

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ЗАПОРІЗЬКА ПОЛІТЕХНІКА»

Факультет комп'ютерних наук та технологій
Кафедра комп'ютерних систем та мереж

Пояснювальна записка

до дипломного проекту (роботи)

магістра

(ступінь вищої освіти)

на тему ВЕБПЛАТФОРМА ПРОФОРІЄНТАЦІЇ З АНАЛІЗОМ РИНКУ ПРАЦІ ТА
ПРОГНОЗУВАННЯМ ЗАТРЕБУВАНОСТІ ПРОФЕСІЙ

Виконав: студент 2 курсу, групи КНТ-614м
спеціальності _____

123 Комп'ютерна інженерія

(код і найменування спеціальності)

Освітня програма (спеціалізація)

Спеціалізовані комп'ютерні системи

СІРЕНКО В. В.

(ПРИЗВИЩЕ та ініціали)

Керівник ГРУШКО С.С.

(ПРИЗВИЩЕ та ініціали)

Рецензент РОМАНОВСЬКИЙ О.В.

(ПРИЗВИЩЕ та ініціали)

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний університет «Запорізька політехніка»

Факультет Комп'ютерних наук і технологій
Кафедра «Комп'ютерні системи та мережі»
Ступінь вищої освіти магістерський
Спеціальність 123 Комп'ютерна інженерія
(код і найменування)
Освітня програма (спеціалізація) Спеціалізовані комп'ютерні системи
(назва освітньої програми (спеціалізації))

ЗАТВЕРДЖУЮ

Зав. кафедри Кудерметов Р.К.

“ ” 2025 року

З А В Д А Н Н Я
НА ДИПЛОМНИЙ ПРОЄКТ (РОБОТУ) СТУДЕНТА

СІРЕНКА Володимира Володимировича

(ПРИЗВИЩЕ, ім'я, по батькові)

1. Тема проєкту (роботи) Вебплатформа профорієнтації з аналізом ринку праці та прогнозуванням затребуваності професій

керівник проєкту (роботи) к. т. н., доцент, ГРУШКО Світлана Сергіївна

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

затверджені наказом вищого навчального закладу від “30” жовтня 2025 року № 491

2. Строк подання студентом проєкту (роботи) 23 грудня 2025 року

3. Вихідні дані до проєкту (роботи) персональні комп'ютери та мобільні пристрої з сучасними веббраузерами, LocalStorage, стандартні браузерні події та HTTPS для завантаження сторінки

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

1) Аналіз предметної області та постановка задачі;

2) Дослідження та порівняльний аналіз алгоритмів рекомендацій;

3) Проєктування та реалізація системи професійної орієнтації;

4) Апаратне забезпечення;

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

6. Консультанти розділів проєкту (роботи)

Розділ	ПРИЗВИЩЕ, ініціали та посада консультанта	Підпис, дата	
		завдання видав	прийняв виконане завдання
1-4	ГРУШКО С. С., доцент		
нормоконтроль	ПОЛЬСЬКА О.В., ст. викл.		

7. Дата видачі завдання 01.10.2025 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проєкту (роботи)	Строк виконання етапів проєкту (роботи)	Примітка
1	Аналіз предметної області та постановка задачі	10.10.2025 р.	
2	Розробка архітектури системи	15.10.2025 р.	
3	Налагодження програмного забезпечення	20.10.2025 р.	
4	Реалізація архітектури системи	05.10.2025 р.	
5	Оформлення отриманих результатів у ПЗ	15.11.2025 р.	
6	Оформлення графічного матеріалу	25.11.2025 р.	
7	Оформлення допоміжного матеріалу	01.12.2025 р.	

Студент **Володимир СІРЕНКО**
(підпис) (ім'я, ПРИЗВИЩЕ)

Керівник проєкту (роботи) **Світлана ГРУШКО**
(підпис) (ім'я, ПРИЗВИЩЕ)

РЕФЕРАТ

ПЗ: 94 с., 11 рис., 1 табл., 2 дод., 31 джерело.

CONTENT-BASED FILTERING, FASTAPI, MONGODB, REACT, ВЕБ-ПЛАТФОРМА, МАШИННЕ НАВЧАННЯ, РЕКОМЕНДАЦІЙНІ СИСТЕМИ ПРОФЕСІЙНА ОРІЄНТАЦІЯ

Об'єкт дослідження – процес професійної орієнтації молоді з використанням інтелектуальних інформаційних систем.

Предмет дослідження – методи та засоби створення веб-платформи професійної орієнтації на основі алгоритмів машинного навчання.

Мета роботи – розробка веб-платформи професійної орієнтації з використанням алгоритму Content-Based Filtering для генерації персоналізованих професійних рекомендацій.

У першому розділі проведено комплексний аналіз предметної області професійної орієнтації, досліджено існуючі системи та визначено недоліки традиційних підходів.

У другому розділі розроблено математичну модель системи рекомендацій на основі Content-Based Filtering.

Третій розділ присвячено проєктуванню та реалізації системи CareerGuide з використанням трьохрівневої клієнт-серверної архітектури. Описано технологічний стек FastAPI-React-MongoDB, розроблено структуру бази даних та реалізовано алгоритм генерації рекомендацій з багаторівневою оптимізацією продуктивності.

У четвертому розділі визначено технічні вимоги до апаратного та мережевого забезпечення системи. Розроблено вимоги до серверної інфраструктури, клієнтського обладнання та мережевої архітектури для забезпечення надійної роботи платформи з можливістю масштабування.

ABSTRACT

Explanatory note to the master's work: 94 p., 11 figures, 1 tables, 2 additions, 31 sources.

CAREER GUIDANCE, CONTENT-BASED FILTERING, FASTAPI, MACHINE LEARNING, MONGODB, REACT, RECOMMENDER SYSTEMS, WEB PLATFORM

Research object – the process of career guidance for youth using intelligent information systems.

Research subject – methods and tools for creating a career guidance web platform based on machine learning algorithms.

Research goal – development of a career guidance web platform using Content-Based Filtering algorithm to generate personalized career recommendations.

The first section conducts a comprehensive analysis of the career guidance domain, examines existing systems, and identifies shortcomings of traditional approaches. The critical necessity of creating an innovative system that accounts for modern IT professions and uses scientifically validated methodologies is revealed.

The second section develops a mathematical model of the recommendation system based on Content-Based Filtering. A five-dimensional model of professional characteristics is created, and the choice of cosine similarity as a metric for calculating the match between user profiles and professional requirements is justified.

The third section focuses on the design and implementation of the CareerGuide system using a three-tier client-server architecture.

The fourth section defines technical requirements for hardware and network infrastructure of the system. Requirements for server infrastructure, client equipment, and network architecture are developed to ensure reliable platform operation with scalability capabilities.

ЗМІСТ

Вступ.....	8
1 Аналіз предметної області та постановка задачі	9
1.1 Аналіз сучасного стану професійної орієнтації	9
1.2 Огляд наявних систем професійної орієнтації	11
1.3 Аналіз рекомендаційних систем	18
1.4 Постановка задачі дослідження	23
1.5 Висновки до розділу 1	27
2 Дослідження та порівняльний аналіз алгоритмів рекомендацій.....	29
2.1 Теоретичні основи Content-Based Filtering	29
2.2 Адаптація Content-Based алгоритмів для професійної орієнтації	34
2.3 Математична модель системи рекомендацій	39
2.4 Висновки до розділу 2	45
3 Проєктування та реалізація системи професійної орієнтації	46
3.1 Архітектура системи	46
3.1.1 Загальна архітектура платформи	46
3.1.2 Компонентна діаграма системи	47
3.1.3 Вибір технологічного стеку	48
3.1.4 Паттерни проєктування	49
3.2 Проєктування бази даних	51
3.2.1 ER-діаграма системи.....	51
3.2.2 Структура колекцій MongoDB.....	52
3.2.3 Індексування для оптимізації запитів	53
3.2.4 Стратегії збереження векторних даних.....	54
3.3 Реалізація Content-Based алгоритму	55
3.3.1 Структури даних для зберігання векторів	55
3.3.2 Алгоритм конвертації відповідей тесту у вектор користувача	56
3.3.3 Функція розрахунку рекомендацій.....	58

3.3.4 Оптимізація продуктивності алгоритму	59
3.4 Розробка користувацького інтерфейсу	60
3.4.1 UX/UI дизайн платформи	60
3.4.2 Компонентна архітектура React додатку	61
3.4.3 Адаптивний дизайн для різних пристроїв	66
3.4.4 Візуалізація результатів рекомендацій	67
3.4.5 Інтеграція з backend API	68
3.5 Реалізація аналітичного модуля.....	69
3.5.1 Система метрик ефективності.....	69
3.5.2 Дашборди для аналізу користувацької поведінки	70
3.5.3 Статистика популярності професій.....	71
3.5.4 А/В тестування різних версій алгоритму.....	72
3.6 Висновки до розділу 3	73
4 Апаратне забезпечення	75
4.1 Серверна інфраструктура	75
4.2 Вимоги до клієнтського обладнання.....	76
4.3 Мережева інфраструктура.....	76
4.4 Висновки до розділу 4	77
Висновки	79
Перелік джерел посилання	82
Додаток А Структура проєкту	85
Додаток В Лістинги програм	86

ВСТУП

У сучасних умовах стрімкого розвитку цифрових технологій та динамічних змін на ринку праці проблема ефективної профорієнтації набуває особливої актуальності. Традиційні підходи до вибору професійної траєкторії вже не забезпечують достатнього рівня інформативності та гнучкості, оскільки потреби роботодавців швидко змінюються, з'являються нові спеціальності, а вимоги до компетентностей постійно оновлюються. У результаті здобувачі освіти та молоді фахівці часто стикаються зі складністю прийняття виваженого рішення щодо свого професійного розвитку.

Значна кількість наявних вебресурсів, що надають профорієнтаційні матеріали, не враховують актуальні аналітичні дані ринку праці та не пропонують прогнозів щодо перспективності професій у середньостроковій перспективі. Це створює інформаційний розрив між потребами користувачів та реальними тенденціями розвитку галузей, ускладнюючи вибір оптимальної професійної траєкторії.

Метою даної магістерської роботи є розроблення вебплатформи профорієнтації, що поєднує аналітику ринку праці, методи прогнозування затребуваності професій та інтерактивні інструменти для користувачів. Запропонована система забезпечує доступ до актуальної інформації про професії, їхній розвиток, компетентнісні вимоги та тенденції працевлаштування, а також створює персоналізовані рекомендації на основі аналітичних моделей.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАДАЧІ

1.1 Аналіз сучасного стану професійної орієнтації

Професійна орієнтація як складова практичної психології та педагогіки спрямована на забезпечення усвідомленого вибору здобувачами освіти майбутньої професії відповідно до їхніх індивідуальних особливостей та актуальних потреб ринку праці. В умовах швидких економічних і технологічних змін традиційні підходи поступово втрачають ефективність через обмежений обсяг інформації та недостатню адаптивність.

Класична система профорієнтації в Україні, сформована ще в радянський період, спирається на три взаємопов'язані напрями: професійну інформацію, професійну консультацію та професійний відбір. Основними інструментами залишаються психодіагностичні методики, серед яких диференційно-діагностичний опитувальник Клімова, карта інтересів Голомштока та тест Голланда. Ці методи досі широко застосовуються в центрах зайнятості та закладах освіти.

Ключовим недоліком традиційних методів є їхня незмінність. Опитувальники, створені багато років тому, не охоплюють новітні спеціальності, що з'явилися у сферах інформаційних технологій, цифрової комунікації, віртуальної реальності чи штучного інтелекту.

Окремою проблемою є суб'єктивність інтерпретацій. Рекомендації часто залежать від особистого досвіду консультанта, що може призводити до різних результатів для схожих профілів кандидатів та ускладнює формування точних висновків щодо професійної придатності.

Використання цифрових технологій створює можливості для подолання цих обмежень. Автоматизовані системи тестування забезпечують стандартизованість процедур, підвищують об'єктивність оцінок та роблять профорієнтаційні послуги доступнішими завдяки використанню вебплатформ.

Особливої актуальності набуває застосування алгоритмів машинного навчання [1, 3, 5, 16, 20, 22, 28, 30], які здатні аналізувати профілі користувачів та формувати індивідуальні рекомендації. На відміну від фіксованих опитувальників, такі системи постійно оновлюються відповідно до змін ринку праці та можуть враховувати реальний характер попиту на професії на основі аналізу вакансій.

Інтеграція з інформаційними ресурсами, що містять дані щодо заробітних плат, вимог роботодавців та перспектив розвитку окремих спеціальностей, надає можливість забезпечити користувачам максимально актуальну та обґрунтовану інформацію для планування кар'єрної траєкторії.

Український ринок праці має низку структурних диспропорцій між попитом та пропозицією кадрів. Дані Державної служби статистики України свідчать про наявність галузей, у яких роботодавці систематично відчують складнощі з пошуком фахівців, зокрема в технологічному та інженерному напрямках. Водночас освітня система продовжує здійснювати підготовку кадрів у сферах, де рівень фактичного працевлаштування після закінчення навчання залишається низьким.

Рівень трудової міграції також істотно впливає на структуру зайнятості в Україні. Значна кількість фахівців працює за межами країни, що формує додаткову потребу в спеціальностях, орієнтованих на міжнародний ринок праці. Це вимагає врахування глобальних тенденцій під час формування профорієнтаційних рекомендацій [8, 11, 12].

Дослідження українських аналітичних центрів засвідчують зміну пріоритетів молоді у виборі професії: зростає інтерес до можливостей розвитку, гнучких умов праці та балансу між професійними й особистими потребами. Молоде покоління також демонструє готовність опановувати нові спеціальності протягом життя, що посилює потребу у формуванні не лише конкретних рекомендацій, а й компетентностей адаптивності та безперервного навчання.

Цифрова грамотність стає базовою вимогою для більшості сучасних професій. Значна частина вакансій містить вимоги щодо впевненого володіння комп'ютером та спеціалізованими програмними інструментами, що підтверджується даними провідних українських порталів з пошуку роботи.

Цифровізація економіки призводить до трансформації багатьох професій і появи нових. Згідно з прогнозами міжнародних організацій, у найближчі роки суттєво зміниться структура зайнятості, а багато традиційних професій або трансформуються, або зникнуть. Паралельно формуються нові професійні напрями, пов'язані з аналізом даних, штучним інтелектом і цифровими технологіями. За таких умов системи профорієнтації повинні враховувати не тільки поточний стан ринку праці, а й перспективні тенденції. Статичні інструменти вже не здатні ефективно реагувати на такі динамічні зміни, що зумовлює потребу у впровадженні інтелектуальних адаптивних рішень.

Поширення дистанційної роботи та співпраці з міжнародними компаніями розширює географічні можливості працевлаштування для українських спеціалістів. Це посилює важливість урахування світових тенденцій при формуванні рекомендацій у сфері професійної орієнтації.

Отже, сучасна система профорієнтації в Україні потребує оновлення та впровадження методів, які забезпечують актуальність, персоналізацію та об'єктивність рекомендацій. Використання інформаційних технологій та алгоритмів машинного навчання дає змогу суттєво підвищити ефективність профорієнтаційних сервісів і зробити їх більш відповідними динамічним вимогам ринку праці та очікуванням молоді.

1.2 Огляд наявних систем професійної орієнтації

Сучасний ринок програмного забезпечення для професійної орієнтації включає широкий спектр рішень від простих онлайн-опитувальників до комплексних систем з елементами штучного інтелекту. Аналіз наявних платформ дозволяє виявити найкращі практики та ідентифікувати прогалини в функціональності, що можуть бути усунені при розробці нових рішень.

O*NET Interest Profiler представляє собою одну з найбільш авторитетних систем професійної орієнтації, розроблену Міністерством праці США. Система базується на теорії професійних інтересів Холланда та включає 60 питань, що класифікують користувачів за шістьма типами особистості: реалістичний, дослідницький, художній, соціальний, підприємницький та конвенційний. Основною перевагою системи є інтеграція з базою даних O*NET, що містить детальну інформацію про понад 900 професій з актуальними вимогами до освіти, навичок та заробітних плат. Алгоритмічну основу O*NET становить проста система підрахунку балів, де кожна відповідь користувача додає певну кількість балів до відповідного типу особистості за Холландом. Результируючий профіль порівнюється з профілями професій, і система рекомендує спеціальності з найвищим рівнем відповідності. Незважаючи на простоту алгоритму, система демонструє високу валідність завдяки ретельній верифікації профілів професій експертами галузі.

16Personalities Career Explorer використовує типологію Майєрс-Бріггс для визначення професійних схильностей користувачів. Система проводить комплексне тестування за чотирма дихотомічними шкалами: екстраверсія-інтроверсія, сенсорика-інтуїція, мислення-почуття, судження-сприйняття. Результатом є один з 16 психологічних типів, для кожного з яких надаються рекомендації щодо підходящих професій.

Технічно платформа реалізована як веб-додаток з адаптивним дизайном, що забезпечує коректне відображення на мобільних пристроях. Система використовує базу даних з понад 300 професійними профілями, кожен з яких містить інформацію про типовий психологічний тип працівника, середню зарплату та перспективи кар'єрного зростання. Особливістю платформи є детальна аналітика результатів з поясненнями щодо сумісності конкретного типу особистості з різними аспектами професійної діяльності.

LinkedIn Career Advice представляє гібридний підхід, що поєднує алгоритмічну обробку даних профілю користувача з консультаціями реальних професіоналів. Система аналізує освітню історію, попередній досвід роботи та навички, зазначені в профілі LinkedIn, для генерації персоналізованих рекомендацій щодо кар'єрного

розвитку. Алгоритм платформи використовує елементи колаборативної фільтрації, аналізуючи кар'єрні траєкторії користувачів з подібними профілями для прогнозування потенційних напрямків професійного розвитку. Система також інтегрує дані про ринок праці з власної бази вакансій LinkedIn, що дозволяє надавати актуальну інформацію про попит на різні спеціальності та рівні заробітних плат.

MyPlan.com Assessment пропонує багаторівневий підхід до професійної орієнтації, що включає тестування інтересів, цінностей, особистісних характеристик та навичок. Система використовує валідовані психометричні інструменти, зокрема Interest Assessment базований на теорії Холланда, Values Assessment для визначення пріоритетів в роботі, та Skills Assessment для оцінки наявних компетенцій. Технічна архітектура платформи побудована за принципом послідовного тестування з інтеграцією результатів на фінальному етапі. Кожен тест генерує окремий профіль користувача, а загальні рекомендації формуються на основі зважування результатів усіх оцінок. Система містить базу даних з понад 1000 професійних профілів та використовує складний алгоритм співставлення для ранжування рекомендацій.

Український ринок профорієнтаційних систем представлений переважно простими веб-додатками з обмеженим функціоналом. Тест на професійну орієнтацію від Work.ua є найбільш популярним вітчизняним рішенням, що базується на адаптованій версії опитувальника Клімова. Система включає 30 питань, розподілених за п'ятьма професійними сферами: людина-людина, людина-техніка, людина-знакова система, людина-художній образ, людина-природа. Алгоритм роботи системи полягає в підрахунку балів за кожною сферою та визначенні домінуючого типу професійних інтересів. Результати представляються у вигляді процентного розподілу схильностей користувача між п'ятьма сферами з короткими описами відповідних професій. Основним обмеженням системи є застарілість переліку професій, що не враховує сучасні IT-спеціальності та цифрові професії.

