ваговими коефіцієнтами. Ваги підібрані таким чином, щоб пріоритетними були найбільш значущі атрибути.

Після отримання всіх коефіцієнтів, метод обчислює загальне значення подібності як зважену суму, що враховує значущість кожного атрибуту для рекомендаційної моделі. Підсумкове значення є числом у діапазоні $[0, 1]$, де більші значення вказують на вищу подібність між користувачами. На основі цього показника система визначає подібних користувачів для нового користувача, формуючи початковий набір релевантних проєктів для подальшого

ранжування.

Таким чином, використання даного методу, також вирішує проблему «холодного старту».

Для користувачів, які навпаки, мають історію взаємодій, система застосовує КФ на основі взаємодії користувачів. У цьому випадку використовується метод `find_similar_users_by_interactions`, який базується на обчисленні схожості між векторами активності двох користувачів. Лістинг даного методу наведено у фрагменті блоку А.3 з додатку А. На відміну від КФ на основі атрибутів, цей метод порівнює реальні взаємодії користувачів, що дозволяє точніше визначити схожість уподобань.

Метод отримує номер користувача та обирає рядок із матриці взаємодій, де кожен елемент відображає числову величину активності щодо конкретного проекту. Далі для кожного іншого користувача система виконує перевірку на схожість по точкам взаємодії. Це важливий етап, оскільки обчислення кореляції має сенс лише тоді, коли в обох користувачів є хоча б кілька проектів, з якими вони взаємодіяли.

Для побудови векторів порівняння відбираються індекси проектів, де хоча б один з користувачів має ненульову взаємодію. Це забезпечує те, що аналіз охоплює всі релевантні дані та не спотворюється нульовими значеннями, які не несуть інформації про інтереси.

Основною метрикою подібності є кореляція Пірсона, що оцінює лінійний зв'язок між двома векторами активностей. Метод намагається обчислити коефіцієнт кореляції для пари користувачів. У разі виникнення помилок або появи некоректних значень система замінює результат нульовим значенням, що запобігає спотворенню подальших обчислень.

Оскільки значення кореляції Пірсона може набувати діапазону від -1 до 1 , метод виконує нормалізацію до інтервалу $[0; 1]$ шляхом лінійного перетворення. Це дозволяє інтерпретувати результат як міру подібності, де 0 означає повну відсутність спільних інтересів, а 1 – максимально можливу схожість у поведінці.

На фінальному етапі відбувається сортування знайдених користувачів у

порядку спадання показника подібності, після чого повертається список із найбільш релевантними користувачами. На основі взаємодії користувачів формується початковий список проєктів, який буде використано на наступних етапах ранжування рекомендацій.

4.1.4 Реалізація гібридного методу

Метод `generate_candidates` у класі `HybridRecommender`, який наведено у фрагменті блоку А.4 з додатку А, виконує етап гібридної рекомендаційної моделі – формування початкового списку проєктів–кандидатів на основі даних колаборативної фільтрації. Метод працює з користувачами, для яких система вже визначила список подібних користувачів разом зі значеннями їхньої схожості.

На першому кроці здійснюється перевірка коректності даних – наявність користувача у списку, валідність індексу та доступність рядка у матриці взаємодій. У разі некоректних даних метод завершує роботу, що забезпечує стабільність системи.

Далі формується масив `project_scores`, який акумулює внесок кожного схожого користувача. Для кожного сусіда беруться його взаємодії з проєктами, нормалізуються та зважуються на квадрат коефіцієнта схожості. Зважування дозволяє значно посилити вплив найбільш подібних користувачів і знизити вагу тих, чия схожість є незначною.

Після сумування внесків усіх сусідів проєкти, з якими цільовий користувач уже взаємодіяв, виключаються зі списку шляхом обнулення їхніх балів. Це гарантує, що система рекомендуватиме для нього нові можливості.

На наступному етапі обираються проєкти з найвищими балами. Для кожного з них здійснюється перевірка валідності та оцінюється відповідність бюджетним обмеженням користувача. Якщо бюджет не збігається, бал проєкту зменшується, але проєкт не вилучається повністю, що дозволяє системі

залишатися гнучкою.

Результатом методу є список пар (проект, оцінка), який передається на наступний етап – ранжування у гібридній моделі.

Метод `recommend`, який наведено у фрагменті блоку A.5 з додатку A – є ядром гібридної рекомендаційної системи, оскільки саме він об'єднує різні підходи до формування персоналізованих рекомендацій у єдиний алгоритм. Його робота охоплює повний цикл – від визначення типу користувача до отримання фінального списку релевантних проектів, що робить цей метод ключовим у реалізації гібридної моделі.

На початку виконується перевірка доступності основних вхідних даних – профілю користувача та списку проектів. За їх відсутності формування рекомендацій неможливе, і метод завершує роботу. Далі система визначає, чи належить користувач до категорії нових чи вже має історію взаємодій. Цей етап є концептуально важливим, адже саме від нього залежить вибір стратегії пошуку схожих користувачів.

Для нових користувачів застосовується КФ на основі атрибутів, який порівнює профілі за характеристиками – навичками, фоновною підготовкою, мовами та іншими параметрами. Якщо ж користувач має достатню історію активності, система використовує КФ на основі взаємодії користувачів, що спирається на аналіз кореляції між векторами взаємодій. Такий підхід до побудування системи дозволяє налаштувати алгоритм до різних сценаріїв використання платформи та забезпечує стабільну роботу незалежно від кількості доступних даних.

На наступному етапі формується список проектів–кандидатів. Якщо сусіди знайдені, система генерує проекти на основі їхньої взаємодії, використовуючи метод `_generate_candidates`. Паралельно виконується контентне ранжування, і отримані результати об'єднуються, уникаючи дублювання. Такий підхід дозволяє забезпечити різноманітність проектів.

Фінальним етапом є переранжування кандидатів, що враховує змістовні атрибути проекту, відповідність бюджету, інтереси користувача. Саме на цьому

етапі формується остаточний рейтинг проектів, який і подається користувачу.

Метод recommend здійснює об'єднання поведінкових, атрибутивних та контентних підходів, забезпечуючи високу гнучкість і точність гібридної моделі рекомендацій.

4.1.5 Демонстрація рекомендаційної системи із застосуванням реалізованого гібридного методу

Знайомство користувача із створеною системою розпочинається з головної сторінки, на якій він може перейти до авторизації або реєстрації. Головна сторінка системи зображена на рисунку 4.1.