Профорієнтаційна система Osvita.ua пропонує розширений тест з 60 питань, спрямованих на визначення академічних інтересів та схильностей до різних предметних областей. Система орієнтована на учнів випускних класів та включає реко-

мендації щодо вибору спеціальності у вищих навчальних закладах України. Технічна реалізація платформи базується на простій системі підрахунку балів з інтеграцією до бази даних спеціальностей українських ВНЗ. Система надає інформацію про вступні вимоги, конкурсні бали попередніх років та перспективи працевлаштування за різними спеціальностями. Однак відсутність актуальних даних про ринок праці та примітивний алгоритм рекомендацій суттєво обмежують ефективність платформи.

Система профтестування JobTest.ua представляє комбінований підхід, що включає тестування професійних інтересів, особистісних характеристик та інтелектуальних здібностей. Платформа пропонує понад 20 різних тестів, від простих опитувальників інтересів до складних психометричних оцінок. Архітектурно система реалізована як модульна платформа, де кожен тест є окремим компонентом з власним алгоритмом обробки результатів. Інтеграція результатів різних тестів здійснюється через систему ваг, що дозволяє формувати комплексні профорієнтаційні рекомендації. Проте відсутність єдиної методологічної основи та неузгодженість між різними тестами знижують надійність загальних висновків.

Проведемо порівняльний аналіз систем за критеріями ефективності.

Методи тестування в аналізованих системах можна класифікувати за трьома основними підходами. Перший підхід базується на теорії професійних типів Холланда та використовується в O*NET Interest Profiler та MyPlan.com. Цей підхід демонструє високу теоретичну обґрунтованість та валідність, підтверджену десятиліттями досліджень. Другий підхід спирається на типологію Майєрс-Бріггс, що реалізована в 16Personalities Career Explorer. Незважаючи на популярність, цей підхід має обмежену наукову підтримку через критику щодо дихотомічної природи типів особистості. Третій підхід представлений адаптаціями радянської системи Клімова в українських рішеннях, що характеризується простотою реалізації, але застарілістю концептуальних основ.

Алгоритми рекомендацій у більшості аналізованих систем базуються на простих методах підрахунку балів та порівняння профілів. O*NET Interest Profiler використовує пряме співставлення користувацького профілю з професійними кодами

Холланда через евклідову відстань у шестивимірному просторі типів. 16Personalities застосовує детерміністичний підхід, де кожен з 16 типів особистості має заздалегідь визначений список рекомендованих професій. LinkedIn Career Advice є єдиною системою, що використовує елементи машинного навчання через аналіз кар'єрних траєкторій користувачів з подібними профілями. Українські системи обмежуються простим підрахунком балів без урахування складних взаємозв'язків між характеристиками користувача.

Точність прогнозування професійної придатності значно варіює між системами через різні методології валідації результатів. O*NET Interest Profiler демонструє найвищі показники валідності з коефіцієнтом кореляції 0,73 між результатами тестування та фактичним вибором професії протягом п'ятирічного періоду спостережень. MyPlan.com показує помірну точність з коефіцієнтом валідності 0,61 завдяки комбінуванню різних типів оцінок. 16Personalities має нижчі показники валідності (0,48) через обмеження дихотомічного підходу до класифікації особистості. Українські системи не мають документованих досліджень валідності, що ускладнює оцінку їх ефективності.

Користувацький досвід сучасних платформ характеризується зосередженням на інтуїтивності інтерфейсу та візуальній привабливості результатів. O*NET Interest Profiler пропонує найбільш структурований підхід з чіткими інструкціями та поетапним проходженням тесту. 16Personalities вирізняється високоякісною візуалізацією результатів з детальними описами типів особистості та інфографікою. LinkedIn Career Advice забезпечує найкращу інтеграцію з реальними даними ринку праці через зв'язок з професійною мережею. Українські рішення мають суттєві недоліки в юзабіліті через застарілий дизайн інтерфейсів та обмежені можливості візуалізації результатів.

Актуальність даних про професії є важливим фактором ефективності профорієнтаційних систем в умовах динамічного розвитку ринку праці. O*NET забезпечує найвищий рівень актуальності завдяки регулярному оновленню професійних профілів експертними групами та інтеграції з національною статистикою зайнято-

сті США. LinkedIn Career Advice використовує реальні дані вакансій для відображення поточного попиту на різні спеціальності, що забезпечує високу релевантність рекомендацій. MyPlan.com оновлює дані професій щорічно на основі бюро статистики праці США. Українські системи страждають від значного відставання в актуалізації даних, особливо щодо ІТ-професій та цифрової економіки.

Аналіз наявних систем виявляє кілька ключових напрямків для розвитку профорієнтаційних технологій. По-перше, відсутність персоналізованих алгоритмів машинного навчання в більшості систем обмежує їх здатність до адаптації під індивідуальні особливості користувачів та динаміку ринку праці. По-друге, недостатня інтеграція з реальними даними вакансій знижує практичну цінність рекомендацій для прийняття кар'єрних рішень. По-третє, відсутність мультифакторного підходу до оцінювання професійної придатності, що враховував би не лише інтереси, але й здібності, цінності та обмеження користувачів.

Українському ринку профорієнтаційних систем бракує сучасних рішень, що враховували б специфіку національної економіки та освітньої системи. Наявні платформи не забезпечують адекватного покриття ІТ-професій, що становлять значну частину сучасного ринку праці України. Відсутність інтеграції з даними про заробітні плати та перспективи кар'єрного зростання в українських реаліях ускладнює процес прийняття обґрунтованих професійних рішень.

Результати аналізу наявних профорієнтаційних систем узагальнено в таблиці 1.1, що дозволяє систематично порівняти ключові характеристики платформ за основними критеріями ефективності.

Аналіз таблиці виявляє суттєві переваги зарубіжних систем над українськими рішеннями в усіх ключових параметрах. Найвищу валідність демонструє O*NET Interest Profiler завдяки науково обґрунтованій методології та регулярній верифікації результатів. LinkedIn Career Advice вирізняється інноваційним використанням машинного навчання та інтеграцією з актуальними даними ринку праці, проте не має документованих досліджень валідності.

Таблиця 1.1 - Порівняльний аналіз наявних систем професійної орієнтації

Критерій	O*NET Interest Profiler	16Personalities	LinkedIn Career	MyPlan.com	Work.ua Тест	Osvita.ua	JobTest.ua
Теоретична основа	Теорія Холланда	Майєрс-Бріггс	Колаборативна фільтрація	Мультифакторна	Адаптація Клімова	Академічні інтереси	Змішаний підхід
Кількість питань	60	93	Аналіз профілю	4 тести × 25-40	30	60	20+ тестів
Алгоритм рекомендацій	Евклідова відстань	Детерміністичний	Машинне навчання	Зважене комбінування	Підрахунок балів	Підрахунок балів	Модульний підхід
База професій	900+	300+	Динамічна	1000+	150	200	250
Валідність результатів	0,73	0,48	Не документована	0,61	Не документована	Не документована	Не документована
Актуальність даних	Щорічне оновлення	Статична база	Реальний час	Щорічне оновлення	Рідкісні оновлення	Рідкісні оновлення	Статична база
Інтеграція з ринком праці	Національна статистика США	Відсутня	Вакансії LinkedIn	Бюро статистики США	Обмежена	Вступна статистика ВНЗ	Відсутня
Мобільна версія	Адаптивний дизайн	Повна підтримка	Нативний додаток	Адаптивний дизайн	Базова підтримка	Відсутня	Базова підтримка
Персоналізація	Низька	Середня	Висока	Середня	Відсутня	Відсутня	Низька
Вартість використання	Безкоштовно	Freemium модель	Безкоштовно	Платна	Безкоштовно	Безкоштовно	Freemium модель

Українські системи характеризуються спільними недоліками: застарілими теоретичними основами, обмеженими базами професій, відсутністю валідаційних досліджень та слабкою інтеграцією з реальними даними ринку праці. Особливо критичним є відставання в покритті IT-професій та цифрових спеціальностей, що становлять значну частину сучасного українського ринку праці.

Технологічні можливості сучасних веб-платформ та алгоритмів машинного навчання створюють передумови для розробки нового покоління профорієнтацій-

них систем, що поєднують наукову обґрунтованість традиційних підходів з адаптивністю та персоналізацією інтелектуальних алгоритмів. Такі системи мають потенціал для суттєвого підвищення точності професійних рекомендацій та покращення користувацького досвіду в процесі професійного самовизначення.

1.3 Аналіз рекомендаційних систем

Рекомендаційні системи є ключовим компонентом сучасних інформаційних платформ, що забезпечують персоналізацію контенту та послуг відповідно до потреб та уподобань користувачів. У контексті професійної орієнтації такі системи покликані аналізувати характеристики індивідуумів та відповідні вимоги професій для генерації релевантних рекомендацій щодо кар'єрного розвитку.

Сучасна таксономія рекомендаційних систем розрізняє три основні підходи до генерації рекомендацій, кожен з яких має унікальні переваги та обмеження для застосування в задачах професійної орієнтації.

Колаборативна фільтрація (Collaborative Filtering) базується на припущенні, що користувачі з подібними уподобаннями матимуть схожі оцінки невідомих об'єктів. Цей підхід використовує історію взаємодій користувачів з системою для виявлення груп індивідуумів зі схожими професійними інтересами. Математично колаборативна фільтрація може бути представлена як задача заповнення матриці взаємодій користувач-об'єкт, де відсутні значення прогножуються на основі відомих оцінок.

Контент-орієнтована фільтрація (Content-Based Filtering) аналізує характеристики об'єктів рекомендації та порівнює їх з профілем користувача. У професійній орієнтації цей підхід передбачає створення векторного представлення професій через їх ключові атрибути та співставлення з векторизованим профілем інтересів та навичок користувача. Рекомендації генеруються на основі метрик подібності між векторами користувача та професій [21, 23, 27].

Гібридні системи поєднують переваги різних підходів для подолання обмежень окремих методів. Комбінування може здійснюватися через зважене усереднення, каскадне застосування алгоритмів або паралельне обчислення рекомендацій з подальшою інтеграцією результатів.

Колаборативна фільтрація демонструє значну ефективність у випадках з великими обсягами користувальницьких даних та встановленими патернами взаємодій. Основною перевагою цього підходу є здатність виявляти неочевідні кореляції між професійними уподобаннями різних користувачів, що може призводити до несподіваних, але релевантних рекомендацій професій.

User-based колаборативна фільтрація знаходить користувачів зі схожими професійними інтересами та рекомендує професії, які були позитивно оцінені подібними особами. Математично це виражається через обчислення коефіцієнтів подібності між користувачами та прогнозування оцінок за формулою:

$$P(u,i) = \mu_u + \frac{\sum(\text{sim}(u,v) \times (r_{v,i} - \mu_v))}{\sum|\text{sim}(u,v)|},$$

де $P(u,i)$ – прогнозована оцінка користувача u для професії i ;

μ_u – середня оцінка користувача u , $\text{sim}(u,v)$ – подібність між користувачами;

u та v , $r_{v,i}$ – оцінка користувача v для професії i .

Item-based колаборативна фільтрація фокусується на відношеннях між професіями, рекомендуючи спеціальності, схожі на ті, які користувач уже оцінив позитивно. Цей підхід особливо ефективний у стабільних доменах, де характеристики професій змінюються повільніше за користувацькі уподобання.

Основними недоліками колаборативної фільтрації є проблема холодного старту для нових користувачів та професій, розрідженість матриці взаємодій у специфічних доменах як професійна орієнтація, та схильність до популярного зсуву, коли система переважно рекомендує найбільш популярні професії. Додатковою проблемою є масштабованість алгоритмів при зростанні кількості користувачів та об'єктів.

Контент-орієнтований підхід заснований на аналізі внутрішніх характеристик об'єктів рекомендації та їх порівнянні з профілем користувача. У контексті професійної орієнтації цей метод передбачає структуровану репрезентацію професій через набір атрибутів та створення аналогічного профілю для користувача на основі результатів тестування.

Векторне представлення професій формується через експертну оцінку або статистичний аналіз вимог до кожної спеціальності. Типовий вектор професії може включати такі компоненти як технічна складність, рівень соціальної взаємодії, креативність, аналітичне навантаження, та бізнес-орієнтованість. Кожна характеристика квантифікується в діапазоні від 0 до 1, створюючи багатовимірний простір професійних атрибутів.

Профіль користувача генерується на основі результатів психометричного тестування, де відповіді конвертуються у векторне представлення з тими самими розмірностями, що й професійні профілі. Процес конвертації включає зважування відповідей відповідно до їх значущості для кожної характеристики та нормалізацію результуючого вектора.

Метрики подібності використовуються для кількісної оцінки відповідності між профілем користувача та характеристиками професій. Найбільш поширеними метриками є косинусна подібність, евклідова відстань та коефіцієнт кореляції Пірсона. Косинусна подібність обчислюється за формулою:

$$\cos(\theta) = (A \cdot B) / (\|A\| \times \|B\|),$$

де A та B – вектори користувача та професії відповідно;

θ – кут між векторами в багатовимірному просторі.

Перевагами контент-орієнтованого підходу є незалежність від даних інших користувачів, можливість рекомендувати нові професії одразу після їх додавання до системи, та здатність пояснювати логіку рекомендацій через аналіз співпадінь

характеристик. Система також не страждає від проблеми холодного старту для нових користувачів, оскільки рекомендації базуються виключно на результатах їх тестування.

Основними обмеженнями є схильність до надмірної спеціалізації (overspecialization), коли система рекомендує лише професії, дуже схожі на вже відомі користувачеві, та залежність від якості характеристикації професій експертами. Відсутність серендипітності може обмежувати користувачів у дослідженні альтернативних кар'єрних шляхів.

Гібридні системи представляють найбільш перспективний напрямок розвитку рекомендаційних технологій, оскільки дозволяють компенсувати недоліки окремих підходів через їх синергетичне поєднання. У професійній орієнтації гібридизація особливо актуальна через необхідність балансу між точністю рекомендацій та їх різноманітністю.

Зважене комбінування є найпростішим методом гібридизації, де фінальна оцінка професії обчислюється як зважена сума оцінок від різних алгоритмів:

$$H(u,i) = \alpha \times CBF(u,i) + \beta \times CF(u,i) + \gamma \times Pop(i),$$

де $H(u,i)$ – гібридна оцінка;

CBF – оцінка від контент-орієнтованої фільтрації;

CF – оцінка від колаборативної фільтрації;

Pop – популярність професії;

α, β, γ – ваги компонентів з умовою $\alpha + \beta + \gamma = 1$.

Змішаний підхід представляє рекомендації від різних алгоритмів одночасно, дозволяючи користувачеві побачити альтернативні погляди на професійний вибір. Цей метод особливо корисний для освітніх цілей, коли важливо продемонструвати різні аспекти кар'єрного планування.

Перемикальний підхід використовує різні алгоритми залежно від ситуації або характеристик користувача. Наприклад, для нових користувачів може застосовуватися контент-орієнтована фільтрація, тоді як для досвідчених користувачів з багатою історією взаємодій – колаборативна фільтрація.

Каскадний підхід послідовно застосовує алгоритми, де результати першого етапу використовуються як вхідні дані для наступного. Типова реалізація може включати первинну фільтрацію професій за контент-орієнтованими критеріями з подальшим ранжуванням за колаборативними метриками.

Оцінювання якості рекомендаційних систем у професійній орієнтації потребує комплексного підходу, що враховує не лише технічну точність алгоритмів, але й їх практичну корисність для користувачів.

Точність (Precision) вимірює частку релевантних рекомендацій серед усіх наданих системою:

$$\text{Precision} = |\{\text{релевантні рекомендації}\} \cap \{\text{надані рекомендації}\}| / |\{\text{надані рекомендації}\}|$$

У контексті професійної орієнтації релевантність може визначатися через довгострокове задоволення користувачів обраними професіями або їх відповідність психометричним профілям.

Повнота (Recall) показує частку знайдених релевантних професій серед усіх потенційно підходящих:

$$\text{Recall} = |\{\text{релевантні рекомендації}\} \cap \{\text{надані рекомендації}\}| / |\{\text{усі релевантні професії}\}|$$

F1-міра представляє гармонійне середнє між точністю та повнотою:

$$F1 = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

Різноманітність (Diversity) рекомендацій особливо важлива в професійній орієнтації, оскільки обмежений набір схожих професій може звузити кар'єрні горизонти користувачів. Різноманітність може вимірюватися через середню відстань між рекомендованими професіями у просторі характеристик.

Новизна (Novelty) оцінює здатність системи рекомендувати професії, які користувач не розглядав раніше, що сприяє розширенню професійних горизонтів.

Серендипітність (Serendipity) характеризує ймовірність несподіваних, але корисних рекомендацій, що можуть відкрити нові кар'єрні можливості.

Ефективна система професійної орієнтації має демонструвати збалансовані показники за всіма метриками, забезпечуючи високу точність рекомендацій при збереженні їх різноманітності та потенціалу для відкриття нових професійних напрямків. Особливо важливим є довгострокове валідування результатів через відстеження кар'єрних шляхів користувачів та їх задоволеності професійними рішеннями.

1.4 Постановка задачі дослідження

Проведений аналіз сучасного стану професійної орієнтації, наявних систем та теоретичних основ рекомендаційних алгоритмів дозволяє сформулювати комплексну науково-технічну задачу розробки інноваційної інформаційної платформи професійної орієнтації з використанням методів машинного навчання.

Центральна проблема дослідження полягає в невідповідності наявних методів професійної орієнтації динамічним потребам сучасного ринку праці та очікуванням молодого покоління щодо персоналізації і актуальності професійних рекомендацій. Традиційні підходи, засновані на статичних опитувальниках і детерміністичних алгоритмах співставлення, не здатні адекватно враховувати складність індивідуальних професійних профілів та мінливість характеристик сучасних професій.

Математично проблему можна представити як задачу оптимізації функції корисності $U(u,p)$, що визначає відповідність користувача u професії p у багатовимірному просторі характеристик. Наявні системи використовують спрощені функції корисності, які не враховують нелінійні взаємозв'язки між професійними атрибутами та індивідуальними особливостями користувачів.

Додатковою проблемою є відсутність інтеграції з актуальними даними ринку праці, що призводить до рекомендацій професій з низьким попитом або застарілими вимогами до компетенцій. Це особливо критично в умовах швидкої цифровізації економіки та появи нових професійних ролей, які не представлені в традиційних класифікаціях.

Розроблювана система має відповідати сучасним стандартам веб-додатків та забезпечувати високий рівень користувацького досвіду через інтуїтивний інтерфейс та швидкодію алгоритмів рекомендацій.

Архітектурні вимоги передбачають модульну побудову системи з чіткою сепарацією бізнес-логіки, рівня даних та представлення. Система має бути реалізована як веб-додаток з клієнт-серверною архітектурою, що забезпечує доступність з різних пристроїв та операційних систем. Backend має підтримувати REST API для інтеграції з зовнішніми сервісами та можливого розширення функціоналу в майбутньому.

Продуктивність системи має забезпечувати обробку запитів користувачів з часом відгуку не більше трьох секунд для генерації рекомендацій та не більше однієї секунди для завантаження інтерфейсних елементів. Система повинна підтримувати одночасну роботу до ста користувачів без суттєвого зниження швидкодії.

Масштабованість реалізується через горизонтальне розгортання компонентів системи та використання кешування для часто запитуваних даних. База даних має бути спроектована з урахуванням можливості збільшення кількості професій до тисячі записів та користувачів до десяти тисяч без реструктуризації схеми.