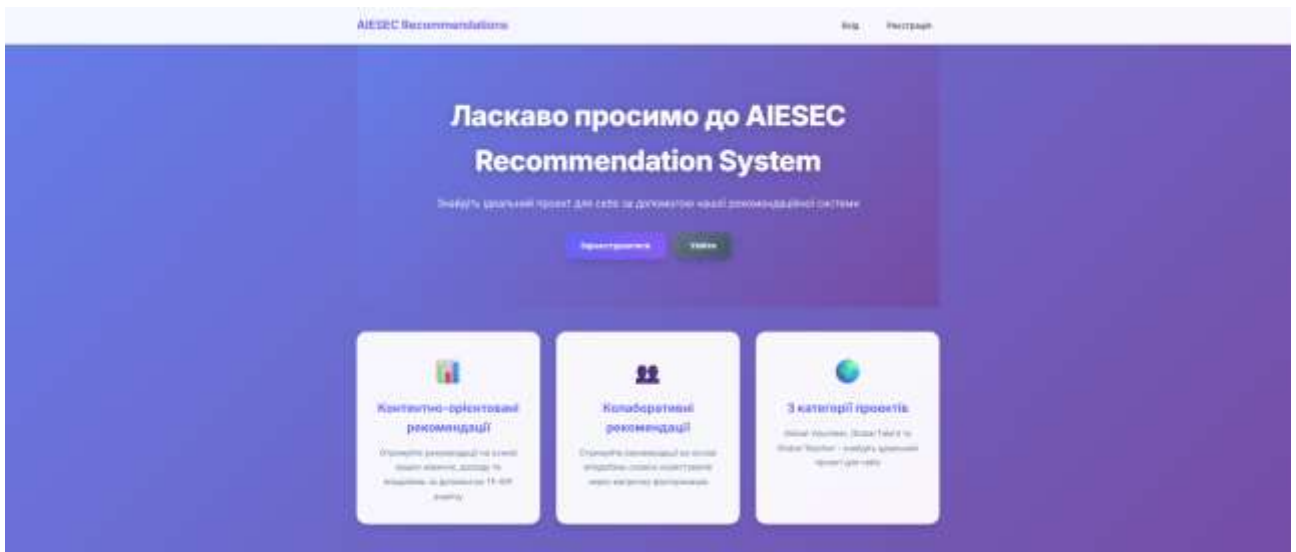


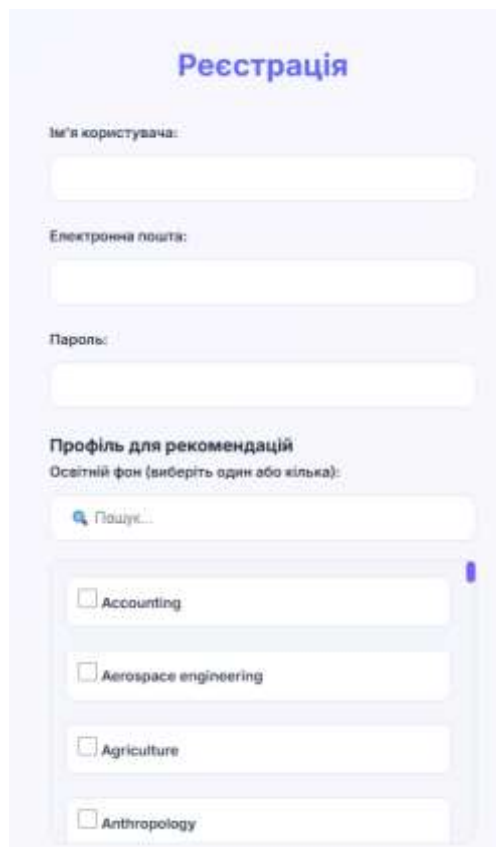
Рисунок 4.1 – Головна сторінка рекомендаційної системи

Розпочнемо з демонстрації процесу реєстрації нового користувача у системі. Даний процес проводиться заповненням користувачем особистої інформації:

- ім'я користувача;
- електронна пошта;

- пароль;
- освітній фон;
- навички;
- мови;
- країни, які цікавлять;
- рівень навчання;
- мінімальна та максимальна кількість тижнів стажування;
- очікуваний бюджет.

Сторінка реєстрації користувача зображена на рисунку 4.2.



The image shows a registration form titled "Регістрація" (Registration). It contains the following elements:

- A header "Регістрація" in blue text.
- A label "Ім'я користувача:" followed by a text input field.
- A label "Електронна пошта:" followed by a text input field.
- A label "Пароль:" followed by a text input field.
- A section titled "Профіль для рекомендацій" (Profile for recommendations).
- A sub-label "Освітній фон (виберіть один або кілька):" (Educational background (select one or more)).
- A search input field with a magnifying glass icon and the text "Пошук...".
- A list of four educational fields, each with a checkbox:
 - Accounting
 - Aerospace engineering
 - Agriculture
 - Anthropology

Рисунок 4.2 – Сторінка реєстрації користувача

Навички (виберіть одну або кілька):

Пошук...

3D Max

.NET

(Visual) FoxPro

АВАР

Мови (виберіть одну або кілька):

Пошук...

Albanian

Arabic

Armenian

Azerbaijani

Рисунок 4.2, аркуш 2

Країни, які вас цікавлять:

Пошук...

Brazil

Colombia

Mexico

Argentina

Типи програм:

Global Volunteer

Global Talent

Global Teacher

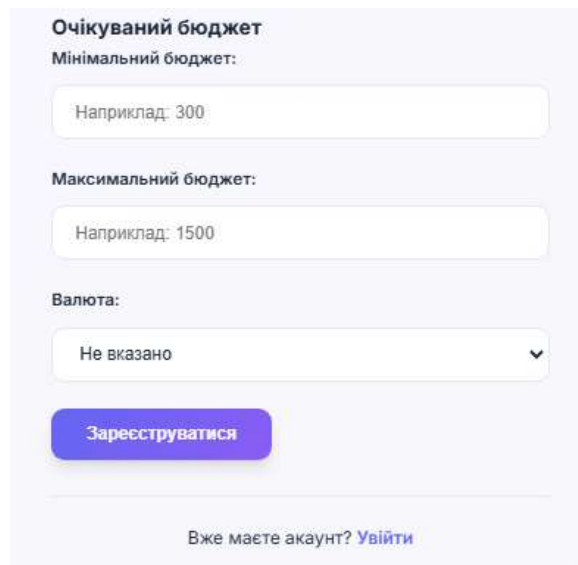
Рівень навчання:

Не вказано

Мінімальна тривалість (тижні):

Максимальна тривалість (тижні):

Рисунок 4.2, аркуш 3



Очікуваний бюджет
Мінімальний бюджет:
Наприклад: 300
Максимальний бюджет:
Наприклад: 1500
Валюта:
Не вказано
Зареєструватися
Вже маєте акаунт? [Увійти](#)

Рисунок 4.2, аркуш 4

Наступним етапом, необхідно продемонструвати роботу рекомендації проектів для користувача. В першу чергу, перевіримо, як працює рекомендація для нового користувача. Для цього підготовлено профіль новоствореного користувача. Переглянемо його профіль та кількість зафіксованих взаємодій із проектами.

Відображення кількості зафіксованих взаємодій нового користувача у системі зображено на рисунку 4.3.

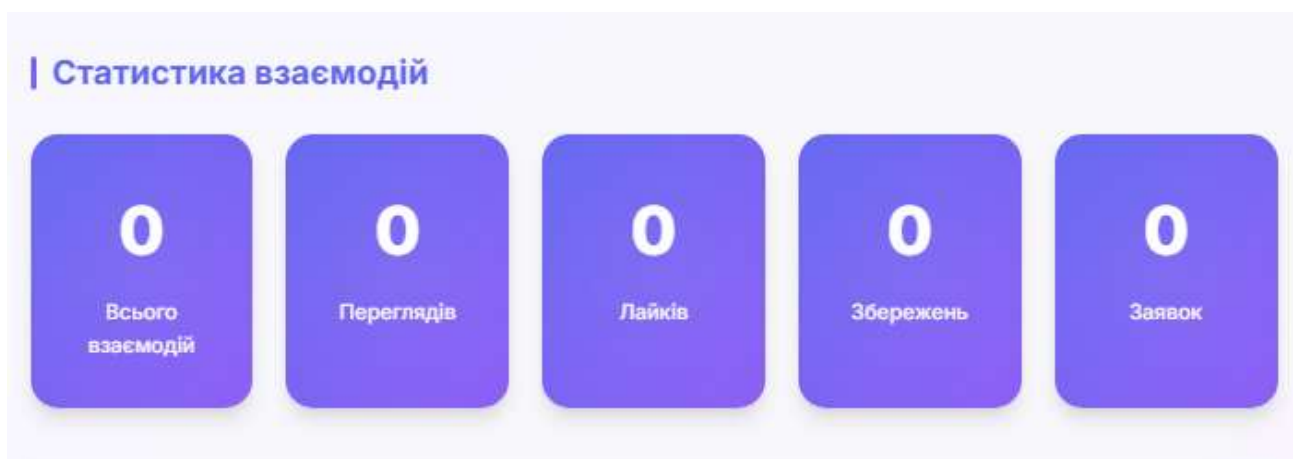


Рисунок 4.3 – Статистка взаємодій нового користувача

Як можна побачити, користувач не мав жодної взаємодії із проектами, а це

означає, що при формування рекомендацій будуть використовуватись КФ на основі атрибутів та контентно-орієнтована модель із використанням TD-IDF. Переглянемо, чи працюють методи формування рекомендації та чи співпадають вони із профілем користувача. Перейдемо на сторінку рекомендації.

Сторінка профілю користувача та сформовані рекомендації зображені на рисунках 4.4-4.5.

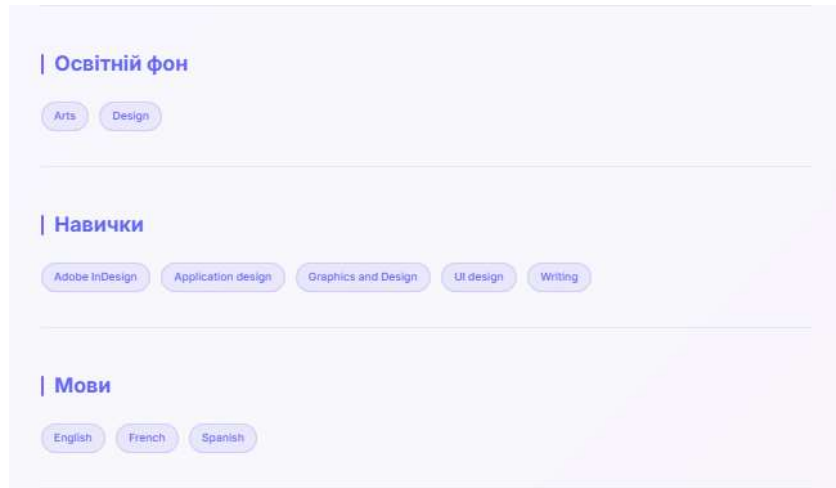


Рисунок 4.4 – Сторінка профілю користувача

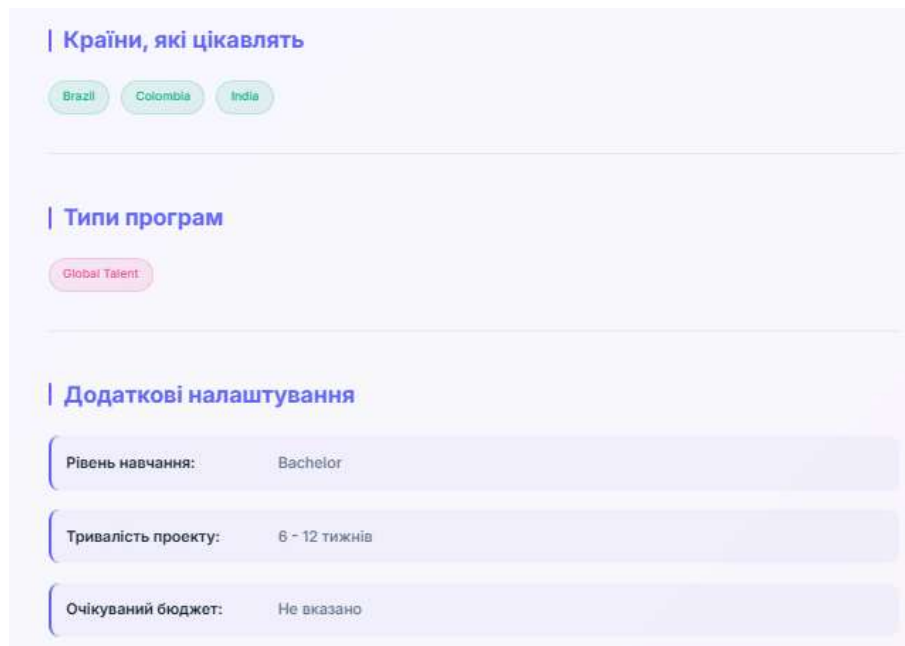


Рисунок 4.4, аркуш 2

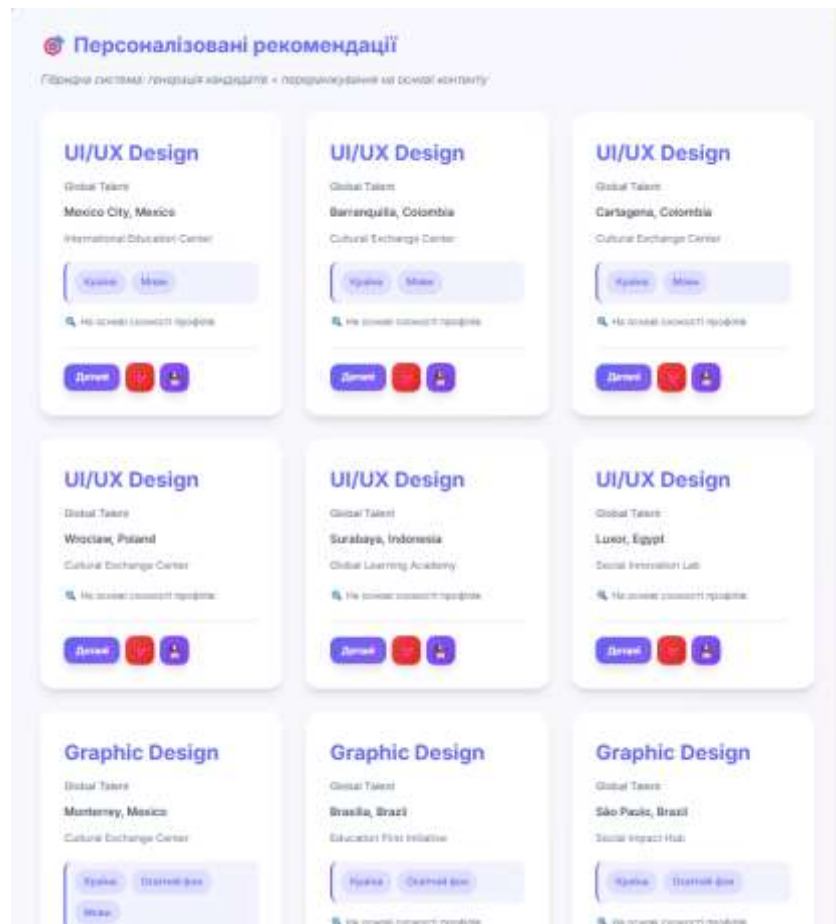


Рисунок 4.5 – Сторінка сформованих рекомендацій

Наведемо повну інформацію однієї з рекомендації. Повна інформація сформованої рекомендації наведено на рисунку 4.6.

– Назад до проєктів

Graphic Design

Global Talent Monterrey, Mexico 100%

Про організацію

Cultural Exchange Center

Опис ролі

In this Graphic Design position, you will gain hands-on experience in a professional environment. Your responsibilities include supporting daily operations, contributing to team projects, and developing practical skills in your field. You'll work alongside experienced professionals and have opportunities to take on meaningful responsibilities.

Проектні активності

Conduct market research and analyze data to support business decisions. Prepare reports and presentations. Support marketing campaigns and social media management. Assist in content creation and brand development. Collaborate with cross-functional teams on various projects. Participate in team meetings and strategy sessions.