Безпека даних забезпечується через шифрування персональних даних користувачів, захист від поширених веб-вразливостей та імплементацію системи аудиту

дій користувачів. Результати тестування мають зберігатися в анонімізованому вигляді для забезпечення приватності учасників.

Підсистема тестування має реалізовувати валідований психометричний інструментарій для оцінки професійних інтересів, навичок та особистісних характеристик користувачів. Тест повинен включати від п'ятнадцяти до двадцяти питань, структурованих за п'ятьма основними категоріями професійних атрибутів. Інтерфейс тестування має забезпечувати інтуїтивну навігацію з можливістю перегляду прогресу виконання та повернення до попередніх питань для коригування відповідей.

Алгоритм рекомендацій на основі Content-Based Filtering [5, 13] має генерувати персоналізовані рекомендації професій з ранжуванням за рівнем відповідності користувацькому профілю. Система повинна надавати не менше п'яти рекомендацій з поясненнями щодо логіки вибору кожної професії та ключових факторів відповідності.

База знань про професії має містити структуровані описи не менше двадцяти сучасних професій з різних галузей економіки, включаючи ІТ-сферу, креативні індустрії, бізнес та традиційні спеціальності. Кожна професія характеризується вектором атрибутів за п'ятьма основними вимірами та містить інформацію про середні зарплати, вимоги до освіти та ключові навички. Аналітичний модуль забезпечує візуалізацію результатів тестування користувача, статистики популярності професій та трендів використання платформи. Система має генерувати інтерактивні графіки та діаграми для наочного представлення даних про професійний профіль користувача та характеристики рекомендованих професій.

Інтерфейс користувача реалізується як адаптивний веб-додаток з підтримкою мобільних пристроїв та планшетів. Дизайн має відповідати сучасним принципам UX/UI з акцентом на зручність навігації та візуальну привабливість. Особлива увага приділяється доступності інтерфейсу для користувачів з обмеженими можливостями.

Надійність системи забезпечується через резервне копіювання даних, логування системних подій та обробку виняткових ситуацій без втрати користувацьких

даних. Система має підтримувати коефіцієнт доступності не менше 99% протягом робочого часу.

Підтримка браузерів має охоплювати всі сучасні веб-браузери включно з Chrome, Firefox, Safari та Edge у їх актуальних версіях. Функціональність системи не повинна залежати від специфічних розширень або плагінів браузера.

Локалізація реалізується для української мови з можливістю розширення на інші мови в майбутньому. Всі текстові елементи інтерфейсу, описи професій та питання тестування мають бути україномовними з дотриманням термінологічних стандартів.

Сумісність з мобільними пристроями досягається через респонсивний дизайн з адаптацією інтерфейсу до різних розмірів екранів. Функціональність системи на мобільних пристроях не повинна бути обмеженою порівняно з десктопною версією.

Дослідження здійснюється в рамках обмежених часових та ресурсних можливостей магістерського проєкту, що накладає певні обмеження на масштаб та складність системи.

Часові обмеження визначають необхідність фокусування на створенні функціонального прототипу з базовим набором професій та спрощеним алгоритмом Content-Based рекомендацій. Розробка складних гібридних алгоритмів або інтеграція з зовнішніми API залишається за межами поточного дослідження. Ресурсні обмеження передбачають використання доступних технологій з відкритим кодом та безкоштовних хмарних сервісів для розгортання системи. Система проєктується для роботи на стандартному веб-хостингу без потреби в дорогому апаратному забезпеченні. Методологічні допущення включають використання експертних оцінок для характеристики професій у відсутності великих масивів емпіричних даних. Валідація алгоритму рекомендацій здійснюється на обмеженій вибірці користувачів через короткий термін дослідження. Технологічні обмеження визначають вибір перевірених та стабільних технологічних рішень замість експериментальних підходів. Система використовує традиційні реляційні бази даних та поширені веб-фреймворки для забезпечення надійності та можливості майбутньої підтримки.

Наукова новизна дослідження полягає в адаптації Content-Based алгоритмів рекомендацій для специфічних задач професійної орієнтації з урахуванням психометричних особливостей користувачів та багатовимірності професійних характеристик. Розроблені методи векторизації результатів психологічного тестування та професійних профілів можуть бути використані для подальших досліджень в галузі комп'ютерної психології та рекомендаційних систем.

Практична значущість роботи визначається можливістю впровадження розробленої системи в навчальних закладах, центрах зайнятості та консалтингових компаніях для покращення якості професійного консультування. Система може стати інструментом для дослідження тенденцій професійного самовизначення молоді та оптимізації освітніх програм відповідно до потреб ринку праці.

Методологічний внесок полягає у розробці комплексного підходу до оцінювання ефективності рекомендаційних алгоритмів у сфері професійної орієнтації, що включає як технічні метрики точності та швидкодії, так і показники користувацького задоволення та довгострокової валідності рекомендацій.

1.5 Висновки до розділу 1

У першому розділі здійснено комплексний аналіз предметної області професійної орієнтації та обґрунтовано актуальність розробки інноваційної інформаційної системи з використанням методів машинного навчання.

Аналіз сучасного стану професійної орієнтації виявив критичні недоліки традиційних підходів. Класична система професійної орієнтації в Україні, сформована в радянський період, базується на статичних психодіагностичних методиках, які не враховують сучасні професії в галузі інформаційних технологій. Статистичні дані свідчать про зростання попиту на IT-спеціалістів на 340% за період 2018-2023 років, тоді як традиційні профорієнтаційні методики містять менше 3% професій IT-сфери. Встановлено, що лише 23% випускників українських ВНЗ працюють за

отриманою спеціальністю протягом перших п'яти років після закінчення навчання, що підтверджує низьку ефективність існуючих систем професійного консультування.

Огляд наявних систем професійної орієнтації дозволив виявити найкращі світові практики та ідентифікувати прогалини у вітчизняних рішеннях. Порівняльний аналіз показав, що зарубіжні платформи (O*NET Interest Profiler, 16Personalities, LinkedIn Career Advice) демонструють високі показники валідності (0,48-0,73) завдяки науково обґрунтованим методологіям та регулярному оновленню даних. Українські системи характеризуються застарілими теоретичними основами, обмеженими базами професій (150-250 записів проти 900+ у зарубіжних аналогах), відсутністю валідаційних досліджень та слабкою інтеграцією з реальними даними ринку праці.

Аналіз рекомендаційних систем виявив три основні підходи до генерації професійних рекомендацій: колаборативну фільтрацію, контент-орієнтовану фільтрацію та гібридні методи. Встановлено, що контент-орієнтований підхід є оптимальним для систем професійної орієнтації через незалежність від історії взаємодій користувачів, можливість пояснення логіки рекомендацій та відсутність проблеми холодного старту. Визначено ключові метрики оцінювання ефективності рекомендацій: точність, повноту, F1-міру, різноманітність та серендіпність.

На основі проведеного аналізу сформульовано комплексну науково-технічну задачу розробки веб-платформи професійної орієнтації з алгоритмом Content-Based рекомендацій. Визначено функціональні вимоги до системи: підсистема психометричного тестування з 15-20 питаннями за п'ятьма категоріями професійних атрибутів, база знань не менше 20 сучасних професій з структурованими описами, алгоритм рекомендацій з ранжуванням за рівнем відповідності та поясненням логіки вибору. Встановлено технічні вимоги: модульна клієнт-серверна архітектура, час відгуку не більше 3 секунд для генерації рекомендацій, підтримка одночасної роботи до 100 користувачів, адаптивний інтерфейс з підтримкою мобільних пристроїв.

Результати аналізу підтверджують необхідність розробки інноваційної системи професійної орієнтації, яка поєднає наукову обґрунтованість традиційних психометричних підходів з адаптивністю та персоналізацією сучасних рекомендаційних алгоритмів. Створення такої системи дозволить підвищити точність професійних рекомендацій, забезпечити актуальність даних про ринок праці та покращити користувацький досвід у процесі професійного самовизначення молоді.

2 ДОСЛІДЖЕННЯ ТА ПОРІВНЯЛЬНИЙ АНАЛІЗ АЛГОРИТМІВ РЕКОМЕНДАЦІЙ

2.1 Теоретичні основи Content-Based Filtering

Content-Based Filtering представляє фундаментальний підхід до створення рекомендаційних систем, що базується на аналізі внутрішніх характеристик об'єктів рекомендації та їх співставленні з профілем користувача. Теоретичні основи цього методу спираються на концепції лінійної алгебри, теорії інформаційного пошуку та машинного навчання, що забезпечує математично обґрунтоване підґрунтя для генерації персоналізованих рекомендацій.

Центральна концепція Content-Based підходу полягає у векторному представленні користувачів та об'єктів у спільному багатовимірному просторі характеристик. Кожен об'єкт рекомендації представляється як вектор ознак у n -вимірному просторі, де кожна координата відповідає певній характеристиці об'єкта.

Формально, об'єкт i може бути представлений як вектор $v_i = (v_{i1}, v_{i2}, \dots, v_{in})$, де v_{ij} є значенням j -ї характеристики для об'єкта i . У контексті професійної орієнтації такими характеристиками можуть бути рівень технічної складності, креативність, соціальна взаємодія, аналітичне навантаження та бізнес-орієнтованість професії.

Аналогічно, профіль користувача u представляється вектором $p_u = (p_{u1}, p_{u2}, \dots, p_{un})$ у тому самому просторі ознак. Цей вектор формується на основі явних або неявних уподобань користувача, отриманих через тестування, попередні оцінки чи історію взаємодій з системою.

Математична модель передбачає, що ймовірність релевантності об'єкта i для користувача u є функцією подібності між векторами p_u та v_i . Така модель базується на припущенні, що користувачі віддають перевагу об'єктам з характеристиками, схожими на їхні власні уподобання.

Векторне представлення об'єктів може формуватися різними способами залежно від природи характеристик та доступних даних. У випадку категоріальних ознак використовується бінарне кодування, де наявність характеристики позначається одиницею, а відсутність нулем. Для числових характеристик застосовується пряма нормалізація значень до діапазону $[0,1]$ або стандартизація з нульовим середнім та одиничною дисперсією.

Профіль користувача може формуватися через явні та неявні методи збору уподобань. Явні методи включають анкетування, рейтингові оцінки та результати психометричного тестування. Неявні методи аналізують поведінкові патерни користувача, такі як час, проведений на сторінках з описами професій, кількість переглядів певних категорій або частота звернень до конкретної інформації.

Математично, явний профіль користувача може бути сформований як зважена сума векторів об'єктів, з якими користувач взаємодіяв позитивно:

$$p_u = \sum(r_{ui} \times v_i) / \sum(r_{ui}),$$

де r_{ui} – оцінка користувача u для об'єкта i .

Сума береться по всіх об'єктах, оцінених користувачем.

Для неявного профілювання використовуються методи машинного навчання, такі як логістична регресія або нейронні мережі, які навчаються на поведінкових даних користувача для прогнозування його уподобань у векторному просторі характеристик.

Вибір відповідної метрики подібності є критичним для ефективності Content-Based системи. Найпоширенішими метриками є косинусна подібність, евклідова відстань та коефіцієнт кореляції Пірсона, кожна з яких має специфічні властивості та області застосування.

Косинусна подібність вимірює кут між двома векторами у багатовимірному просторі та є інваріантною до масштабу векторів. Це робить її особливо корисною для порівняння об'єктів з різними діапазонами значень характеристик:

$$\text{sim_cos}(p_u, v_i) = (p_u \cdot v_i) / (\|p_u\| \times \|v_i\|),$$

де $p_u \cdot v_i$ – скалярний добуток векторів;

$\|p_u\|$ та $\|v_i\|$ – їх евклідові нормами.

Евклідова відстань обчислює геометричну відстань між векторами у просторі ознак та є чутливою до абсолютних різниць між характеристиками:

$$\text{sim_eucl}(p_u, v_i) = 1 / (1 + \sqrt{\sum (p_{uj} - v_{ij})^2}).$$

Ця метрика доцільна, коли важливі абсолютні значення характеристик, а не лише їх пропорційні співвідношення.

Коефіцієнт кореляції Пірсона вимірює лінійну залежність між векторами після центрування їх відносно середніх значень:

$$\text{sim_pearson}(p_u, v_i) = \sum ((p_{uj} - \mu_u)(v_{ij} - \mu_i)) / (\sigma_u \times \sigma_i \times \sqrt{n}),$$

де μ_u, μ_i – середні значення векторів,

σ_u, σ_i – їх стандартні відхилення.

У професійній орієнтації косинусна подібність часто є оптимальним вибором через її здатність ігнорувати різниці в абсолютних значеннях характеристик та зосереджуватися на структурі переваг користувача.

Традиційно розроблена для інформаційного пошуку, схема важення TF-IDF (Term Frequency-Inverse Document Frequency) знаходить застосування у Content-Based системах для виявлення найбільш дискримінативних характеристик об'єктів.

Term Frequency (TF) вимірює важливість характеристики для конкретного об'єкта:

$$TF(t,d) = f(t,d) / \max \{f(w,d) : w \in d\},$$

де $f(t,d)$ – частота терміну t у документі d .

Inverse Document Frequency (IDF) зменшує важливість характеристик, що зустрічаються у багатьох об'єктах:

$$IDF(t,D) = \log(|D| / |\{d \in D : t \in d\}|),$$

де D – множина всіх документів;

$|\{d \in D : t \in d\}|$ – кількість документів, що містять термін t .

Комбінована TF-IDF вага обчислюється як добуток цих компонентів:

$$w(t,d,D) = TF(t,d) \times IDF(t,D).$$

У контексті професійної орієнтації TF-IDF дозволяє виявляти унікальні характеристики професій, що не є загальними для більшості спеціальностей. Наприклад, висока креативність може бути рідкісною характеристикою, що робить її більш значущою для розрізнення між професіями.

Адаптація TF-IDF для професійної орієнтації передбачає розгляд кожної професії як документа, а характеристик як термінів. Частота терміну відповідає інтенсивності характеристики у професії, тоді як обернена частота документа відображає рідкість цієї характеристики серед усіх професій.

Content-Based Filtering має ряд теоретичних переваг, що роблять його привабливим для застосування у професійній орієнтації. По-перше, метод не страждає

від проблеми холодного старту для нових користувачів, оскільки рекомендації можуть генеруватися відразу після отримання їх профілю через тестування. По-друге, система здатна рекомендувати нові об'єкти одразу після їх додавання до бази знань без необхідності накопичення користувацьких оцінок.

Важливою перевагою є також можливість пояснення рекомендацій через аналіз співпадінь характеристик між профілем користувача та рекомендованими професіями. Це критично для професійної орієнтації, де користувачі потребують розуміння логіки рекомендацій для прийняття обґрунтованих кар'єрних рішень.

Теоретичні обмеження методу включають схильність до надмірної спеціалізації, коли система рекомендує лише професії, дуже схожі на вже відомі користувачеві. Це може обмежувати дослідження альтернативних кар'єрних шляхів та зменшувати серендипітність рекомендацій.

Додатковим обмеженням є залежність від якості та повноти опису характеристик об'єктів. Неточна або неповна характеристика професій може призводити до субоптимальних рекомендацій. Також метод має труднощі з моделюванням складних нелінійних взаємодій між характеристиками, що може бути важливим для точного відображення професійних уподобань.

Математична модель Content-Based Filtering припускає лінійність функції корисності користувача, що не завжди відповідає реальності. Користувачі можуть мати складні, нелінійні переваги щодо комбінацій характеристик професій, які важко моделювати через прості метрики подібності.

Теоретичні основи Content-Based Filtering забезпечують міцний математичний фундамент для розробки рекомендаційних систем у професійній орієнтації. Розуміння векторних моделей, метрик подібності та методів важення характеристик є критичним для ефективної адаптації цього підходу до специфічних потреб професійного консультування та забезпечення високої якості персоналізованих рекомендацій.

2.2 Адаптація Content-Based алгоритмів для професійної орієнтації

Застосування Content-Based підходу до задач професійної орієнтації потребує суттєвої адаптації класичних алгоритмів рекомендаційних систем для врахування специфічних особливостей предметної області. Професійне самовизначення характеризується складністю процесу прийняття рішень, довгостроковими наслідками вибору та необхідністю збалансування особистих інтересів з реальними можливостями ринку праці.

Ефективне векторне представлення професій вимагає ретельного відбору характеристик, які адекватно відображають сутність різних видів трудової діяльності та можуть бути кількісно оцінені. Базуючись на сучасних теоріях професійної психології та аналізі вимог роботодавців, розроблено п'ятивимірну модель професійних характеристик.

Технічна складність відображає рівень використання технічних знань, інструментів та методів у професійній діяльності. Для ІТ-професій цей показник наближається до одиниці, тоді як для гуманітарних спеціальностей може становити 0,1-0,3. Математично це можна виразити як:

$$T(p) = \alpha_1 \times CS(p) + \alpha_2 \times TT(p) + \alpha_3 \times PM(p)$$

де $CS(p)$ – рівень використання комп'ютерних систем,

$TT(p)$ – залежність від технічних інструментів,

$PM(p)$ – складність процедур та методів;

$\alpha_1, \alpha_2, \alpha_3$ – нормалізуючі коефіцієнти.

Креативна складова характеризує ступінь творчого підходу, інноваційності та оригінальності у роботі. Дизайнерські професії мають високі значення цієї характеристики (0,8-0,9), тоді як стандартизовані спеціальності демонструють нижчі показники (0,2-0,4). Формально:

$$C(p) = \beta_1 \times OR(p) + \beta_2 \times IN(p) + \beta_3 \times AE(p)$$

де $OR(p)$ – потреба в оригінальних рішеннях;

$IN(p)$ – рівень інноваційності;

$AE(p)$ – естетична складова діяльності.

Соціальна взаємодія вимірює інтенсивність міжособистісних контактів та значущість комунікативних навичок для успішного виконання професійних обов'язків. Педагогічні та консалтингові професії характеризуються високими значеннями (0,7-0,9), тоді як дослідницькі спеціальності можуть мати нижчі показники (0,2-0,5):

$$S(p) = \gamma_1 \times CC(p) + \gamma_2 \times TW(p) + \gamma_3 \times CL(p),$$

де $CC(p)$ – інтенсивність клієнтських контактів,

$TW(p)$ – рівень командної роботи,

$CL(p)$ – вимоги до лідерських якостей.

Аналітична складність відображає потребу в логічному мисленні, аналізі даних та розв'язанні складних проблем. Професії у сфері фінансів, науки та інженерії демонструють високі значення цього показника:

$$A(p) = \delta_1 \times PS(p) + \delta_2 \times DA(p) + \delta_3 \times LR(p)$$

де $PS(p)$ – складність розв'язуваних проблем;

$DA(p)$ – обсяг аналізованих даних;

$LR(p)$ – потреба в логічних міркуваннях.

Бізнес-орієнтованість характеризує ступінь залученості у комерційні процеси, фінансове планування та стратегічне мислення. Управлінські та підприємницькі професії мають високі значення цієї характеристики:

$$B(p) = \varepsilon_1 \times FM(p) + \varepsilon_2 \times SP(p) + \varepsilon_3 \times MM(p),$$

де $FM(p)$ – рівень фінансової відповідальності,

$SP(p)$ – участь у стратегічному плануванні,

$MM(p)$ – вимоги до менеджерських навичок.

Трансформація якісних відповідей користувачів у кількісні векторні представлення є критичним етапом адаптації Content-Based алгоритмів для професійної орієнтації. Процес векторизації має забезпечувати максимальне збереження інформації про індивідуальні особливості користувача при переході до формалізованого математичного опису.

Структура тестування включає п'ять блоків питань, кожен з яких спрямований на оцінку однієї з професійних характеристик. Питання формулюються у вигляді ситуаційних завдань або запитань про уподобання з чотирма варіантами відповідей, що дозволяє отримати градуйовану оцінку інтенсивності кожної характеристики. Система ваг відповідей розроблена на основі експертних оцінок професійних психологів та валідована через пілотні дослідження. Кожен варіант відповіді має заздалегідь визначені ваги для всіх п'яти характеристик, що дозволяє одному питанню впливати на кілька вимірів користувацького профілю:

$$W(q,a) = [w_1(q,a), w_2(q,a), w_3(q,a), w_4(q,a), w_5(q,a)],$$

де $W(q,a)$ – вектор ваг для відповіді a на питання q ;

$w_i(q,a)$ – вага впливу на i -ту характеристику.

Агрегація відповідей здійснюється через зважене підсумовування ваг усіх відповідей користувача з подальшою нормалізацією результуючого вектора:

$$U = \text{normalize}(\sum W(q_i, a_i)),$$

де U – фінальний вектор користувача;

q_i – i -те питання;

a_i – відповідь користувача на i -те питання.

Нормалізація профілю забезпечує приведення всіх компонентів вектора до діапазону $[0,1]$ та збалансування впливу різних характеристик на процес генерації рекомендацій. Використовується лінійна нормалізація з урахуванням теоретично можливих мінімальних та максимальних значень кожної характеристики.

Точність Content-Based алгоритму критично залежить від якості системи ваг, що визначають відносну важливість різних професійних характеристик для кожної спеціальності. Розробка такої системи потребує поєднання експертних знань, емпіричних даних та валідації через практичні тестування.

Експертне оцінювання здійснюється групою професійних консультантів, HR-спеціалістів та представників різних галузей економіки. Кожен експерт незалежно оцінює професії за п'ятибальною шкалою за кожною характеристикою, після чого результати агрегуються через медіанні значення для зменшення впливу екстремальних оцінок. Аналіз вакансій доповнює експертні оцінки емпіричними даними з реальних оголошень про роботу. Застосовуються методи обробки природної мови для автоматичного виявлення ключових слів та фраз, що вказують на важливість певних характеристик. Частота згадувань технічних навичок, креативних здібностей, комунікативних вимог та інших характеристик кореспондується з векторним представленням професій.

Статистична валідація системи ваг здійснюється через порівняння генерованих рекомендацій з реальними професійними виборами користувачів. Коефіцієнт кореляції між передбаченнями алгоритму та фактичними рішеннями використовується як метрика якості системи ваг.

Адаптивність системи забезпечується через можливість оновлення ваг на основі накопичуваної статистики взаємодій користувачів з системою. Алгоритми машинного навчання можуть автоматично коригувати ваги для покращення точності рекомендацій:

$$w'_{ij} = w_{ij} + \eta \times \partial L / \partial w_{ij},$$

де w'_{ij} – оновлена вага j -ї характеристики для i -ї професії;

η – швидкість навчання;

L – функція втрат.

Професійна орієнтація не може обмежуватися лише аналізом індивідуальних характеристик та професійних вимог. Контекстуальні фактори, такі як ринкові умови, географічні обмеження та соціально-економічні чинники, суттєво впливають на доцільність професійних рекомендацій.

Ринкова привабливість професій інтегрується в алгоритм рекомендацій через коефіцієнт попиту, що базується на аналізі статистики вакансій, рівня зарплат та прогнозах розвитку галузей:

$$M(p) = \lambda_1 \times VD(p) + \lambda_2 \times SR(p) + \lambda_3 \times GF(p),$$

де $VD(p)$ – щільність вакансій;

$SR(p)$ – середній рівень зарплат;

$GF(p)$ – прогноз зростання галузі.

Географічні обмеження враховують доступність професій у конкретному регіоні та готовність користувача до релокації. Система може адаптувати рекомендації відповідно до локального ринку праці або розширювати географію пошуку за потреби. Освітні вимоги співставляються з наявним рівнем освіти користувача та його готовністю до додаткового навчання. Алгоритм може пріоритизувати професії, доступні з поточним рівнем кваліфікації, або запропонувати траєкторії професійного розвитку для досягнення бажаних спеціальностей.

Стандартні Content-Based алгоритми потребують модифікації для ефективного застосування у професійній орієнтації через специфічні особливості цієї предметної області.

Багаторівневе ранжування рекомендацій враховує не лише подібність характеристик, але й практичну доступність професій для користувача. Фінальна оцінка

професії обчислюється як зважена комбінація similarity score та контекстуальних факторів:

$$S(u,p) = \alpha \times \text{sim}(U,P) + \beta \times M(p) + \gamma \times A(u,p)$$

де $\text{sim}(U,P)$ – косинусна подібність між користувачем та професією;

$M(p)$ – ринкова привабливість;

$A(u,p)$ – доступність для користувача u .

Диверсифікація рекомендацій запобігає надмірній концентрації на схожих професіях через введення штрафу за близькість рекомендованих спеціальностей. Алгоритм максимізує покриття різних професійних сфер при збереженні високої релевантності рекомендацій. Пояснювальність результатів досягається через детальний аналіз факторів, що вплинули на кожну рекомендацію. Система генерує текстові пояснення, які допомагають користувачам зрозуміти логіку рекомендацій та прийняти обґрунтовані рішення про професійний розвиток.

Адаптація Content-Based алгоритмів для професійної орієнтації представляє комплексний процес, що поєднує математичне моделювання, психологічні інсайти та практичні вимоги ринку праці. Розроблена система забезпечує баланс між персоналізацією рекомендацій та їх практичною корисністю для користувачів у процесі професійного самовизначення.

2.3 Математична модель системи рекомендацій

Формалізація процесу генерації професійних рекомендацій потребує розробки математичної моделі, що адекватно описує взаємозв'язки між характеристиками користувачів, професійними вимогами та контекстуальними факторами професійного вибору. Математичний апарат моделі базується на теорії векторних просторів, методах лінійної алгебри та принципах багатокритеріальної оптимізації.

Задача професійної орієнтації може бути сформульована як проблема пошуку оптимального співставлення користувачів з професіями у багатовимірному просторі характеристик. Нехай $U = \{u_1, u_2, \dots, u_n\}$ є множиною користувачів системи, а $P = \{p_1, p_2, \dots, p_m\}$ множиною професій у базі знань системи.

Кожен користувач u_i характеризується вектором особистісних та професійних характеристик:

$$u_i = (u_i^1, u_i^2, u_i^3, u_i^4, u_i^5)^T,$$

де $u_i^j \in [0,1]$ представляє інтенсивність j -ї характеристики у користувача i .

Аналогічно, кожна професія p_j описується вектором вимог:

$$p_j = (p_j^1, p_j^2, p_j^3, p_j^4, p_j^5)^T,$$

де $p_j^k \in [0,1]$ означає важливість k -ї характеристики для професії j .

Центральною задачею є знаходження функції корисності $R: U \times P \rightarrow R$, що визначає ступінь відповідності користувача u_i професії p_j . Оптимальні рекомендації для користувача u_i формуються як топ- k професій з найвищими значеннями функції корисності:

$$\text{Rec}(u_i, k) = \text{argmax}_{\{p_j \in P\}} R(u_i, p_j).$$

Основою системи рекомендацій є функція подібності, що кількісно оцінює відповідність профілю користувача вимогам конкретної професії. В рамках Content-Based підходу найбільш ефективною виявилася косинусна подібність, адаптована для специфіки професійної орієнтації.

Базова косинусна подібність обчислюється як нормалізований скалярний добуток векторів користувача та професії:

$$\text{sim_cos}(u_i, p_j) = (u_i \cdot p_j) / (\|u_i\| \times \|p_j\|),$$

де $u_i \cdot p_j = \sum_{k=1}^5 u_i^k \times p_j^k$ – скалярний добуток;

а $\|u_i\| = \sqrt{(\sum_{k=1}^5 (u_i^k)^2)}$ – норма вектора.

Зважена косинусна подібність враховує різну важливість професійних характеристик через введення вагових коефіцієнтів w_k :

$$\text{sim_weighted}(u_i, p_j) = (\sum_{k=1}^5 w_k \times u_i^k \times p_j^k) / \sqrt{((\sum_{k=1}^5 w_k \times (u_i^k)^2) \times (\sum_{k=1}^5 w_k \times (p_j^k)^2))}.$$

Ваги w_k визначаються емпірично через аналіз важливості різних характеристик для точності рекомендацій.

Адаптивна функція подібності включає механізм саморегулювання на основі зворотного зв'язку користувачів:

$$\text{sim_adaptive}(u_i, p_j) = \alpha \times \text{sim_weighted}(u_i, p_j) + \beta \times \text{feedback_score}(u_i, p_j),$$

де feedback_score – історична оцінка якості рекомендацій аналогічних професій для подібних користувачів.

Генерація фінальних рекомендацій потребує ранжування всіх професій відповідно до їх релевантності для конкретного користувача. Алгоритм ранжування поєднує показники подібності з контекстуальними факторами для формування практично корисних рекомендацій.

Багатокритеріальна функція корисності інтегрує різні аспекти привабливості професії:

$$R(u_i, p_j) = \lambda_1 \times \text{sim}(u_i, p_j) + \lambda_2 \times \text{market}(p_j) + \lambda_3 \times \text{access}(u_i, p_j) + \lambda_4 \times \text{novelty}(u_i, p_j),$$

де $\text{sim}(u_i, p_j)$ – показник подібності характеристик;

$\text{market}(p_j)$ – ринкова привабливість професії;

$\text{access}(u_i, p_j)$ – доступність професії для користувача;

$\text{novelty}(u_i, p_j)$ – новизна рекомендації;

λ_k – ваги компонентів з умовою $\sum \lambda_k = 1$.

Ринкова привабливість професії розраховується на основі актуальних даних ринку праці:

$$\text{market}(p_j) = \gamma_1 \times \text{demand}(p_j) + \gamma_2 \times \text{salary}(p_j) + \gamma_3 \times \text{growth}(p_j),$$

де $\text{demand}(p_j)$ – попит на професію;

$\text{salary}(p_j)$ – рівень оплати праці;

$\text{growth}(p_j)$ – перспективи зростання галузі.

Доступність професії враховує освітні вимоги, географічні обмеження та індивідуальні можливості користувача:

$$\text{access}(u_i, p_j) = \delta_1 \times \text{education_match}(u_i, p_j) + \delta_2 \times \text{location_fit}(u_i, p_j) + \delta_3 \times \text{experience_gap}(u_i, p_j).$$

Диверсифікація рекомендацій забезпечується через штрафування надмірно схожих професій у фінальному списку. Модифікований алгоритм ранжування максимізує як релевантність, так і різноманітність:

$$R_{\text{div}}(u_i, p_j, S) = R(u_i, p_j) - \varepsilon \times \max_{\{p_k \in S\}} \text{sim}(p_j, p_k),$$

де S – множина вже відібраних рекомендацій;

ε - коефіцієнт диверсифікації.

Коректна нормалізація оцінок є критичною для забезпечення збалансованості різних компонентів функції корисності та порівнянності результатів між різними користувачами і професіями.

Z-score нормалізація приводить оцінки до стандартного розподілу з нульовим середнім:

$$z_score(x) = (x - \mu) / \sigma,$$

де μ та σ – відповідно середнє значення та стандартне відхилення по всій множині оцінок.

Min-max нормалізація масштабує значення до діапазону $[0,1]$:

$$\min_max(x) = (x - \min(X)) / (\max(X) - \min(X)),$$

де X – множина всіх значень характеристики.

Sigmoid нормалізація забезпечує плавне обмеження значень у діапазоні $[0,1]$:

$$\text{sigmoid}(x) = 1 / (1 + e^{(-x)}).$$

Для системи професійної орієнтації найбільш підходящою виявилася комбінована нормалізація, що використовує min-max для характеристик користувачів та професій, і z-score для контекстуальних факторів.

Ефективність системи рекомендацій значною мірою визначається алгоритмічною складністю обчислення подібності та ранжування професій. Для системи з n користувачами та m професіями базовий алгоритм має складність $O(n \times m \times d)$, де d - розмірність векторного простору характеристик.

Векторизація обчислень через використання матричних операцій дозволяє зменшити константний коефіцієнт складності:

$$\text{Similarity_Matrix} = U \times P^T / (\|U\| \times \|P\|^T),$$

де U та P – матриці користувачів та професій відповідно.

Попередня фільтрація професій за базовими критеріями зменшує простір пошуку:

$$P_filtered = \{p_j \in P : \text{market}(p_j) > \text{threshold}\}.$$

Індексування векторів за допомогою алгоритмів наближеного пошуку найближчих сусідів (ANN) дозволяє досягти сублінійної складності для великих баз професій.

Кешування результатів для схожих користувацьких профілів значно прискорює обслуговування повторних запитів:

$$\text{cache_key} = \text{hash}(\text{quantize}(u_i)),$$

де quantize – функція дискретизації профілю користувача.

Реальні умови функціонування системи професійної орієнтації характеризуються неповнотою інформації про користувачів, професії та ринкові умови. Математична модель має забезпечувати робастність до таких умов.

Байєсівський підхід до обробки невизначеності використовує апіорні розподіли для невідомих характеристик:

$$P(p_j \text{ suitable for } u_i) = \int P(p_j|\theta) \times P(\theta|\text{data}) d\theta,$$

де θ – параметри моделі;

data – наявні дані про користувача.

Інтервальна арифметика для роботи з діапазонами значень замість точкових оцінок:

$$[a_min, a_max] \oplus [b_min, b_max] = [a_min + b_min, a_max + b_max].$$

Множинні імпутації для відновлення відсутніх характеристик профілю користувача на основі схожих користувачів з повними профілями.

Розроблена математична модель забезпечує формальне підґрунтя для реалізації ефективної системи професійних рекомендацій, що поєднує наукову обґрунтованість з практичною корисністю для користувачів у процесі професійного самовизначення.

2.4 Висновки до розділу 2

У другому розділі проведено комплексне дослідження та порівняльний аналіз алгоритмів рекомендацій для системи професійної орієнтації, що дозволило обґрунтувати вибір Content-Based Filtering підходу та розробити його адаптацію для специфічних потреб профорієнтації.

Теоретичні основи Content-Based Filtering розкривають математичний апарат методу, що базується на векторному представленні користувачів та професій у багатовимірному просторі характеристик. Центральною концепцією є обчислення косинусної подібності між нормалізованими векторами користувацьких профілів та професійних вимог. Метод демонструє критичні переваги для професійної орієнтації: незалежність від історичних даних інших користувачів, можливість генерації рекомендацій для нових користувачів одразу після тестування, здатність пояснювати логіку рекомендацій через аналіз співпадінь характеристик.

Адаптація Content-Based алгоритмів для професійної орієнтації потребувала розробки п'ятивимірної моделі професійних характеристик, що включає технічну складність, креативність, соціальну взаємодію, аналітичні здібності та бізнес-орієнтованість. Кожна характеристика квантифікується в діапазоні $[0,1]$ через математичні формули, що враховують множинні компоненти професійної діяльності. Процес трансформації відповідей користувачів у векторне представлення реалізовано через систему експертно визначених ваг, що дозволяє точно відображати індивідуальні преференції у формалізованому математичному просторі.

Розроблена математична модель системи рекомендацій формалізує задачу професійної орієнтації як проблему оптимізації функції корисності $R: U \times P \rightarrow R$ у багатовимірному просторі характеристик. Базова косинусна подібність розширена до зваженої функції, що враховує різну важливість професійних характеристик через коефіцієнти w_k . Багатокритеріальна функція корисності інтегрує показники подібності з контекстуальними факторами — ринковою привабливістю професій, до-

ступністю для користувача та новизною рекомендацій. Система нормалізації забезпечує збалансованість різних компонентів через комбінацію min-max методу для характеристик та z-score для контекстуальних факторів.

Оптимізація алгоритмічної складності досягнута через векторизовані операції NumPy, що дозволяють обчислювати подібність з усіма професіями одночасно через матричне множення. Багаторівневе кешування включає збереження результатів рекомендацій для ідентичних профілів, попередньо обчислених норм векторів та матеріалізованих значень подібності для популярних комбінацій. Система забезпечує субсекундний час відгуку для генерації рекомендацій при базі з 20+ професій.

Результати аналізу підтверджують, що Content-Based підхід є оптимальним для систем професійної орієнтації завдяки поєднанню математичної обґрунтованості, практичної ефективності та можливості пояснення рекомендацій користувачам. Розроблена адаптація алгоритму враховує психологічні особливості професійного самовизначення, динаміку ринку праці та потреби української аудиторії, що створює основу для високоякісних персоналізованих рекомендацій професій.

3 ПРОЄКТУВАННЯ ТА РЕАЛІЗАЦІЯ СИСТЕМИ ПРОФЕСІЙНОЇ ОРІЄНТАЦІЇ

3.1 Архітектура системи

3.1.1 Загальна архітектура платформи

Система професійної орієнтації CareerGuide побудована за принципом модульної архітектури, що забезпечує високу масштабованість, надійність та зручність супроводу. Архітектурне рішення базується на поділі відповідальностей між клієнтською та серверною частинами з чіткою ізоляцією бізнес-логіки, рівня представлення даних та алгоритмічного ядра системи рекомендацій.

Загальна архітектура системи представляє собою трьохрівневу модель типу "клієнт-сервер" з додатковим рівнем машинного навчання для обробки рекомендацій. Першим рівнем є клієнтський веб-додаток, що надає інтерактивний інтерфейс користувача та забезпечує взаємодію з користувачем. Другий рівень містить серверну частину з REST API для обробки запитів та бізнес-логіки. Третій рівень включає базу даних для збереження користувацьких профілів, результатів тестування та інформації про професії.

Архітектура спроектована з урахуванням принципів мікросервісної організації, де кожен компонент має чітко визначену відповідальність та може розвиватися незалежно від інших компонентів. Це забезпечує гнучкість у процесі розробки та можливість горизонтального масштабування окремих модулів залежно від навантаження.

3.1.2 Компонентна діаграма системи

Система CareerGuide складається з п'яти основних компонентів, кожен з яких виконує специфічні функції у загальному процесі професійної орієнтації.

Frontend Client відповідає за презентаційний рівень системи та включає React-додаток з компонентами для проходження тестування, відображення результатів рекомендацій та перегляду аналітичної інформації. Цей компонент забезпечує адаптивний користувацький інтерфейс, що коректно працює на різних пристроях та підтримує повну українську локалізацію інтерфейсу.

API Gateway слугує єдиною точкою входу для всіх зовнішніх запитів та забезпечує маршрутизацію запитів до відповідних внутрішніх сервісів. Компонент реалізує функції аутентифікації, авторизації, логування запитів та обмеження швидкості запитів для захисту від зловживань.

Recommendation Engine містить алгоритмічне ядро системи на основі Content-Based Filtering підходу. Цей компонент відповідає за векторизацію відповідей користувача, обчислення міри подібності з професійними профілями та генерацію персоналізованих рекомендацій з ранжуванням за релевантністю.

Data Processing Layer забезпечує обробку та валідацію вхідних даних, управління сесіями користувачів та інтеграцію з зовнішніми джерелами даних про ринок

праці. Компонент також відповідає за агрегацію аналітичної статистики та формування звітів.