Робочі години

Monday to Saturday 5:30 pm - 9:00 pm, Sunday 8:00 am - 11:00 am

Доступні дати

03 Dec, 2026 - 28 Jan, 2027, 25 Jan, 2026 - 22 Mar, 2026, 30 Sep, 2026 - 23 Dec, 2026

Мови проєкту

Spanish

Вимоги

Освітній фон:
50%

Graphic design Arts

Навички:
Graphic Design Adobe Illustrator
Photoshop

Мови (обов'язкові):
50%

Spanish

Мінімальний рівень навчання:
Masters

Національність:
Non-EU

Умови

Проживання:
Only Provided — Housing will be provided at a cost.

Транспорт:
Only Provided — Transportation will be provided at a cost.

Харчування:
Only Provided — Food will be provided at a cost.

Медичне страхування:
Optional

Віза:
Work permit

Лідерські можливості:
Yes

Вартість

Внесок:
771,69 USD

Рисунок 4.6 – Детальна інформація сформованої рекомендації

Якщо порівнювати профіль користувача та сформовані рекомендації, можна сказати, що пропозиції сформовані успішно.

Тепер можемо провести перевірку із активним користувачем, котрий має історію взаємодій.

Історія взаємодій, профіль користувача та сформовані рекомендації для нього зображені на рисунках 4.7 – 4.10.