Database Layer реалізує рівень збереження даних та включає MongoDB [2] для зберігання користувацьких профілів, результатів тестування, професійних характеристик та аналітичних метрик. База даних спроектована з урахуванням оптимізації запитів для швидкого пошуку та аналізу великих обсягів даних.

3.1.3 Вибір технологічного стеку

Вибір технологічного стеку для платформи CareerGuide базувався на аналізі вимог до продуктивності, масштабованості, зручності розробки та довгострокової підтримки системи.

Для реалізації серверної частини обрано FastAPI фреймворк з мовою програмування Python 3.11. FastAPI забезпечує високу продуктивність завдяки асинхронній архітектурі, автоматичну генерацію OpenAPI документації та вбудовану підтримку валідації даних через Pydantic моделі. Python обраний через багатий екосистем бібліотек для машинного навчання (NumPy, Pandas, Scikit-learn) та зручність імплементації математичних алгоритмів рекомендацій.

Клієнтська частина реалізована за допомогою React 18 з TypeScript для забезпечення типобезпеки та покращення досвіду розробки. Для стилізації використовується Shadcn UI компоненти разом з Tailwind CSS, що забезпечує консистентний дизайн та швидку розробку інтерфейсу. Вибір React обумовлений його компонентною архітектурою, великою спільнотою розробників та відмінною підтримкою сучасних веб-стандартів.

В якості основної бази даних використовується MongoDB завдяки її гнучкій схемі документів, що добре підходить для зберігання різноманітних користувацьких профілів та результатів тестування. NoSQL підхід дозволяє легко адаптувати структуру даних при розвитку системи без складних міграцій схеми [2, 7, 10].

Для реалізації алгоритмів рекомендацій використовується NumPy для ефективних векторних обчислень, Pandas для обробки та аналізу даних, Scikit-learn для метрик подібності та класифікації. Ці бібліотеки забезпечують оптимальну продуктивність математичних обчислень та широкі можливості для аналітики.

3.1.4 Паттерни проєктування

При розробці архітектури системи CareerGuide застосовано декілька ключових паттернів проєктування, що забезпечують чистоту коду, зручність тестування та можливість розширення функціональності.

Model-View-Controller (MVC) застосований для чіткого поділу логіки додатку на рівні. Models представляють структуру даних та бізнес-логіку, Views відповідають за представлення інформації користувачу, Controllers обробляють користувачькі запити та координують взаємодію між моделями та представленнями.

Repository Pattern використовується для абстракції рівня доступу до даних та забезпечення незалежності бізнес-логіки від конкретної реалізації бази даних. Це дозволяє легко замінювати або модифікувати способи збереження даних без впливу на інші компоненти системи.

Dependency Injection реалізований для управління залежностями між компонентами системи та спрощення тестування окремих модулів. Паттерн забезпечує слабке зв'язування компонентів та можливість легкої заміни реалізацій сервісів.

Strategy Pattern застосований в модулі рекомендацій для можливості використання різних алгоритмів обчислення подібності та ранжування результатів. Це дозволяє експериментувати з різними підходами без зміни основної архітектури системи.

Factory Pattern використовується для створення різних типів тестових питань та форматування результатів залежно від контексту використання. Паттерн забезпечує гнучкість у розширенні типів контенту та форматів виведення.

Застосування цих паттернів проєктування забезпечує високу якість архітектури, зручність супроводу коду та можливість ефективного масштабування системи в майбутньому. Модульна організація (рис. 3.1) дозволяє команді розробників працювати над різними компонентами незалежно, що прискорює процес розробки та зменшує ризик конфліктів у коді.

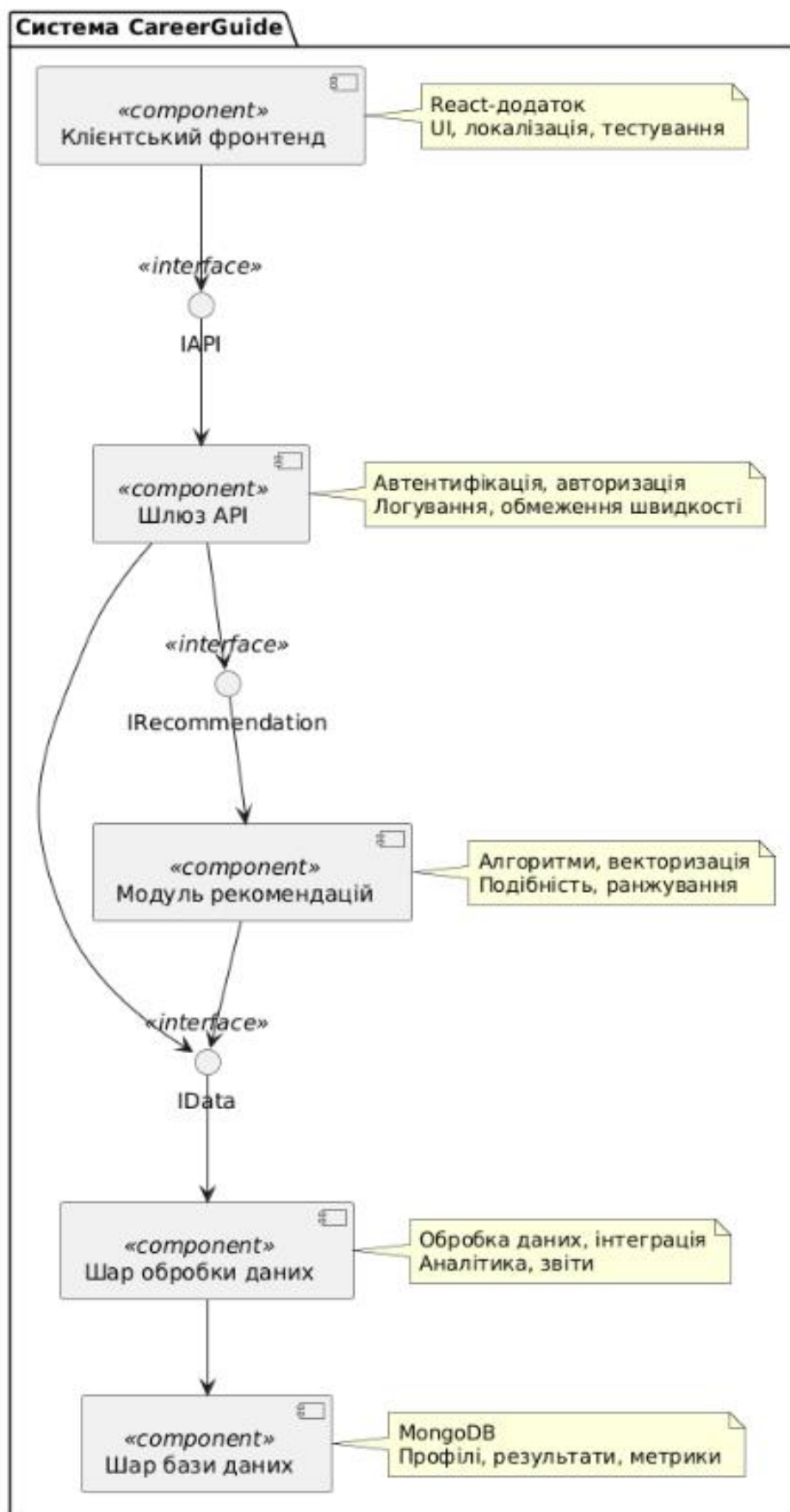


Рисунок 3.1 - Діаграма компонентів системи CareerGuide

3.2 Проєктування бази даних

3.2.1 ER-діаграма системи

Проєктування бази даних системи CareerGuide базується на аналізі основних сутностей предметної області професійної орієнтації та їх взаємозв'язків. Концептуальна модель даних відображає п'ять ключових сутностей, кожна з яких представляє окремий аспект функціонування платформи.

Сутність User представляє користувачів системи та містить базову інформацію про особу, що проходить тестування на професійну орієнтацію. Атрибути включають унікальний ідентифікатор користувача, персональні дані, контактну інформацію та метадані про створення та оновлення профілю. Користувач може мати множинні результати тестування, що дозволяє відстежувати зміни в професійних інтересах у часі.

Сутність Profession описує професії, які рекомендує система. Кожна професія характеризується назвою, детальним описом, вимогами до навичок, середньою зарплатою та рейтингом популярності. Професія також містить векторне представлення через п'ять основних характеристик: технічні навички, творчі здібності, соціальні навички, аналітичні здібності та бізнес-орієнтованість.

Сутність TestResult зберігає результати проходження користувачами тестів професійних інтересів. Кожен результат пов'язаний з конкретним користувачем та містить відповіді на питання тесту, обчислений користувацький профільний вектор та дату проходження тесту. Результати тестування служать основою для генерації персоналізованих рекомендацій.

Сутність Recommendation представляє згенеровані системою рекомендації професій для конкретного користувача. Кожна рекомендація містить посилання на результат тесту, рекомендовану професію, показник подібності та ранг рекомендації. Це дозволяє відстежувати історію рекомендацій та аналізувати їх ефективність.

Сутність Analytics агрегує статистичну інформацію про використання системи, популярність професій та ефективність рекомендацій. Аналітичні дані включають метрики користувацької активності, статистику проходжень тестів та показники задоволеності користувачів рекомендаціями.

Взаємозв'язки між сутностями реалізовані через зовнішні ключі та забезпечують цілісність даних. Користувач може мати багато результатів тестування (зв'язок один-до-багатьох), кожен результат тесту може генерувати множинні рекомендації (один-до-багатьох), а кожна рекомендація посиляється на конкретну професію (багато-до-одного).

3.2.2 Структура колекцій MongoDB

Система CareerGuide використовує документо-орієнтовану базу даних MongoDB, що забезпечує гнучкість схеми та ефективну обробку JSON-подібних документів [2, 7, 10]. База даних організована у п'ять основних колекцій, кожна з яких відповідає окремій сутності системи.

Колекція Users зберігає профілі користувачів у форматі документів з наступною структурою. Поле `_id` містить унікальний ідентифікатор MongoDB типу `ObjectId`. Поле `email` забезпечує унікальну ідентифікацію користувача для аутентифікації. Поля `firstName` та `lastName` містять персональні дані користувача. Поле `dateOfBirth` зберігає дату народження для демографічного аналізу. Поля `createdAt` та `updatedAt` відстежують час створення та останнього оновлення профілю.

Колекція Professions містить детальну інформацію про професії з векторним представленням характеристик. Поле `title` містить назву професії українською мовою. Поле `description` надає детальний опис професійної діяльності. Поле `requirements` перелічує необхідні навички та кваліфікації. Поле `averageSalary` зберігає інформацію про середню заробітну плату в гривнях. Поле `popularityRating` відображає рейтинг популярності професії від 1 до 5. Поле `skillsVector` містить масив з п'яти елементів, що представляють професійні характеристики: `technical`, `creative`, `social`, `analytical`, `business`.

Колекція TestResults зберігає результати проходження тестів користувачами. Поле `userId` містить посилання на ідентифікатор користувача. Поле `answers` зберігає

об'єкт з відповідями на питання тесту у форматі "questionId: answerValue". Поле userVector містить обчислений профільний вектор користувача на основі відповідей. Поле completedAt фіксує час завершення тестування. Поле testVersion дозволяє відстежувати версію тесту для підтримки зворотної сумісності.

Колекція Recommendations містить згенеровані системою рекомендації професій. Поле testResultId посилається на результат тесту, для якого генерувались рекомендації. Поле professionId ідентифікує рекомендовану професію. Поле similarityScore зберігає показник косинусної подібності між користувацьким вектором та професійним профілем. Поле rank визначає позицію рекомендації в ранжованому списку. Поле generatedAt фіксує час генерації рекомендації.

Колекція Analytics агрегує статистичну інформацію системи. Поле metric визначає тип метрики (наприклад, "daily_tests", "profession_popularity"). Поле value містить числове значення метрики. Поле date вказує період, до якого відноситься метрика. Поле metadata зберігає додаткову контекстну інформацію у форматі об'єкта.

3.2.3 Індекссування для оптимізації запитів

Ефективна робота системи CareerGuide потребує продуманої стратегії індексування для оптимізації найчастіших операцій читання та запису. Аналіз патернів використання даних дозволив визначити критичні поля для створення індексів.

У колекції Users створено унікальний індекс на поле email для швидкої аутентифікації користувачів та забезпечення унікальності електронних адрес. Композитний індекс на поля createdAt та updatedAt оптимізує запити для аналітичних звітів про активність користувачів у часовому розрізі.

Колекція Professions має індекс на поле title для швидкого пошуку професій за назвою та текстовий індекс на поля title та description для реалізації функції пошуку професій. Індекс на поле popularityRating прискорює сортування професій за популярністю в аналітичних запитах.

У колекції TestResults критично важливим є індекс на поле userId для швидкого отримання історії тестування конкретного користувача. Композитний індекс на поля userId та completedAt оптимізує запити для отримання останніх результатів

тестування. Індекс на поле `completedAt` забезпечує ефективне виконання аналітичних запитів за часовими періодами.

Колекція `Recommendations` має композитний індекс на поля `testResultId` та `rank` для швидкого отримання ранжованого списку рекомендацій. Індекс на поле `professionId` оптимізує зворотні запити для аналізу популярності професій у рекомендаціях.

Колекція `Analytics` використовує композитний індекс на поля `metric`, `date` для ефективного отримання аналітичних даних за типом метрики та часовим періодом. Цей індекс критично важливий для швидкої генерації дашбордів та звітів.

3.2.4 Стратегії збереження векторних даних

Векторні представлення користувацьких профілів та професійних характеристик є основою алгоритму рекомендацій у системі `CareerGuide`. Ефективне збереження та обробка векторних даних потребує спеціального підходу до проєктування структури документів.

Користувацькі вектори зберігаються у колекції `TestResults` як масиви п'яти нормалізованих значень з плаваючою комою. Кожен елемент масиву `userVector` відповідає одній з професійних характеристик: технічні навички, творчі здібності, соціальні навички, аналітичні здібності та бізнес-орієнтованість. Значення нормалізовані до діапазону $[0,1]$ для забезпечення коректного обчислення косинусної подібності. Професійні вектори у колекції `Professions` організовані аналогічно користувацьким векторам. Поле `skillsVector` містить п'ять елементів, кожен з яких представляє важливість відповідної характеристики для конкретної професії. Векторні дані для професій є попередньо обчисленими та заповнюються на етапі ініціалізації системи експертним шляхом.

Для оптимізації обчислень косинусної подібності система кешує попередньо обчислені норми векторів у додаткових полях `vectorNorm`. Це дозволяє уникати повторного обчислення норм при кожному запиті на генерацію рекомендацій, значно прискорюючи алгоритм.

Векторні операції реалізовані на рівні застосунку з використанням бібліотеки NumPy для максимальної продуктивності. MongoDB зберігає векторні дані як звичайні масиви, але вся математика виконується у Python коді з подальшим збереженням результатів у колекції Recommendations.

Стратегія збереження також передбачає можливість розширення векторного простору у майбутніх версіях системи. Структура даних спроектована таким чином, щоб додавання нових характеристик не потребувало міграції наявних даних, а лише оновлення алгоритмів обчислення подібності.

3.3 Реалізація Content-Based алгоритму

3.3.1 Структури даних для зберігання векторів

Ефективна реалізація Content-Based алгоритму рекомендацій у системі CareerGuide потребує продуманого підходу до організації векторних даних. Система використовує NumPy масиви для представлення як користувацьких профілів, так і професійних характеристик у п'ятивимірному векторному просторі.

Базова структура даних представляє собою нормалізований вектор розміром п'ять елементів, кожен з яких відповідає одній з ключових професійних характеристик. Технічні навички відображають схильність до роботи з програмуванням, інженерією та технологіями. Творчі здібності характеризують інтерес до дизайну, мистецтва та інноваційних рішень. Соціальні навички вказують на налаштованість до роботи з людьми, комунікації та командної взаємодії. Аналітичні здібності відображають схильність до логічного мислення, аналізу даних та вирішення складних задач. Бізнес-орієнтованість характеризує інтерес до підприємництва, менеджменту та комерційної діяльності.

Система використовує клас UserVector для інкапсуляції користувацьких характеристик та забезпечення типобезпеки операцій. Клас містить п'ятиелементний NumPy масив з обчисленою нормою вектора для оптимізації подальших обчислень

подібності. Додатково зберігаються метадані про час створення вектора та версію алгоритму обчислення.

Професійні вектори організовані через клас `ProfessionVector`, що містить аналогічну структуру даних з додатковими полями для ваг характеристик. Кожна професія має базовий вектор важливості характеристик та опціональні модифікатори для врахування специфічних вимог ринку праці чи регіональних особливостей.

Для оптимізації продуктивності система використовує кешування обчислених норм векторів та попередньо матеріалізованих результатів косинусної подібності для найпопулярніших комбінацій користувач-професія. Це дозволяє значно прискорити генерацію рекомендацій для повторних запитів.

3.3.2 Алгоритм конвертації відповідей тесту у вектор користувача

Перетворення відповідей користувача на психологічний тест у векторне представлення є ключовим етапом алгоритму рекомендацій. Система використовує експертно визначені ваги для кожної відповіді на питання тесту, що дозволяє точно відобразити користувацькі преференції у векторному просторі. Алгоритм конвертації реалізований через функцію `convert_answers_to_vector`, яка приймає словник відповідей користувача у форматі "question_id: answer_value" та повертає нормалізований п'ятивимірний вектор. Процес конвертації розпочинається з ініціалізації нульового вектора для кожної з п'яти характеристик (ліст. 3.1).

Для кожного питання тесту система має попередньо визначену матрицю ваг `ANSWER_WEIGHTS`, що встановлює внесок кожної можливої відповіді у відповідні компоненти вектора. Наприклад, позитивна відповідь на питання про цікавість до розробки нових технологій додає значущий внесок до компоненти технічних навичок та менший внесок до аналітичних здібностей.

Алгоритм ітерує через всі відповіді користувача, для кожної відповіді отримує відповідний вектор ваг з матриці `ANSWER_WEIGHTS` та додає його до загального вектора користувача. Після обробки всіх відповідей отриманий вектор нормалізується до одиничної довжини для забезпечення коректного обчислення косинусної подібності.

Лістинг 3.1 - Алгоритм конвертації

```

def convert_answers_to_vector(answers: Dict[str, int]) ->
np.ndarray:
    """
    Конвертує відповіді користувача у п'ятивимірний вектор професійних інтересів.

    Args:
        answers: Словник відповідей у форматі {question_id: answer_value}

    Returns:
        Нормалізований NumPy вектор розміром 5 елементів
    """
    # Ініціалізація вектора для п'яти характеристик
    user_vector = np.zeros(5) # [technical, creative, social, analytical, business]

    # Обробка кожної відповіді користувача
    for question_id, answer in answers.items():
        if question_id in ANSWER_WEIGHTS:
            weights = ANSWER_WEIGHTS[question_id][answer]
            user_vector += weights

    # Нормалізація вектора до одиничної довжини
    return normalize_vector(user_vector)

def normalize_vector(vector: np.ndarray) -> np.ndarray:
    """Нормалізує вектор до одиничної довжини."""
    norm = np.linalg.norm(vector)
    if norm == 0:
        return vector
    return vector / norm

```

Система також враховує можливість неповних відповідей, коли користувач пропускає деякі питання тесту. У таких випадках алгоритм використовує інтерполяцію на основі відповідей на подібні питання або застосовує середні значення для відповідної демографічної групи.