Рисунок 4.7 – Статистика взаємодій активного користувача

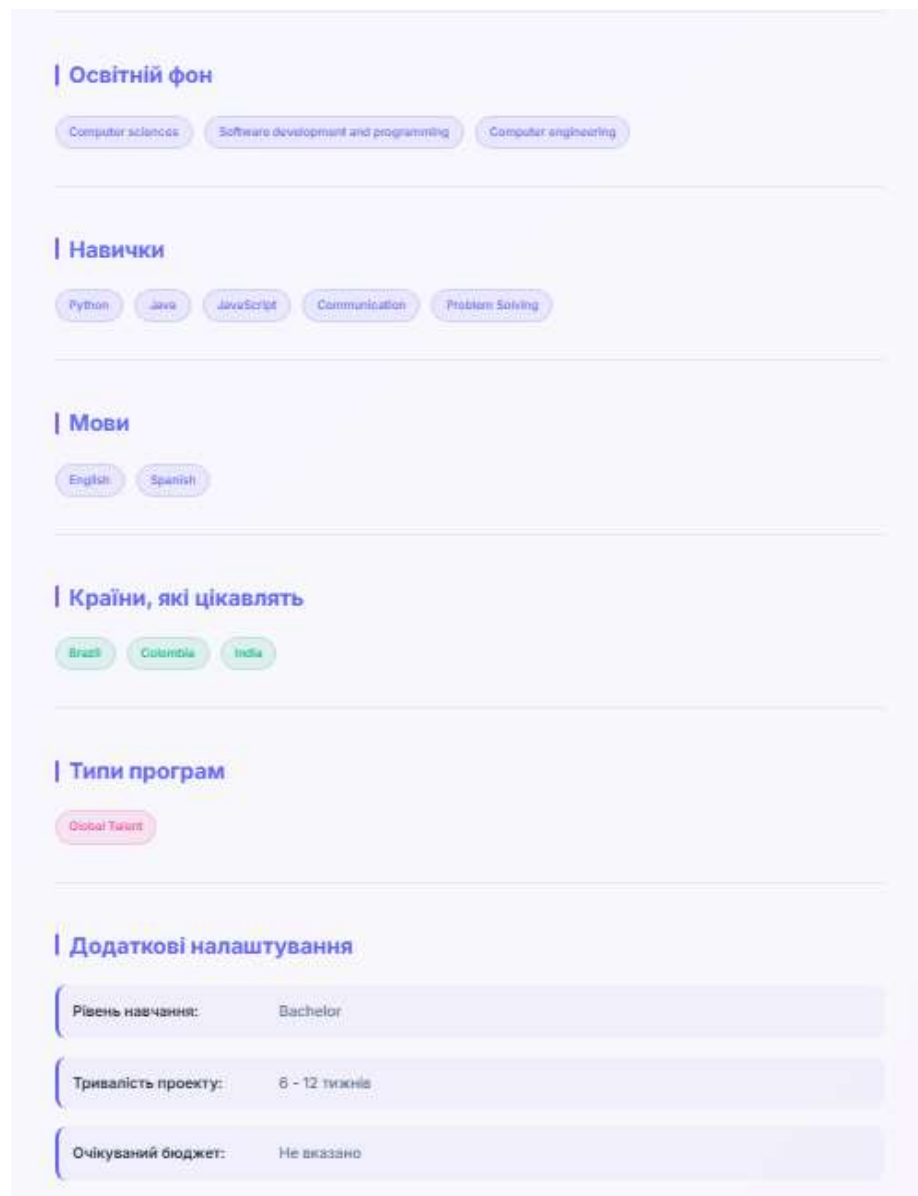


Рисунок 4.8 – Профіль активного користувача

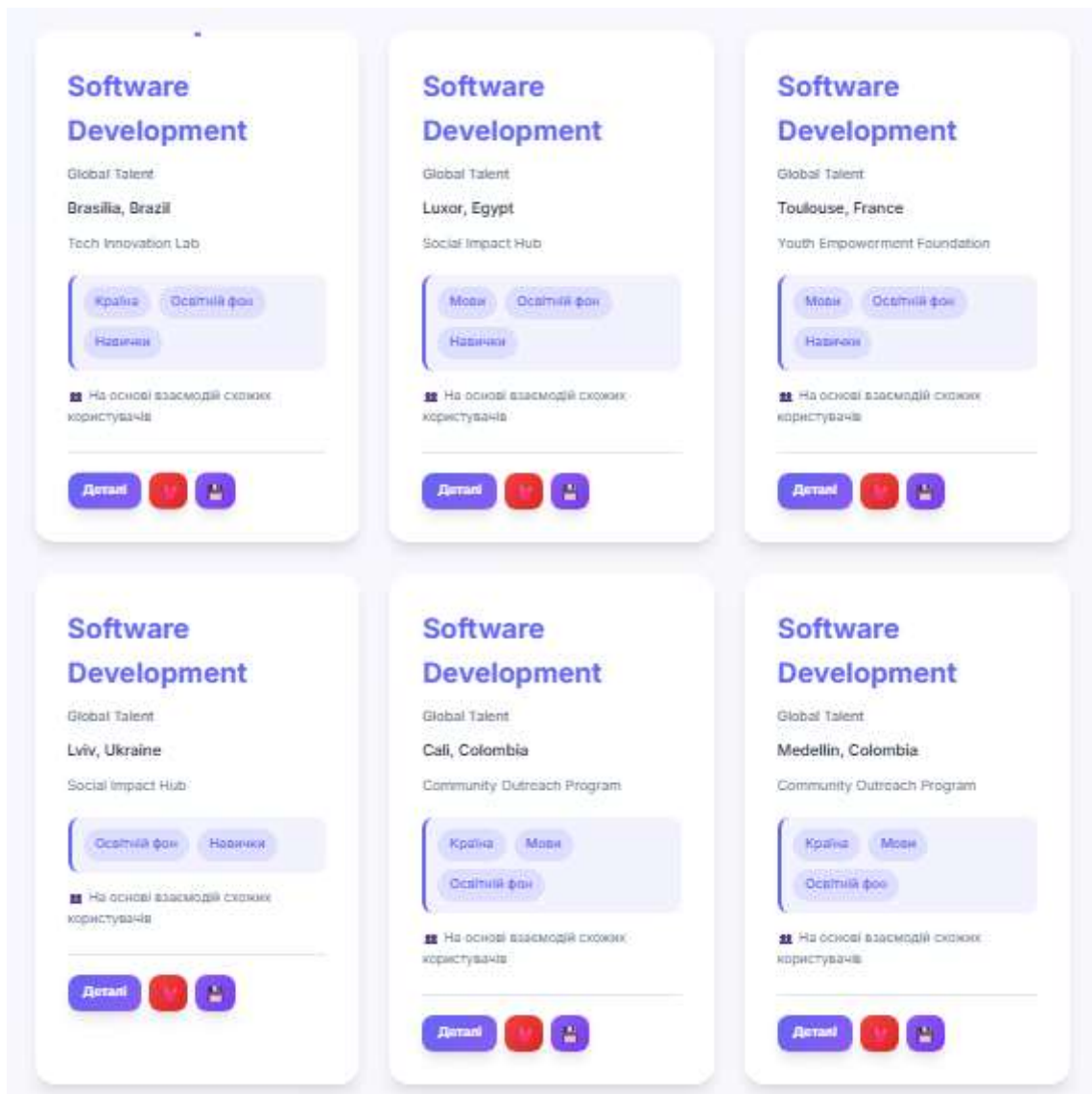


Рисунок 4.9 – Сформовані рекомендації для активного користувача

← Назад до проєктів

Software Development

Global Talent 100% 📍 Brasília, Brazil 100%

Про організацію

Tech Innovation Lab

Опис ролі

In this Software Development position, you will gain hands-on experience in a professional environment. Your responsibilities include supporting daily operations, contributing to team projects, and developing practical skills in your field. You'll work alongside experienced professionals and have opportunities to take on meaningful responsibilities.

Проектні активності

Conduct market research and analyze data to support business decisions. Prepare reports and presentations. Collaborate with cross-functional teams on various projects. Participate in team meetings and strategy sessions. Develop and implement new processes to improve efficiency. Support daily operations and administrative tasks.

Робочі години

Monday to Friday 10:00 am - 8:00 pm

Доступні дати

07 Jun, 2026 - 02 Aug, 2026, 02 Apr, 2026 - 28 May, 2026

Мови проєкту

Portuguese

Вимоги

Освітній фон:
100%
Computer engineering
Computer sciences
Software development

Навички:
100%
Programming
Software Development
Problem Solving

Мови (обов'язкові):
Portuguese

Мінімальний рівень навчання:
Bachelor

Національність:
EU

Умови

Проживання:
Only Provided — Housing will be provided at a cost.

Транспорт:
Only Provided — Transportation will be provided at a cost.

Харчування:
Only Provided — Food will be provided at a cost.

Медичне страхування:
Optional

Віза:
Work permit

Лідерські можливості:
Opportunities available

Рисунок 4.10 – Детальна інформація сформованої рекомендації

У висновку, можна сказати, що сформований гібридний метод успішно

працює у рекомендаційній системі та надає якісні пропозиції користувачеві. Проблема «холодного старту» також вирішена за допомогою впровадження КФ на основі атрибутів.

4.2 Експериментальна оцінка гібридного методу

4.2.1 Опис метрик для експериментальної оцінки

Мета експерименту – перевірити якість роботи сформованої гібридної система. Порівняти її з трьома основними методами – КФ на основі взаємодії користувачів, КФ на основі атрибутів, контентно–орієнтованою моделлю. Також необхідно підтвердити, чи дійсно вирішується проблема «холодного старту». Не менш важливо, оцінити, наскільки точно ця модель формує персоналізовані пропозиції для користувачів.

Для перевірки якості пропозицій застосовано метод відкладеної вибірки. Дані про взаємодії користувачів поділили на дві групи:

- навчальний набір – це 80% минулих даних; на основі яких створювали матриці взаємодій та навчали алгоритми;
- тестовий набір – 20% даних, прихованих від моделі, які використовувалися для перевірки точності прогнозів.

Для кожного учасника тестового набору створювали $K=20$ рекомендацій та зіставляли з справжніми даними про взаємодії.

Для аналізу ефективності системи обрано комплекс метрик, що оцінюють аспекти якості: точність, повноту, якість ранжування та різноманітність.

$Precision@K$ дозволяє оцінити, наскільки користувач може довіряти топ–списку рекомендацій. Для волонтерської платформи це критично важливо, оскільки користувач рідко переглядає більше 10 пропозицій за раз [25]. Дана метрика виконується за формулою:

$$Precision@K = \frac{|Rel(u) \cap TopK(u)|}{K}, \quad (4.1)$$

де u – користувач, для якого обчислюється метрика;

$Rel(u)$ – множина релевантних проєктів для користувача (в тестових даних);

$TopK(u)$ – множина проєктів, рекомендованих користувачу у топ- K ;

$|Rel(u) \cap TopK(u)|$ – кількість релевантних проєктів у списку рекомендацій;

K – довжина списку рекомендацій.

$Recall@K$ відображає здатність системи "не губити" цікаві можливості. Високий $Recall$ означає, що система ефективно покриває інтереси користувача [25]. Дана метрика виконується за формулою:

$$Recall@K = \frac{|Rel(u) \cap TopK(u)|}{|Rel(u)|}, \quad (4.2)$$

де u – користувач, для якого обчислюється метрика;

$Rel(u)$ – множина релевантних проєктів для користувача (в тестових даних);

$TopK(u)$ – множина проєктів, рекомендованих користувачу у топ- K ;

$|Rel(u) \cap TopK(u)|$ – кількість релевантних проєктів у списку рекомендацій;

$|Rel(u)|$ – загальна кількість релевантних проєктів у тестовій вибірці для користувача.

$AP@K$ враховує не просто число потрібних рекомендацій, але й те, де саме вони знаходяться у списку. Якщо корисні проєкти стоять вище – значення AP зростає. Такий підхід допомагає перевірити, наскільки добре метод сортує результати за важливістю [26]. Дана метрика працює за формулою:

$$AP@K(u) = \frac{1}{|Rel(u)|} \sum_{i=1}^K Precision@i \cdot rel(i), \quad (4.3)$$

де i – позиція у списку рекомендацій;

$Precision@i$ – точність на зрізі до позиції i ;

$rel(i)$ – бінарний індикатор релевантності;

$|Rel(u)|$ – кількість релевантних проєктів для користувача;

$AP@K(u)$ – середня точність рекомендацій для користувача u .

$MAP@K$ – це усереднена точність ранжування по всіх користувачах, яка використовується для оцінки якості моделі. Високе значення MAP означає, що система в середньому ставить релевантні проєкти вище незалежно від типу користувача [26]. Дана метрика працює за формулою:

$$MAP@K = \frac{1}{N} \sum_{u=1}^N AP@K(u), \quad (4.4)$$

де N – загальна кількість користувачів у вибірці;

u – індекс користувача;

$AP@K(u)$ – середня точність рекомендацій для користувача u ;

$MAP@K$ – середнє значення $AP@K$ для всіх користувачів (mean average precision).

$NDCG@K$ – ця метрика ідеально підходить для задач, де релевантність є небінарною. Вона використовує логарифмічне дисконтування, штрафуючи систему за розміщення високорелевантних проєктів у кінці списку [27]. Формули з якої складається дана метрика, виглядають як:

$$DCG_k = \sum_{i=1}^k \frac{rel_i}{\log_2(i+1)}, \quad (4.5)$$

де i – позиція елемента у списку рекомендацій;

rel_i – релевантність елемента на позиції i ;

$\log_2(i + 1)$ – логарифмічний штраф за віддаленість позиції;

DCG_k – зважена сума релевантностей у топ- k .

$$IDCG_k = \sum_{i=1}^k \frac{rel_i}{\log_2(i+1)}, \quad (4.6)$$

де i – позиція у списку;

rel_i – релевантність при ідеальному впорядкуванні;

$\log_2(i + 1)$ – штраф за позицію;

$IDCG_k$ – максимальне можливе значення DCG при ідеальному сортуванні.

$$NDCG_k = \frac{DCG_k}{IDCG_k}, \quad (4.7)$$

де DCG_k – фактичний показник упорядкованості рекомендацій;

$IDCG_k$ – ідеальне DCG для того самого списку;

$NDCG_k$ – нормалізована якість ранжування у межах [0; 1].