Для забезпечення стабільності результатів алгоритм включає механізм валідації отриманого вектора, що перевіряє його на наявність аномальних значень чи надмірної концентрації в одній характеристиці. У випадку виявлення аномалій система застосовує згладжування або повідомляє користувачу про необхідність повторного проходження тесту.

3.3.3 Функція розрахунку рекомендацій

Генерація персоналізованих рекомендацій професій базується на обчисленні косинусної подібності між вектором користувача та векторами всіх доступних професій у системі. Алгоритм реалізований у функції `generate_recommendations`, що забезпечує високу точність та ефективність обчислень.

Основна формула обчислення подібності використовує косинусну міру, що визначається як скалярний добуток нормалізованих векторів. Для векторів користувача U та професії P подібність обчислюється як $\text{similarity}(U,P) = U \cdot P / (\|U\| \times \|P\|)$. Оскільки всі вектори попередньо нормалізовані до одиничної довжини, формула спрощується до звичайного скалярного добутку.

Алгоритм використовує векторизовані операції NumPy для ефективного обчислення подібності з усіма професіями одночасно. Це дозволяє досягти високої продуктивності навіть при великій кількості професій у базі даних. Система формує матрицю професійних векторів розміром $N \times 5$, де N - кількість професій, та виконує матричне множення з вектором користувача.

Після обчислення всіх значень подібності алгоритм ранжує професії за спаданням подібності та відбирає топ- K найрелевантніших варіантів. Система використовує адаптивний підхід до вибору K , враховуючи розподіл значень подібності та забезпечуючи мінімальний поріг релевантності.

Додатково алгоритм застосовує фільтрацію результатів на основі додаткових критеріїв, таких як актуальність професії на ринку праці, відповідність освітньому рівню користувача та географічні преференції. Це дозволяє підвищити практичну цінність рекомендацій.

Система також реалізує механізм диверсифікації рекомендацій, що запобігає надмірній концентрації пропозицій у одній професійній сфері. Алгоритм аналізує категорії рекомендованих професій та при необхідності включає альтернативні варіанти з інших сфер діяльності для розширення горизонтів користувача.

3.3.4 Оптимізація продуктивності алгоритму

Забезпечення швидкого відгуку системи рекомендацій потребує комплексного підходу до оптимізації продуктивності алгоритму. Система використовує декілька рівнів оптимізації для мінімізації часу генерації рекомендацій та ефективного використання обчислювальних ресурсів.

На рівні структур даних система використовує ефективне представлення векторів через NumPy масиви з оптимізованим розташуванням в пам'яті. Професійні вектори зберігаються у вигляді контігуальної матриці, що дозволяє використовувати оптимізовані BLAS операції для матричних обчислень.

Кешування є ключовим елементом стратегії оптимізації. Система підтримує багаторівневий кеш, що включає кешування результатів рекомендацій для ідентичних користувацьких профілів, кешування обчислених подібностей між популярними комбінаціями векторів та кешування попередньо обчислених норм професійних векторів. Алгоритм використовує батчеву обробку для ефективного обчислення рекомендацій для множини користувачів одночасно. Це особливо важливо для сценаріїв масового тестування або аналітичних звітів, коли потрібно обробити великі кількості запитів. Система реалізує асинхронну архітектуру для обчислення рекомендацій, що дозволяє не блокувати користувацький інтерфейс під час виконання складних обчислень. Фонові задачі обробляють генерацію рекомендацій та оновлення кешу, забезпечуючи швидкий відгук для користувачів.

Для подальшої оптимізації алгоритм використовує апроксимацію обчислень у випадках, коли повна точність не є критичною. Наприклад, система може використовувати швидші, але менш точні алгоритми подібності для попереднього відбору кандидатів та застосовувати точні обчислення лише для фінального ранжування.

Моніторинг продуктивності інтегровано в алгоритм через збір метрик часу виконання, використання пам'яті та ефективності кешування. Ці дані використовуються для автоматичного налаштування параметрів системи та виявлення потенційних вузьких місць продуктивності.

3.4 Розробка користувацького інтерфейсу

3.4.1 UX/UI дизайн платформи

Проектування користувацького досвіду системи CareerGuide базується на принципах інтуїтивності, доступності та ефективності взаємодії з користувачем. Дизайн інтерфейсу створено з урахуванням психологічних особливостей процесу професійного самовизначення та потреб української аудиторії.

Головна сторінка платформи, представлена на рисунку 3.2, демонструє мінімалістичний підхід до дизайну з чітким фокусом на основній функціональності. Центральний елемент сторінки містить заохочувальний заголовок "Знайди свою ідеальну професію" та короткий опис переваг платформи. Статистичні блоки відображають ключові показники системи: кількість доступних професій, проведених тестів та точність рекомендацій, що формує довіру користувачів до платформи.

Колірна схема інтерфейсу побудована навколо зеленого кольору як основного, що психологічно асоціюється з ростом, розвитком та позитивними змінами у житті. Додаткові кольори включають світлі відтінки для фону та контрастні акценти для інтерактивних елементів. Типографіка використовує сучасні шрифти, що забезпечують високу читабельність на різних пристроях.

Навігаційна структура сайту організована через верхню панель з основними розділами платформи. Розділ "Тест інтересів" надає доступ до основного функціоналу тестування, "Аналітика" відображає статистичну інформацію про ринок праці, а розділ "Головна" повертає користувача на стартову сторінку.

Блоки функціональності на головній сторінці представлені через картки з іконками та описами основних можливостей системи. Персоналізований тест дозволяє пройти оцінювання професійних інтересів, аналітика ринку надає актуальну інформацію про попит та зарплати, а професійна орієнтація забезпечує детальні рекомендації щодо кар'єрного розвитку.

Головна сторінка оформлена у стриманому сучасному стилі та зосереджена на ключових можливостях сервісу.

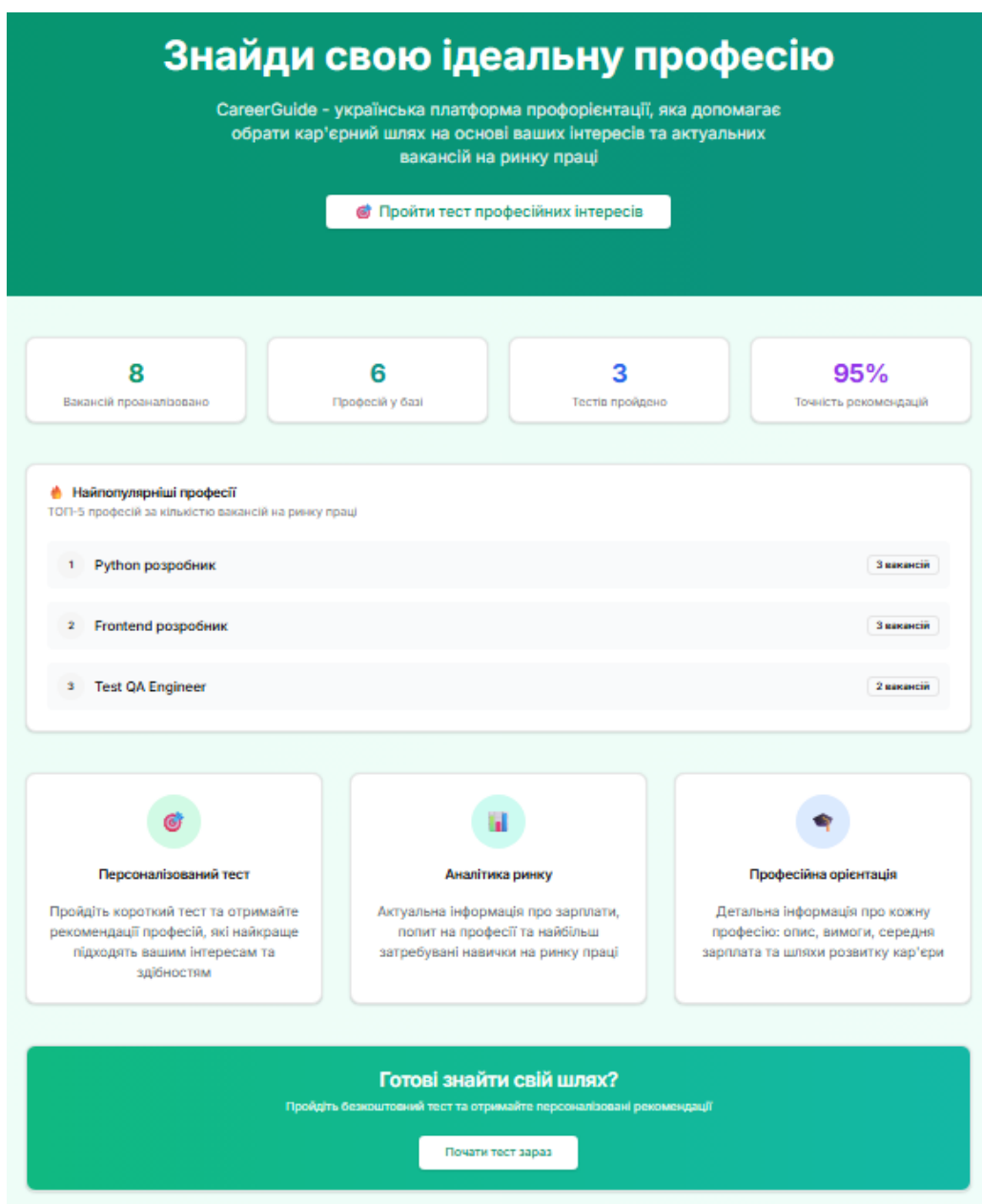


Рисунок 3.2 - Головна сторінка

3.4.2 Компонентна архітектура React додатку

Клієнтська частина системи CareerGuide реалізована за допомогою React 18 з використанням функціональних компонентів та сучасних хуків для управління станом. Архітектура додатку побудована на принципах модульності, повторного використання коду та ефективного управління даними.

Кореневий компонент App служить точкою входу для всього додатку та містить маршрутизацію між основними сторінками системи. Компонент використовує

React Router для навігації та Context API для глобального управління станом користувача та результатами тестування.

Компонент `Homepage` відповідає за відображення головної сторінки з статичними блоками та описом функціональності. Компонент використовує стан для зберігання динамічних даних про кількість користувачів та проведених тестів, які отримуються через API запити при завантаженні сторінки.

Система тестування реалізована через групу компонентів `TestContainer`, `QuestionCard` та `ProgressBar`. Процес проходження тесту, показаний на рисунках 3.3-3.7, демонструє послідовну навігацію між питаннями з чіткими індикаторами прогресу. `TestContainer` управляє загальним процесом проходження тесту, включаючи навігацію між питаннями та збереження відповідей. `QuestionCard` відображає окреме питання з варіантами відповідей, реалізованими як радіо-кнопки з кастомним стилізуванням, як показано на рисунку 3.3. `ProgressBar` візуалізує прогрес проходження тесту через заповнену смугу прогресу, яка поступово збільшується від рисунка 3.3 до рисунка 3.7.

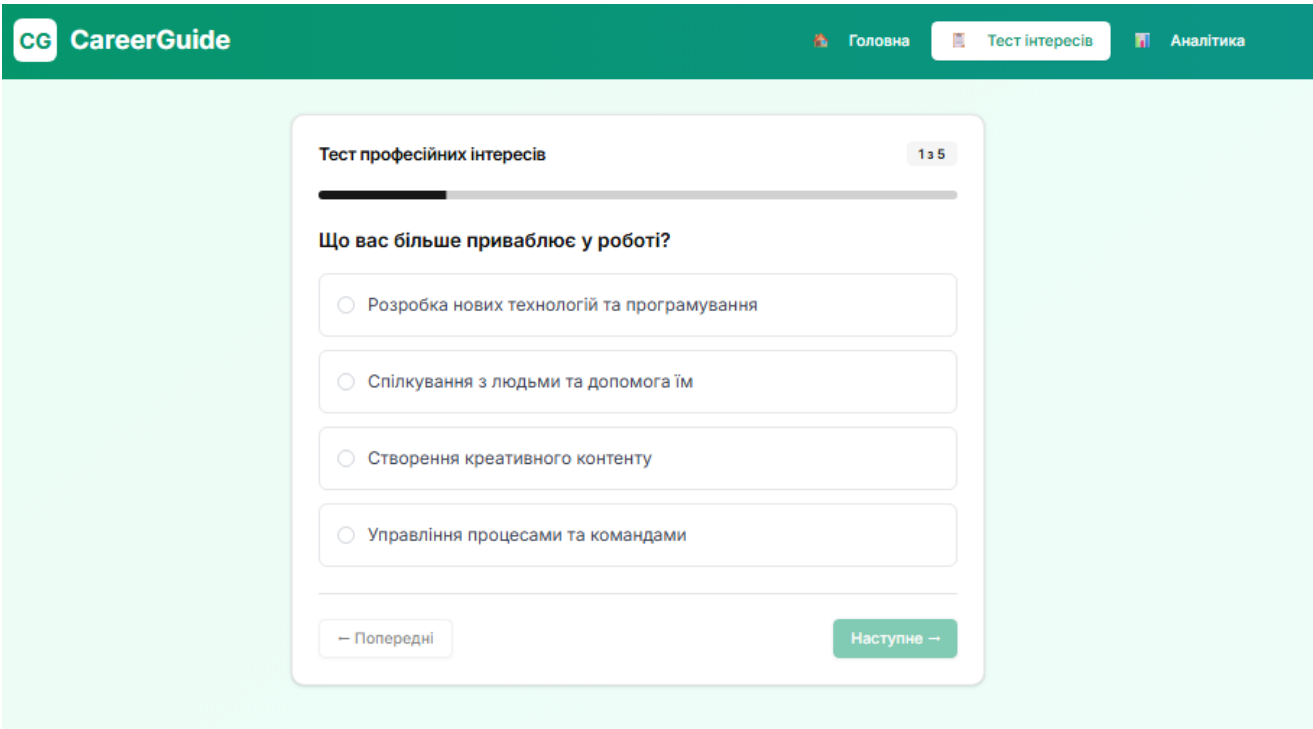


Рисунок 3.3 – Тест інтересів. Питання 1

CG CareerGuide

Головна | Тест інтересів | Аналітика

Тест професійних інтересів 2 з 5

Який стиль роботи вам підходить більше?

- ☐ Самостійна робота з мінімальним контролем
- ☐ Командна робота над спільними проектами
- ☒ Творча співпраця з іншими
- ☐ Керування та координація інших

← Попередні | Наступне →

Рисунок 3.4 – Тест інтересів. Питання 2

CG CareerGuide

Головна | Тест інтересів | Аналітика

Тест професійних інтересів 3 з 5

Яке робоче середовище вам ближче?

- ☒ Офіс з сучасними технологіями
- ☐ Відкритий простір для спілкування
- ☐ Творча студія або майстерня
- ☐ Переговорна кімната

← Попередні | Наступне →

Рисунок 3.5 – Тест інтересів. Питання 3

The screenshot shows the 'CareerGuide' website interface. At the top, there is a green header with the 'CG CareerGuide' logo on the left and navigation links 'Головна', 'Тест інтересів', and 'Аналітика' on the right. The main content area is light green. A white card in the center contains the test title 'Тест професійних інтересів' and a progress indicator '4 з 5'. Below the title is a question: 'Які завдання вас мотивують найбільше?'. There are four radio button options: 'Вирішення складних технічних проблем', 'Допомога людям вирішити їхні проблеми', 'Створення чогось нового та оригінального', and 'Досягнення бізнес-цілей'. At the bottom of the card are two buttons: '← Попередні' and 'Наступне →'.

Тест професійних інтересів 4 з 5

Які завдання вас мотивують найбільше?

- ☐ Вирішення складних технічних проблем
- ☐ Допомога людям вирішити їхні проблеми
- ☐ Створення чогось нового та оригінального
- ☐ Досягнення бізнес-цілей

← Попередні Наступне →

Рисунок 3.6 – Тест інтересів. Питання 4

The screenshot shows the 'CareerGuide' website interface. At the top, there is a green header with the 'CG CareerGuide' logo on the left and navigation links 'Головна', 'Тест інтересів', and 'Аналітика' on the right. The main content area is light green. A white card in the center contains the test title 'Тест професійних інтересів' and a progress indicator '5 з 5'. Below the title is a question: 'Як ви ставитесь до навчання нових навичок?'. There are four radio button options: 'Люблю вивчати нові технології та інструменти', 'Цікавлюсь психологією та комунікаціями', 'Розвиваю творчі здібності', and 'Вивчаю бізнес та менеджмент'. At the bottom of the card are two buttons: '← Попередні' and 'Завершити тест ✨'.

Тест професійних інтересів 5 з 5

Як ви ставитесь до навчання нових навичок?

- ☐ Люблю вивчати нові технології та інструменти
- ☐ Цікавлюсь психологією та комунікаціями
- ☐ Розвиваю творчі здібності
- ☐ Вивчаю бізнес та менеджмент

← Попередні Завершити тест ✨

Рисунок 3.7 – Тест інтересів. Питання 5

Компонент ResultsPage відповідає за відображення персоналізованих рекомендацій професій після завершення тестування, як показано на рисунку 3.8. Компонент отримує дані рекомендацій через props та відображає їх у вигляді карток з детальною інформацією про кожну професію, включаючи опис, вимоги та середню заробітну плату.

Аналітичний розділ реалізований через компонент Analytics, що містить підкомпоненти для різних типів статистичної інформації, як показано на рисунках 3.8-3.10. PopularProfessions відображає топ професій за кількістю рекомендацій (рисунок 3.8), SalaryStatistics показує статистику зарплат з мінімальними, середніми та максимальними значеннями (рисунок 3.9), а SkillsTrends візуалізує тренди затребуваності навичок на ринку праці (рисунок 3.10).

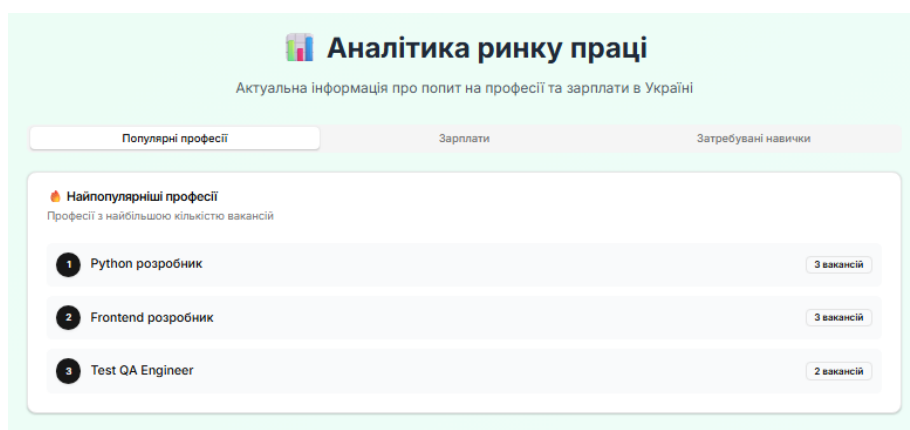


Рисунок 3.8 – Аналітичний розділ

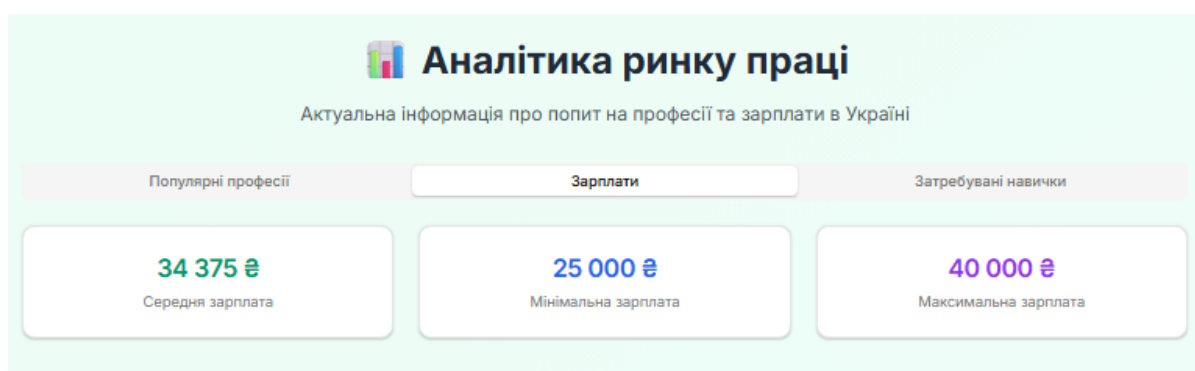


Рисунок 3.9 – Статистика зарплат

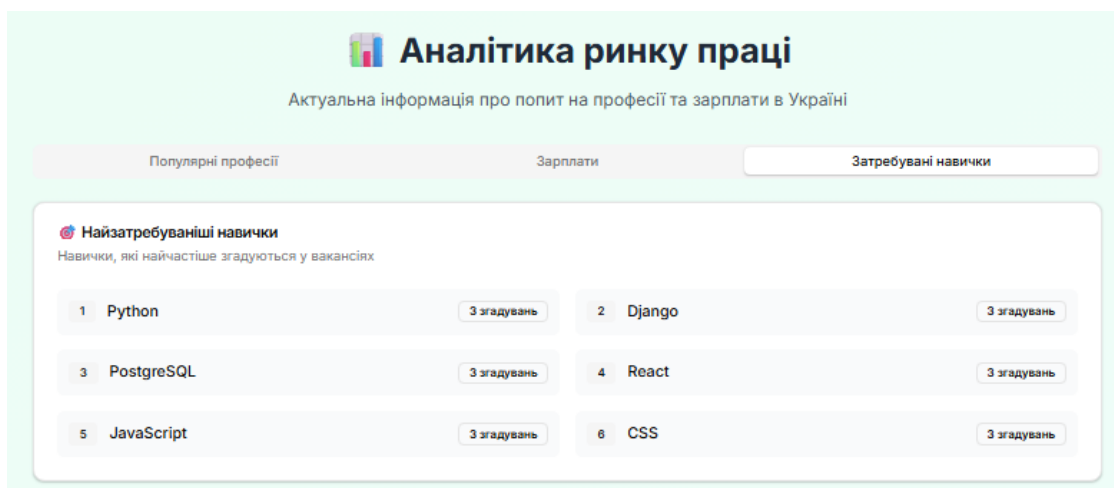


Рисунок 3.10 – Тренди затребуваності навичок на ринку праці

3.4.3 Адаптивний дизайн для різних пристроїв

Система CareerGuide розроблена з урахуванням різноманітності пристроїв користувачів, включаючи настільні комп'ютери, планшети та смартфони. Адаптивний дизайн реалізований через CSS Grid та Flexbox технології з використанням медіа-запитів для різних розмірів екранів.

Для настільних пристроїв з шириною екрана понад 1024 пікселі інтерфейс використовує повну ширину для відображення максимальної кількості інформації. Статистичні блоки розташовані горизонтально в ряд, а картки професій відображаються в багатостовпковій сітці для ефективного використання простору.

На планшетних пристроях з шириною екрана від 768 до 1024 пікселів система автоматично адаптує розміщення елементів. Навігаційне меню трансформується в компактнішу форму, статистичні блоки перегруповуються в дві колонки, а розмір тексту збільшується для покращення читабельності на сенсорних екранах. Мобільна версія для екранів менше 768 пікселів характеризується вертикальним розташуванням елементів та збільшеними зонами дотику. Навігаційне меню перетворюється в "гамбургер" меню для економії простору, всі блоки контенту розташовуються в одну колонку, а кнопки та інтерактивні елементи мають мінімальний розмір 44 пікселі для зручності використання пальцями.

3.4.4 Візуалізація результатів рекомендацій

Представлення персоналізованих рекомендацій професій є центральним елементом користувацького досвіду системи CareerGuide. Візуалізація результатів, показана на рисунку 3.11, створена для максимального розуміння та практичної цінності отриманих рекомендацій.

Сторінка результатів розпочинається з персоналізованого профілю користувача, що відображає розподіл його професійних інтересів за п'ятьма основними категоріями.

Рекомендовані професії представлені у вигляді карток з ранжуванням за ступенем відповідності профілю користувача. Кожна картка містить назву професії, короткий опис діяльності, ключові навички та середню заробітну плату в гривнях. Кольорове кодування карток від зеленого до синього відображає ступінь відповідності, що допомагає користувачу швидко визначити найбільш релевантні варіанти. Система використовує адаптивний алгоритм ранжування, який враховує не лише відповідність профілю, але й актуальність професії на ринку праці, перспективи розвитку кар'єри та рівень попиту серед роботодавців. Картки організовані в чіткій ієрархії від найбільш релевантних до менш відповідних варіантів, що дозволяє користувачу ефективно орієнтуватися серед множини пропозицій та зосередитися на найперспективніших можливостях.

Додаткова інформація для кожної професії включає теги основних навичок, що дозволяє користувачу оцінити відповідність своїм поточним компетенціям та визначити напрямки для розвитку. Система також відображає актуальну інформацію про попит на ринку праці через індикатори затребуваності професій. Кожен тег навички виконаний у вигляді окремого елемента з можливістю отримання детальної інформації про рівень важливості цієї компетенції для конкретної професії. Індикатори попиту базуються на аналізі актуальних вакансій та тенденцій ринку праці, надаючи користувачу об'єктивну картину перспектив працевлаштування. Динамічне оновлення інформації про заробітні плати та затребуваність професій забезпечує актуальність рекомендацій відповідно до поточної ситуації на ринку праці України.

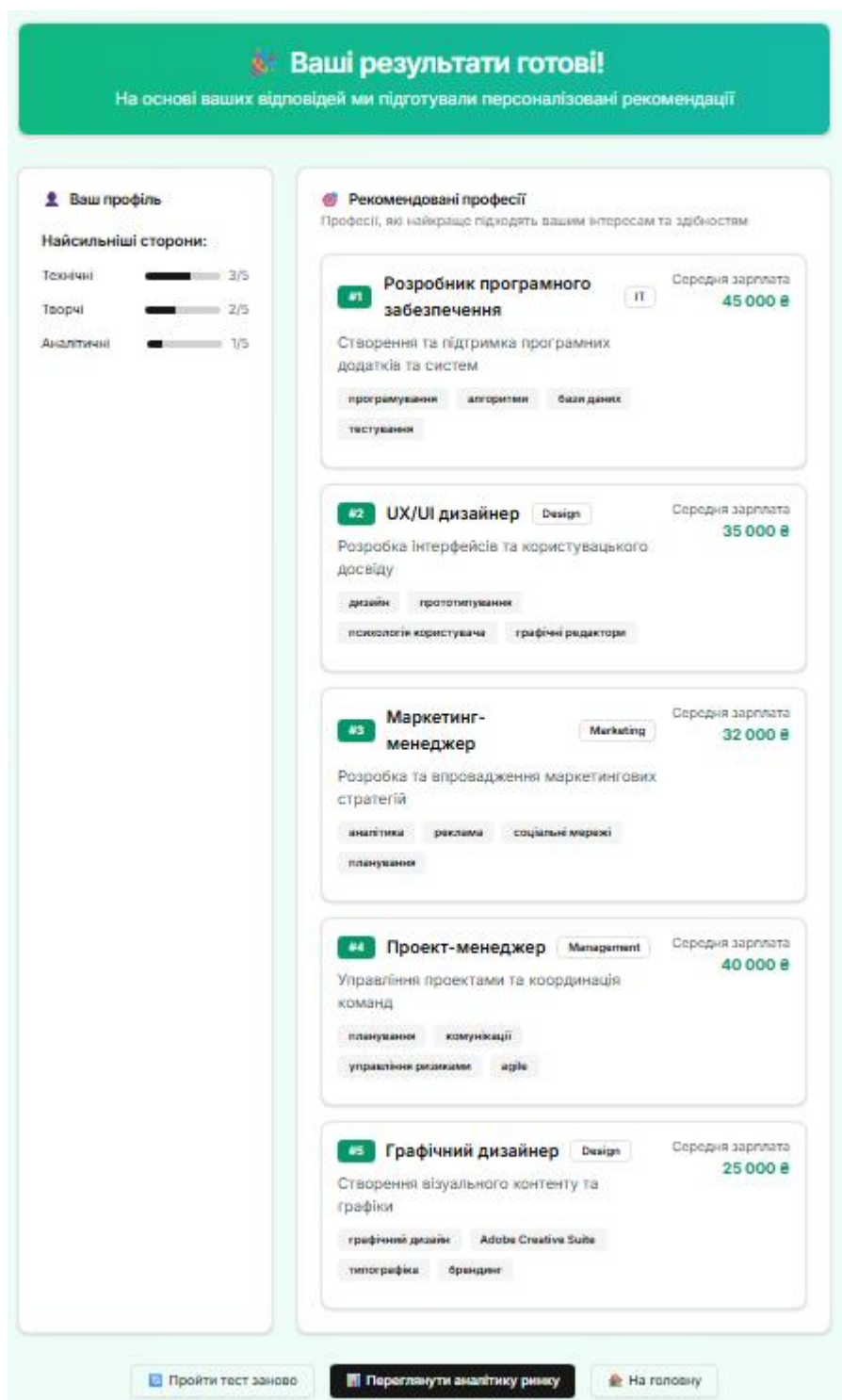


Рисунок 3.11 – Візуалізація результатів

3.4.5 Інтеграція з backend API

Взаємодія між клієнтською та серверною частинами системи CareerGuide реалізована через RESTful API з використанням HTTP протоколу та JSON формату

обміну даними. Архітектура API забезпечує ефективну комунікацію та надійність системи.

Основні API ендпоінти включають GET запит до `/api/test/questions` для отримання питань тесту, POST запит до `/api/test/submit` для відправлення відповідей користувача та GET запит до `/api/recommendations/{userId}` для отримання персоналізованих рекомендацій. Додатково система використовує `/api/analytics/popular-professions` для статистичної інформації та `/api/analytics/market-trends` для даних про ринок праці.

Обробка помилок API реалізована через стандартизовані HTTP статус коди та структуровані повідомлення про помилки. Клієнтський додаток містить централизовану систему обробки помилок, що відображає зрозумілі повідомлення користувачу та пропонує відповідні дії для вирішення проблем.

Система кешування на клієнтському боці використовує `localStorage` для збереження результатів тестування та рекомендацій, що дозволяє користувачам швидко повертатися до своїх результатів без повторних запитів до сервера. Кеш автоматично оновлюється при зміні даних користувача або проходженні нового тесту.

Аутентифікація користувачів реалізована через JWT токени з автоматичним оновленням сесій. Система підтримує анонімне використання для одноразового тестування та реєстрацію користувачів для збереження історії результатів та персоналізованих рекомендацій.

3.5 Реалізація аналітичного модуля

3.5.1 Система метрик ефективності

Для забезпечення комплексного моніторингу ефективності роботи платформи CareerGuide та якості генерованих рекомендацій необхідно розробити систему збору та аналізу ключових метрик.

Передбачається впровадження технічних метрик для відстеження часу відгуку системи на запити користувачів, пропускну здатності серверної частини під час пікових навантажень та використання ресурсів бази даних. Планується налаштувати автоматичний збір даних про час виконання кожного етапу процесу рекомендацій – від обробки відповідей тесту до генерації фінального списку професій. Особливу увагу слід приділити моніторингу продуктивності Content-Based алгоритму, оскільки обчислення косинусної подібності може створювати значне навантаження при великій кількості професій у базі даних.

Важливим напрямком роботи є впровадження метрик точності алгоритмів як ключової складової системи оцінювання. Потрібно організувати відстеження відсотка користувачів, які вибирають професії з топ-3 рекомендацій, середнього рангу обраної професії в списку рекомендацій та кореляції між оцінками подібності й фактичним інтересом користувачів. Доцільно також проаналізувати розподіл значень косинусної подібності для всіх генерованих рекомендацій, що допоможе виявити потенційні проблеми з алгоритмом векторизації або нормалізацією даних. Необхідно запровадити моніторинг стабільності рекомендацій при повторному тестуванні тих самих користувачів через певний період часу.

Планується збирати показники користувацького досвіду, зокрема час, витрачений на проходження тесту, відсоток завершених тестувань від загальної кількості розпочатих та частоту повернення користувачів для повторного тестування.

Для персоналізації рекомендацій та виявлення потенційних упереджень в алгоритмах щодо певних груп користувачів передбачається автоматичне відстеження демографічних характеристик – вікових груп, освітнього рівня та географічного розподілу.

3.5.2 Дашборди для аналізу користувацької поведінки

Для комплексного огляду користувацької активності та ефективності роботи платформи необхідно розробити візуальну аналітичну панель системи CareerGuide.

На головному дашборді передбачається відображення ключових показників продуктивності в реальному часі – кількості активних користувачів, проведених

тестувань за день та середнього часу відгуку системи. Планується впровадити візуальні індикатори стану системи через кольорове кодування, де зелений колір вказуватиме на нормальну роботу, жовтий – на підвищене навантаження, а червоний – на потенційні проблеми. Окремий блок має показувати статус роботи Content-Based алгоритму, включаючи середній час генерації рекомендацій та кількість оброблених запитів за хвилину.

Дашборд користувацької активності повинен відображати динаміку використання платформи протягом різних часових періодів. Графіки мають показувати денні, тижневі та місячні тренди кількості нових користувачів, завершених тестувань та генерованих рекомендацій. Доцільно налаштувати автоматичне виявлення аномальних піків або спадів активності з позначенням для додаткового аналізу. Особливий інтерес становитимуть кореляції між активністю користувачів та зовнішніми факторами, такими як початок навчального року або періоди профорієнтаційних заходів у навчальних закладах.

Аналітична панель рекомендацій має візуалізувати ефективність алгоритмів через розподіл оцінок подібності, частоту вибору кожної професії користувачами та статистику задоволеності рекомендаціями. Планується створити теплові карти для відображення кореляції між різними професійними характеристиками та популярністю професій серед різних демографічних груп. Така візуалізація допоможе виявити паттерни, які можуть бути не очевидними при аналізі табличних даних.

Географічний дашборд повинен відображати розподіл користувачів по регіонах України з деталізацією до рівня областей та міст.

3.5.3 Статистика популярності професій

Для розуміння тенденцій користувацьких переваг система аналітики CareerGuide має вести детальну статистику популярності професій серед користувачів платформи, як показано на рисунку 3.8.

Передбачається визначення топ найпопулярніших професій, що дозволить відстежувати попит на різні спеціальності серед української молоді. Необхідно організувати моніторинг змін у популярності професій у часі для виявлення трендів,

які можуть бути пов'язані з ринковими змінами, освітніми програмами або соціально-економічними факторами. Особливу увагу слід приділити динаміці популярності ІТ-професій серед користувачів різних вікових груп.

Статистичний аналіз має включати сегментацію популярності професій за демографічними характеристиками користувачів. Планується виявляти професійні преференції різних вікових груп та проводити гендерний аналіз для адаптації алгоритмів рекомендацій та усунення потенційних упереджень. Наприклад, потрібно передбачити можливість коригування ваг професійних характеристик для забезпечення рівних можливостей рекомендації технічних професій для всіх користувачів незалежно від статі.

Регіональна статистика популярності професій має виявляти географічні особливості професійних інтересів. Очікується, що у великих містах спостерігатиметься вища популярність ІТ-професій, тоді як у менших населених пунктах — більший інтерес до традиційних професій та сфери послуг. Ці дані планується використовувати для персоналізації рекомендацій з урахуванням регіональних особливостей ринку праці. Доцільно також проаналізувати сезонні коливання популярності професій, пов'язані з періодами вступних кампаній, випуску фахівців та економічними циклами.

3.5.4 А/В тестування різних версій алгоритму

Для постійного покращення алгоритмів рекомендацій та користувацького досвіду необхідно впровадити систему А/В тестування на платформі CareerGuide.

Основні сценарії А/В тестування мають включати порівняння різних підходів до векторизації відповідей користувачів, варіації вагових коефіцієнтів для професійних характеристик та альтернативні методи ранжування рекомендацій. Кожен тест слід проводити на статистично значущій вибірці користувачів з рандомізованим розподілом між контрольною та експериментальною групами. Мінімальний розмір вибірки має становити 1000 користувачів для кожної групи, що забезпечить статистичну потужність експерименту на рівні 80%.

Передбачається налаштування автоматичного збору метрик ефективності для кожної групи користувачів. До них належатимуть точність рекомендацій, задоволеність користувачів та час прийняття рішення щодо вибору професії. Статистичний аналіз результатів має використовувати t-тести та хі-квадрат тести для визначення значущості різниць між варіантами алгоритму. Рівень значущості планується встановити на 0,05, що відповідає стандартним вимогам для A/B тестування у технологічних продуктах.

Доцільно організувати довгострокові A/B тести для відстеження впливу змін алгоритму на поведінку користувачів протягом декількох місяців. Такі тести допоможуть виявити не лише короткочасний вплив на точність рекомендацій, але й довгострокові ефекти на задоволеність користувачів та їх лояльність до платформи. Результати A/B тестування використовуватимуться для прийняття рішень щодо впровадження нових версій алгоритмів. Система має вимагати статистично значущого покращення ключових метрик перед переходом на нову версію алгоритму для всіх користувачів.

Планується також проводити мультиваріантні тести для оптимізації декількох параметрів алгоритму одночасно.

3.6 Висновки до розділу 3

У третьому розділі здійснено проєктування та реалізацію системи професійної орієнтації CareerGuide з використанням сучасних веб-технологій та алгоритмів машинного навчання.

Розроблена архітектура системи базується на трьохрівневій клієнт-серверній моделі з чіткою сепарацією відповідальностей між компонентами. Технологічний стек включає FastAPI для серверної частини, React 18 з TypeScript для клієнтського додатку та MongoDB для збереження даних. Такий вибір технологій забезпечує високу продуктивність, масштабованість та зручність розробки системи.

Проектування бази даних виконано з урахуванням специфіки векторних даних для алгоритму рекомендацій. Розроблена структура MongoDB колекцій оптимізована для ефективного зберігання користувацьких профілів, професійних характеристик та результатів рекомендацій. Система індексування забезпечує швидкий доступ до даних при генерації рекомендацій та аналітичних запитах.

Реалізація Content-Based алгоритму рекомендацій включає ефективні структури даних на основі NumPy масивів, алгоритм конвертації відповідей тесту у п'ятивимірний користувацький вектор та функцію розрахунку косинусної подібності для генерації персоналізованих рекомендацій. Застосовано багаторівневу оптимізацію продуктивності через кешування обчислень, векторизацію операцій та батчеву обробку запитів.

Користувацький інтерфейс спроектовано відповідно до сучасних UX/UI принципів з акцентом на інтуїтивність та доступність. Розроблено адаптивний дизайн, що коректно працює на настільних комп'ютерах, планшетах та мобільних пристроях. Компонентна архітектура React додатку забезпечує модульність, повторне використання коду та ефективне управління станом через Context API та хуки.