4.2.2 Результати експерименту

Після застосування вищезазначених метрик та проведення експерименту отримано результати оцінювання гібридного методу. Результати експерименту наведено у таблиці 4.1.

Таблиця 4.1 – Результати оцінювання гібридного методу

Метод	Метрика	K=5	K=10	K=20
Контентна–орієнтована фільтрація	Precision@K	0.0909	0.0727	0.0535
	Recall@K	0.1287	0.1599	0.1599
	MAP@K	0.0469	0.0515	0.0515
	NDCG@K	0.1282	0.1194	0.1166
КФ на основі атрибутів	Precision@K	0.0182	0.0273	0.0219
	Recall@K	0.0083	0.0256	0.0559
	MAP@K	0.0083	0.0119	0.0141
	NDCG@K	0.0308	0.0329	0.0423

Кінець таблиці 4.1

Метод	Метрика	K=5	K=10	K=20
КФ на основі взаємодій	Precision@K	0.0182	0.0273	0.0219
	Recall@K	0.0083	0.0256	0.0559
	MAP@K	0.0083	0.0119	0.0141
	NDCG@K	0.0308	0.0329	0.0423
Гібридний метод	Precision@K	0.1273	0.1000	0.0699
	Recall@K	0.1529	0.1932	0.1932
	MAP@K	0.0565	0.0637	0.0637
	NDCG@K	0.1630	0.1524	0.1497

Як видно з таблиці 4.1, гібридний метод стабільно демонструє найкращі результати за всіма метриками та на всіх рівнях.

Графіки результатів експериментів наведені на рисунках 4.11 – 4.14.

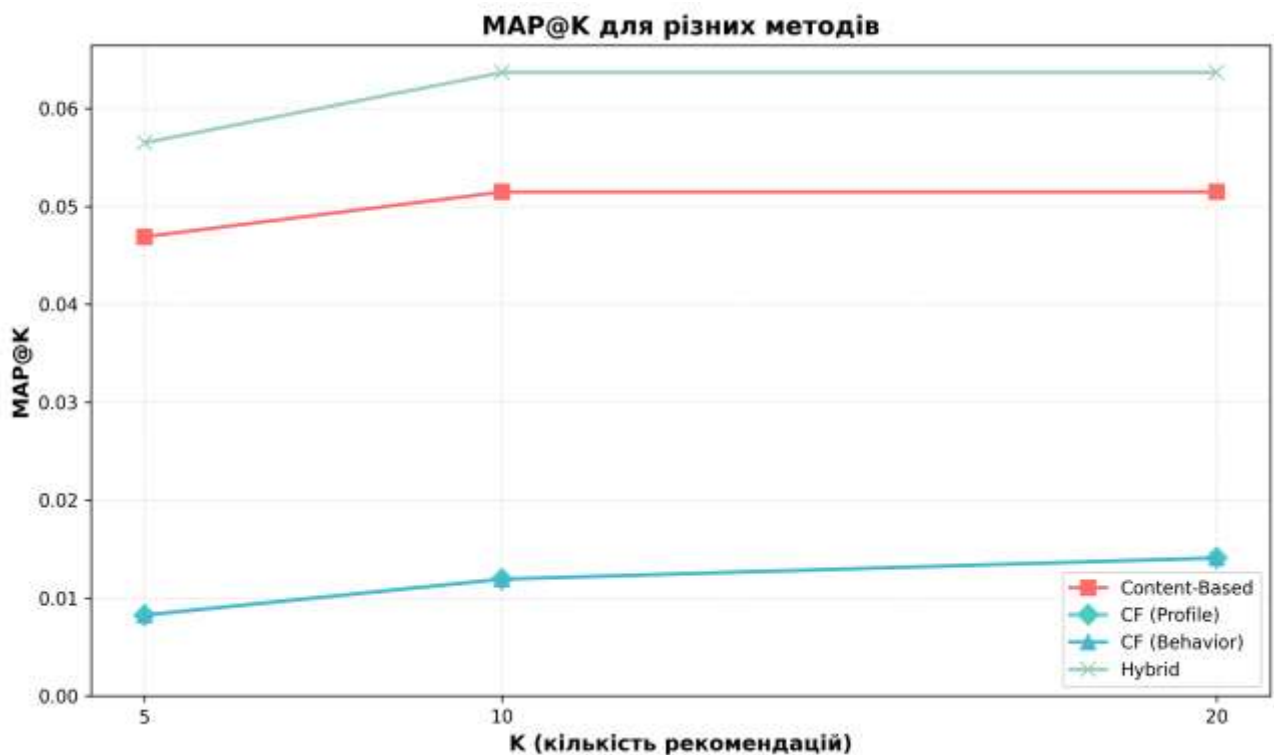


Рисунок 4.11 – Результати порівняння методів за метрикою MAP@K

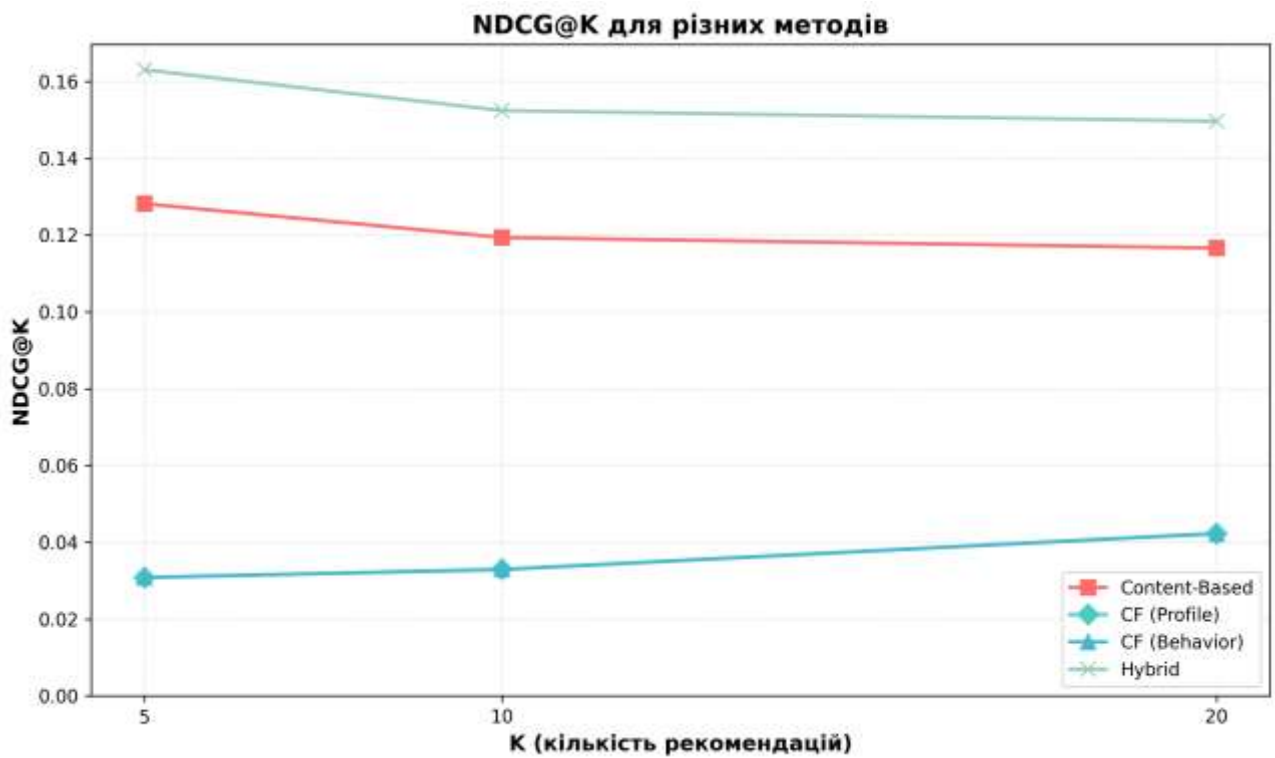


Рисунок 4.12 – Результати порівняння методів за метрикою NDCG@K

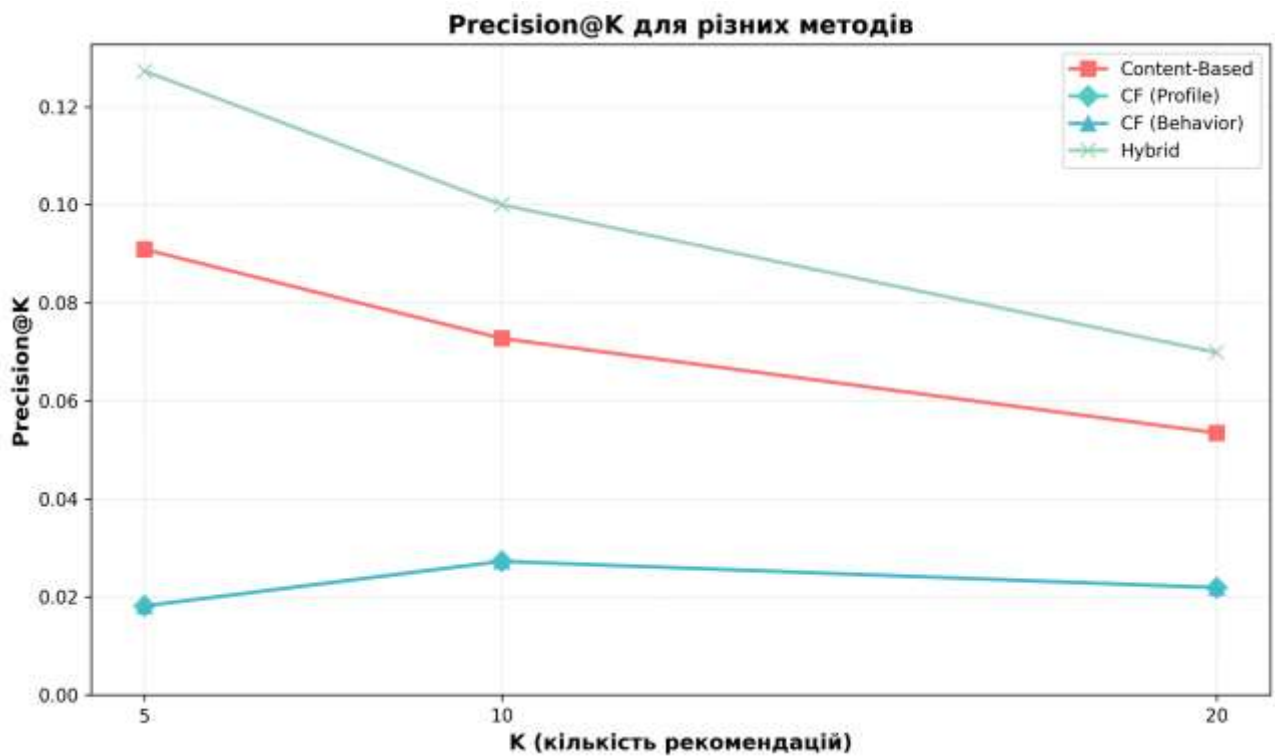


Рисунок 4.13 – Результати порівняння методів за метрикою Precision@K

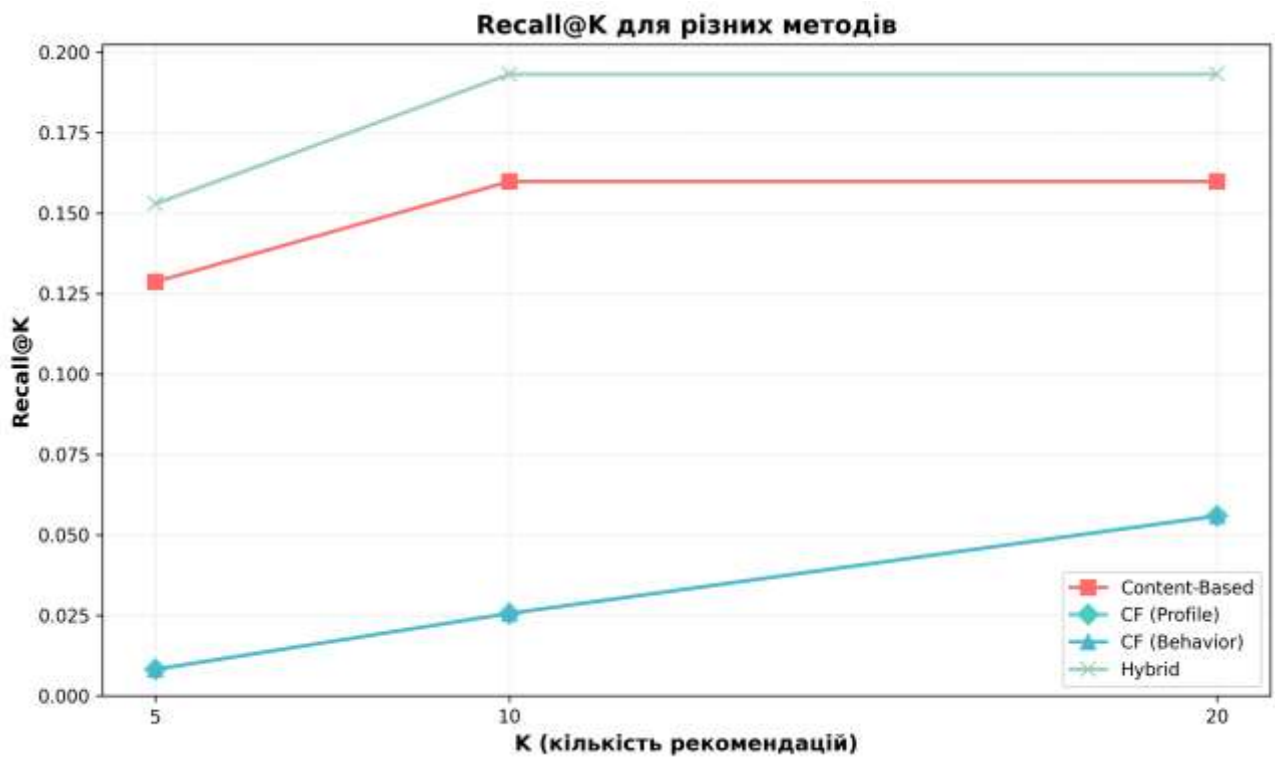


Рисунок 4.14 – Результати порівняння методів за метрикою Recall@K

ВИСНОВКИ

У ході кваліфікаційної роботи детально вивчено методи побудови рекомендацій у інформаційній системі AIESEC. На основі аналізу запропоновано гібридну модель, яка покращує точність рекомендацій щодо проєктів для молоді.

Аналіз предметної області допоміг чітко окреслити специфіку роботи AIESEC і основні функції, потрібні для рекомендаційної системи. Вивчення наявних систем-аналогів разом із науковими матеріалами виявило слабкі й сильні сторони класичних методів.

Контентно-орієнтована модель ефективна за структурованими описами, проте не враховує взаємодію користувачів; і навпаки, колаборативна фільтрація спирається на минулі взаємодії, однак страждає від проблеми початкового періоду.

Досліджено КФ на основі взаємодій користувачів й атрибутів. На основі цього окреслено ключові вимоги до алгоритмів. Виявилось, що поєднання різних стратегій робить систему гнучкішою. З урахуванням отриманих даних створено гібридний метод, який об'єднує дані контентного аналізу з оцінками подібності між учасниками. Упровадження методики дало можливість значно підвищити точність рекомендацій і врахувати ширший спектр волонтерських напрямів.

Експериментальна оцінка показала, що запропонований метод працює краще. Точність виявилась вищою за результати окремих моделей. Такі дані свідчать про те, що поєднання різних підходів має сенс у рекомендаційних системах.

Таким чином, результати дослідження можна застосувати для модернізації платформ AIESEC – зокрема, щоб оптимізувати підбір міжнародних волонтерських проєктів. Це також дає основу для тестування оригінальних алгоритмів у контексті персоналізованих пропозицій.

Роботу виконано відповідно до вимог методичних вказівок щодо розробки та оформлення кваліфікаційних робіт [28], методичних рекомендацій з проходження передатестаційної практики [29], а також із дотриманням положень державних стандартів ДСТУ 8302:2015 [30] і ДСТУ 3008:2015 [31].

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. About us. *AIESEC*. URL: <https://aiesec.org/about-us> (date of access: 22.10.2025).
2. Volunteering. *AIESEC*. URL: <https://aiesec.org/global-volunteer> (date of access: 22.10.2025).
3. Global Talent. *AIESEC*. URL: <https://aiesec.ua/global-talent> (дата звернення: 22.10.2025).
4. Local Volunteer. *AIESEC*. URL: <https://aiesec.ua/local-volunteer> (дата звернення: 22.10.2025).
5. Platform basics & overview. *AIESEC*. URL: <https://support.aiesec.org/hc/en-us/categories/4403785357588-Platform-basics-overview> (date of access: 22.10.2025)
6. Resnick P., Varian H. Recommender systems. *dl.acm.org/doi*. URL: <https://dl.acm.org/doi/10.1145/245108.245121> (date of access: 13.11.2025).
7. Adomavicius G., Tuzhilin A. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Xplore*. URL: <https://ieeexplore.ieee.org/document/1423975> (date of access: 13.11.2025).
8. Covington P., Adams J., Sargin E. Deep Neural Networks for YouTube Recommendations. *https://dl.acm.org/*. URL: <https://dl.acm.org/doi/10.1145/2959100.2959190> (date of access: 14.11.2025).
9. From Predictive to Generative – How Michelangelo Accelerates Uber’s AI Journey / K. Wang et al. *Uber Blog*. URL: <https://www.uber.com/en-UA/blog/from-predictive-to-generative-ai/> (date of access: 14.11.2025).
10. Content Based Recommendation System on Netflix Data. *ResearchGate*. URL: https://www.researchgate.net/publication/378337926_Content_Based_Recommendation_System_on_Netflix_Data (date of access: 14.11.2025).
11. Gupta R. Near real-time features for near real-time personalization. *LinkedIn*.

URL: <https://www.linkedin.com/blog/engineering/member-customer-experience/near-real-time-features-for-near-real-time-personalization> (date of access: 14.11.2025).

12. Linden G., Smith B., York J. Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Xplore*.

URL: <https://ieeexplore.ieee.org/document/1167344> (date of access: 15.11.2025).

13. Redesigning VolunteerMatch's Ranking Algorithm: Toward More Equitable Access to Volunteers / V. Manshadi et al. *ResearchGate*.

URL: https://www.researchgate.net/publication/372145696_Redesigning_Volunteer_Match's_Ranking_Algorithm_Toward_More_Equitable_Access_to_Volunteers (date of access: 15.11.2025).

14. Frequently Asked Questions. *European Youth Portal*. URL: https://youth.europa.eu/solidarity/faq_en (date of access: 15.11.2025).

15. IAESTE A.s.b.l. FAQ. *IAESTE A.s.b.l.* URL: <https://iaeste.org/student-faqs> (date of access: 15.11.2025).

16. How It Works. *International Volunteer HQ*. URL: <https://www.volunteerhq.org/how-it-works/> (date of access: 15.11.2025).

17. How it works. *Volunteer Abroad & Internship Programs | Projects Abroad UK*. URL: <https://www.projects-abroad.co.uk/how-it-works/> (date of access: 15.11.2025).

18. Ricci F., Rokach L., Shapira B. Recommender Systems Handbook. *ResearchGate*.

URL: https://www.researchgate.net/publication/227268858_Recommender_Systems_Handbook (date of access: 16.11.2025).

19. Li X., Zhang T. Collaborative Filtering Recommendation Algorithm Based on User Attributes and Item Score. *Wiley Online Library*.

URL: <https://onlinelibrary.wiley.com/doi/10.1155/2022/4544152> (date of access: 16.11.2025).

20. Collaborative Filtering-Based Recommender Systems: A Deep Dive. *Futureweb AI*. URL: <https://futurewebai.com/blogs/collaborative-filtering->

[based-recommendation](#) (date of access: 25.11.2025).

21. ramzanzdemir. Recommendation Systems: Content-Based TF-IDF. *Checking your browser - reCAPTCHA.*

URL: <https://www.kaggle.com/code/ramzanzdemir/recommendation-systems-content-based-tf-idf> (date of access: 25.11.2025).

22. What is Python? - Python Language Explained - AWS. *Amazon Web Services, Inc.* URL: <https://aws.amazon.com/what-is/python/> (date of access: 02.12.2025).

23. What is NumPy? – NumPy v2.3 Manual. *NumPy.* URL: <https://numpy.org/doc/stable/user/whatisnumpy.html> (date of access: 02.12.2025).

24. SciPy User Guide – SciPy v1.16.2 Manual. *NumPy and SciPy Documentation – Numpy and Scipy documentation.* URL: <https://docs.scipy.org/doc/scipy/tutorial/> (date of access: 02.12.2025).

25. Precision and recall at K in ranking and recommendations. *Evidently AI - AI Evaluation & LLM Observability Platform.* URL: <https://www.evidentlyai.com/ranking-metrics/precision-recall-at-k> (date of access: 05.12.2025).

26. Rink K. Mean Average Precision at K (MAP@K) clearly explained. *Toward Data Science.* URL: <https://towardsdatascience.com/mean-average-precision-at-k-map-k-clearly-explained-538d8e032d2/> (date of access: 05.12.2025).

27. Normalized Discounted Cumulative Gain (NDCG): Where To Use It. *Arize AI.* URL: <https://arize.com/blog-course/ndcg/> (date of access: 05.12.2025).

28. Методичні вказівки щодо розробки та оформлення кваліфікаційної роботи другого (магістерського) рівня вищої освіти за освітньо-професійною програмою «Інформаційні управляючі системи та технології» спеціальності 122 Комп'ютерні науки / Упоряд.: К.Е. Петров, В.М. Левикін, С.Ф. Чалий, М.В. Євланов, В.І. Саєнко, Д.К. Міхнов, А.В. Міхнова, О.В. Чала. ХНУРЕ: Харків, 2024. 24 с.

29. Методичні вказівки до передатестаційної практики для здобувачів

вищої освіти усіх форм навчання другого (магістерського) рівня вищої освіти спеціальності 122 Комп'ютерні науки освітньо-професійної програми «Інформаційні управляючі системи та технології». / Упоряд.: К.Е. Петров, С.Ф. Чалий, О.Д. Міхнова. Харків: ХНУРЕ, 2024. 19 с.

30. ДСТУ 8302:2015. Бібліографічне посилання. Загальні положення та правила складання. К.: ДП «УкрНДНЦ», 2016. 20 с.

31. ДСТУ 3008:2015. Державний стандарт України. Інформація та документація. Звіти у сфері науки і техніки. Структура та правила оформлювання. К.: ДП «УкрНДНЦ», 2016. 31 с.