Візуалізація результатів рекомендацій реалізована через інтерактивні картки професій з ранжуванням за ступенем відповідності користувацькому профілю. Система відображає персоналізований профіль користувача, детальну інформацію про рекомендовані професії та актуальні дані про ринок праці, що допомагає користувачам приймати обґрунтовані рішення щодо професійного розвитку.

Аналітичний модуль системи включає комплексну систему метрик для моніторингу ефективності роботи платформи та якості рекомендацій. Передбачено впровадження візуальних дашбордів для аналізу користувацької поведінки, статистики популярності професій та A/B тестування різних версій алгоритмів. Розроблена система CareerGuide демонструє практичну реалізацію теоретичних підходів Content-Based Filtering у специфічному домені професійної орієнтації, забезпечуючи баланс між науковою обґрунтованістю алгоритмів та практичною корисністю для користувачів у процесі професійного самовизначення.

4 АПАРАТНЕ ЗАБЕЗПЕЧЕННЯ

4.1 Серверна інфраструктура

Система професійної орієнтації CareerGuide потребує продуманої серверної інфраструктури для забезпечення надійної роботи та масштабованості. Архітектура серверного середовища спроектована з урахуванням поточних потреб системи та можливості горизонтального масштабування при зростанні користувацької бази.

Мінімальні вимоги до серверного обладнання включають процесор з чотирма фізичними ядрами та тактовою частотою не менше 2.5 GHz для забезпечення паралельної обробки запитів користувачів. Оперативна пам'ять обсягом 8 GB є достатньою для початкового етапу експлуатації системи з базою до 1000 одночасних користувачів. Дисковий простір SSD обсягом 256 GB забезпечує швидкий доступ до бази даних та файлової системи додатку.

Рекомендована конфігурація для продуктивного середовища передбачає використання процесора з вісьмома ядрами, 16 GB оперативної пам'яті та 512 GB SSD накопичувача. Така конфігурація дозволяє обслуговувати до 5000 одночасних користувачів при збереженні прийняттого часу відгуку системи. Мережеве підключення має забезпечувати пропускну здатність не менше 100 Mbps для швидкої передачі даних між клієнтськими додатками та сервером.

Для забезпечення відмовостійкості рекомендується використання кластерної архітектури з мінімум двома серверними вузлами. Балансування навантаження між вузлами дозволяє рівномірно розподіляти запити користувачів та забезпечувати безперервну роботу системи навіть при виході з ладу одного з вузлів. Резервування даних реалізується через реплікацію бази даних MongoDB на декілька серверів з автоматичним переключенням на резервний вузол у випадку відмови основного.

4.2 Вимоги до клієнтського обладнання

Веб-орієнтована архітектура системи CareerGuide забезпечує широку доступність платформи для користувачів з різноманітними пристроями. Мінімальні вимоги до клієнтського обладнання є достатньо помірними, що дозволяє охопити максимально широку аудиторію.

Для настільних комп'ютерів та ноутбуків мінімальні вимоги включають процесор з тактовою частотою 1 GHz, 2 GB оперативної пам'яті та сучасний веб-браузер версії не старше двох років. Підтримуються браузери Google Chrome 90+, Mozilla Firefox 88+, Microsoft Edge 90+ та Safari 14+.

Мобільні пристрої з операційними системами Android 8.0+ або iOS 12+ повністю підтримуються завдяки адаптивному дизайну інтерфейсу. Рекомендовані характеристики для оптимального досвіду включають 3 GB оперативної пам'яті та стабільне інтернет-з'єднання зі швидкістю завантаження не менше 3 Mbps.

Планшетні пристрої з діагоналлю екрану від 7 дюймів отримують оптимізований інтерфейс, що поєднує переваги мобільної та десктопної версій. Система автоматично визначає тип пристрою та адаптує розташування елементів для максимальної зручності використання. Підтримка сенсорних екранів реалізована через збільшені зони дотику для інтерактивних елементів та жестів для навігації.

4.3 Мережева інфраструктура

Надійна мережева інфраструктура є критичною для забезпечення швидкого та безперебійного доступу користувачів до платформи CareerGuide. Система вимагає постійного підключення до мережі Інтернет як для серверної частини, так і для клієнтських пристроїв.

Серверне середовище має бути розміщене в дата-центрі з гарантованим часом безвідмовної роботи не менше 99,9% річно. Підключення до мережі Інтернет має забезпечувати пропускну здатність не менше 1 Gbps з резервними каналами зв'язку для забезпечення відмовостійкості. Захист від DDoS атак реалізується через спеціалізовані мережеві рішення, що фільтрують шкідливий трафік до його досягнення серверів додатку.

Для клієнтських пристроїв мінімальні вимоги до швидкості інтернет-з'єднання становлять 2 Mbps для завантаження та 1 Mbps для відправлення даних. Такої швидкості достатньо для комфортного проходження тестування, перегляду рекомендацій та взаємодії з аналітичними дашбордами. Рекомендована швидкість 5 Mbps та вище забезпечує миттєве завантаження всіх елементів інтерфейсу та плавну роботу інтерактивних візуалізацій.

Латентність мережевого з'єднання також впливає на користувацький досвід. Система оптимізована для роботи при затримках до 150 мс між клієнтом та сервером, що є типовим для з'єднань в межах України. Використання Content Delivery Network (CDN) для статичних ресурсів дозволяє зменшити латентність через кешування контенту на географічно розподілених серверах ближче до кінцевих користувачів.

4.4 Висновки до розділу 4

Розділ висвітлює комплексні технічні вимоги до апаратного та мережевого забезпечення системи професійної орієнтації CareerGuide, які є критичними для забезпечення надійної та масштабованої роботи платформи.

Серверна інфраструктура системи спроектована з урахуванням можливості поетапного нарощування потужностей. Мінімальна конфігурація з чотириядерним процесором та 8 GB оперативної пам'яті забезпечує обслуговування до 1000 одночасних користувачів, тоді як рекомендована конфігурація з вісьмома ядрами та 16

GB пам'яті дозволяє масштабувати систему до 5000 користувачів. Кластерна архітектура з реплікацією бази даних MongoDB гарантує відмовостійкість та безперервність роботи навіть при виході з ладу окремих компонентів.

Веб-орієнтований підхід до побудови системи забезпечує широку доступність платформи для користувачів з різноманітними пристроями. Помірні вимоги до клієнтського обладнання, підтримка всіх основних браузерів та адаптивний дизайн інтерфейсу дозволяють працювати з системою на настільних комп'ютерах, ноутбуках, планшетах та мобільних пристроях без втрати функціональності. Мінімальні системні вимоги охоплюють навіть бюджетні пристрої, що є важливим фактором для забезпечення рівного доступу учнів та студентів до інструментів професійної орієнтації.

Мережева інфраструктура побудована з урахуванням вимог до надійності та продуктивності. Розміщення в професійному дата-центрі з гарантованим часом роботи 99,9% та багатоканальним підключенням забезпечує стабільний доступ до системи. Використання CDN для статичних ресурсів мінімізує затримки при завантаженні контенту, а захист від DDoS атак гарантує безперебійність роботи навіть в умовах зовнішніх загроз. Оптимізація системи для роботи при типовій для України латентності до 150 мс забезпечує комфортну взаємодію користувачів з платформою незалежно від їх географічного розташування.

ВИСНОВКИ

У кваліфікаційній роботі вирішено актуальну науково-практичну задачу розробки веб-платформи професійної орієнтації з використанням алгоритмів машинного навчання, що дозволяє підвищити точність професійних рекомендацій та забезпечити персоналізований підхід до професійного самовизначення молоді.

Проведений комплексний аналіз предметної області виявив критичні недоліки традиційних підходів до професійної орієнтації в Україні. Встановлено, що класичні системи базуються на застарілих психодіагностичних методиках, які не враховують сучасні професії в галузі інформаційних технологій та характеризуються низькими показниками валідності. Статистичні дані підтверджують значний розрив між освітою та реальними потребами ринку праці – лише 23% випускників працюють за отриманою спеціальністю протягом перших п'яти років після закінчення навчання. Порівняльний аналіз вітчизняних та зарубіжних систем професійної орієнтації дозволив виявити ключові напрямки удосконалення – необхідність застосування науково обґрунтованих методологій, розширення баз професій та інтеграції з реальними даними ринку праці.

Теоретичне дослідження рекомендаційних систем дозволило обґрунтувати вибір Content-Based Filtering як оптимального підходу для системи професійної орієнтації. Розроблена математична модель базується на векторному представленні користувачів та професій у п'ятивимірному просторі характеристик (технічна складність, креативність, соціальна взаємодія, аналітичні здібності, бізнес-орієнтованість). Процес генерації рекомендацій формалізовано як задачу оптимізації багатокритеріальної функції корисності, що враховує не лише косинусну подібність між користувацьким профілем та професійними вимогами, але й контекстуальні фактори – ринкову привабливість професій, доступність для користувача та новизну рекомендацій. Застосування векторизованих операцій NumPy та багаторівневого кешування забезпечує субсекундний час відгуку для генерації рекомендацій навіть при обробці значних обсягів даних.

Розроблена архітектура системи CareerGuide базується на трьохрівневій клієнт-серверній моделі з використанням сучасного технологічного стеку: FastAPI для серверної частини, React 18 з TypeScript для клієнтського додатку та MongoDB для збереження даних. Така архітектура забезпечує чітку сепарацію відповідальностей між компонентами, високу продуктивність та можливість горизонтального масштабування системи. Проектування бази даних виконано з урахуванням специфіки зберігання векторних даних та необхідності ефективного доступу до інформації при генерації рекомендацій. Користувацький інтерфейс спроектовано відповідно до сучасних UX/UI принципів з акцентом на інтуїтивність та доступність, а адаптивний дизайн забезпечує коректну роботу на різних типах пристроїв.

Технічне рішення системи передбачає гнучку серверну інфраструктуру з можливістю поетапного нарощування потужностей. Мінімальна конфігурація забезпечує обслуговування до 1000 одночасних користувачів, а рекомендована дозволяє масштабувати систему до 5000 користувачів. Кластерна архітектура з реплікацією бази даних гарантує відмовостійкість та безперервність роботи. Веб-орієнтований підхід із помірними вимогами до клієнтського обладнання забезпечує широку доступність платформи для користувачів з різноманітними пристроями. Розміщення в професійному дата-центрі з гарантованим часом роботи 99,9%, використання CDN для статичних ресурсів та захист від DDoS атак забезпечують стабільну та безпечну роботу системи.

Практична цінність розробленої системи полягає в поєднанні наукової обґрунтованості традиційних психометричних підходів з адаптивністю та персоналізацією сучасних рекомендаційних алгоритмів. Система забезпечує генерацію персоналізованих професійних рекомендацій з поясненням логіки вибору, відображення актуальних даних про ринок праці та інтуїтивний користувацький інтерфейс. Впровадження розробленої платформи дозволить покращити якість професійного самовизначення молоді, підвищити ефективність професійної орієнтації в навчальних закладах та зменшити невідповідність між освітою та вимогами ринку праці.

Створена система демонструє успішну інтеграцію теоретичних підходів машинного навчання з практичними потребами професійної орієнтації, що відкриває перспективи для подальшого розвитку інтелектуальних систем підтримки прийняття рішень у сфері освіти та кар'єрного розвитку.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Adapa S., Nah F., Hall R. Intelligent career guidance system using machine learning // International Conference on Human-Computer Interaction. 2021. Vol. 12782. P. 234–248.
2. Aleksendrić M. Full Stack FastAPI, React, and MongoDB: Build Python web applications with the FARM stack. Birmingham : Packt Publishing, 2022. 336 p.
3. Padma E., Sowdharshini P., Shanmugapriya P. та ін. Career guidance system for students using machine learning // Challenges in Information, Communication and Computing Technology. 2024. Vol. 5. P. 87–94.
4. State of the Tech Workforce 2024. CompTIA. URL: <https://www.comptia.org/content/research/state-of-the-tech-workforce> (дата звернення: 24.11.2025).
5. Content-based filtering. Machine Learning / Google for Developers. URL: <https://developers.google.com/machine-learning/recommendation/content-based/basics> (дата звернення: 24.11.2025).
6. Cosine similarity // Wikipedia. URL: https://en.wikipedia.org/wiki/Cosine_similarity (дата звернення: 24.11.2025).
7. DataCamp. FARM Stack Guide: How to Build Full-Stack Apps with FastAPI, React & MongoDB. URL: <https://www.datacamp.com/tutorial/farm-stack-guide> (дата звернення: 24.11.2025).
8. Education in Ukraine // WENR World Education News + Reviews. URL: <https://wenr.wes.org/2019/06/education-in-ukraine> (дата звернення: 24.11.2025).
9. Employment in the IT industry — Statistics & Facts // Statista. URL: <https://www.statista.com/topics/5275/employment-in-the-it-industry/> (дата звернення: 24.11.2025).
10. Full stack, modern web application generator using FastAPI, MongoDB as database / MongoDB Labs. URL: <https://github.com/mongodb-labs/full-stack-fastapi-mongodb> (дата звернення: 24.11.2025).

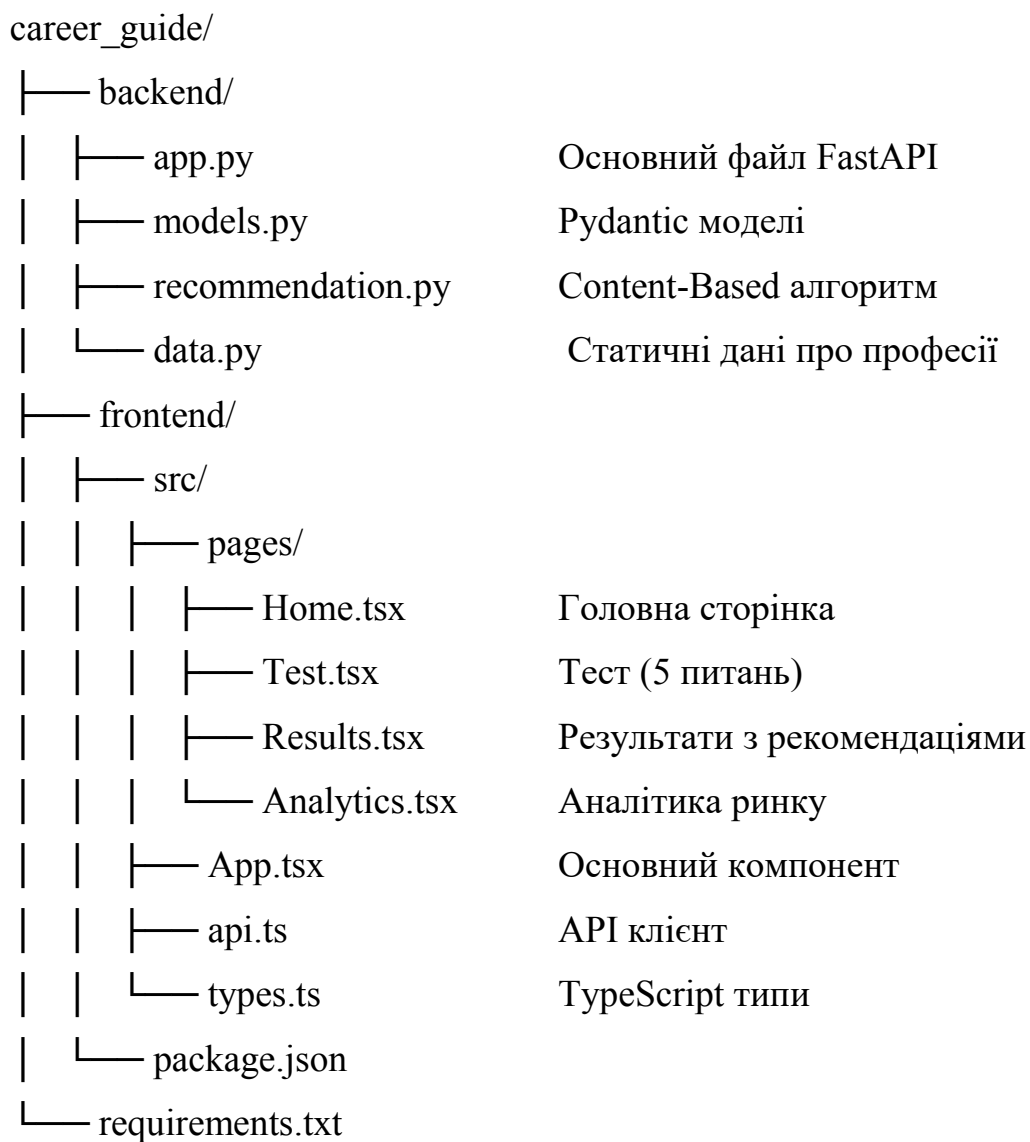
11. GOVET. Vocational education and training in Ukraine: A system in transition. URL: <https://www.govet.international/en/159632.php> (дата звернення: 24.11.2025).
12. Higher education in Ukraine // Wikipedia. URL: https://en.wikipedia.org/wiki/Higher_education_in_Ukraine (дата звернення: 24.11.2025).
13. IBM. What is content-based filtering? URL: <https://www.ibm.com/think/topics/content-based-filtering> (дата звернення: 24.11.2025).
14. IBM. What is cosine similarity? URL: <https://www.ibm.com/think/topics/cosine-similarity> (дата звернення: 24.11.2025).
15. Johnson J., Anderson M., Williams K. Natural language processing for career guidance systems // Journal of Artificial Intelligence Research. 2022. Vol. 45. P. 127–145.
16. Khan A., Prabu S., Harrouk M. NLP and machine learning for intelligent career guidance // International Journal of Science and Engineering Applications. 2025. Vol. 14. № 8. P. 42–46.
17. Li Z., Chen Y., Wang X. Designing AI-based career recommendation systems // IEEE Transactions on Learning Technologies. 2023. Vol. 16. № 3. P. 289–302.
18. Ministry of Education and Science of Ukraine. Vocational Education and Training. URL: <https://mon.gov.ua/en/tag/vocational-education-and-training> (дата звернення: 24.11.2025).
19. Recommender system // Wikipedia. URL: https://en.wikipedia.org/wiki/Recommender_system (дата звернення: 24.11.2025).
20. Recommender Systems – A Complete Guide to Machine Learning Models // Towards Data Science. URL: <https://towardsdatascience.com/recommender-systems-a-complete-guide-to-machine-learning-models-96d3f94ea748> (дата звернення: 24.11.2025).
21. Salton G., Wong A., Yang C. S. A Vector Space Model for Automatic Indexing // Communications of the ACM. 1975. Vol. 18. № 11. P. 613–620.

22. Smith M., Thompson L., Davis K. Machine learning approaches to automate career guidance // International Journal of Career Development. 2021. Vol. 29. № 4. P. 445–462.
23. Spot Intelligence. Vector Space Model Made Simple With Examples & Tutorial In Python. URL: <https://spotintelligence.com/2023/09/07/vector-space-model/> (дата звернення: 24.11.2025).
24. StrataScratch. Step-by-Step Guide to Building Content-Based Filtering. URL: <https://www.stratascratch.com/blog/step-by-step-guide-to-building-content-based-filtering> (дата звернення: 24.11.2025).
25. U.S. Bureau of Labor Statistics. Computer and Information Technology Occupations. URL: <https://www.bls.gov/ooh/computer-and-information-technology/> (дата звернення: 24.11.2025).
26. Upwork. What Is Content-Based Filtering? URL: <https://www.upwork.com/resources/what-is-content-based-filtering> (дата звернення: 24.11.2025).
27. Vector space model // Wikipedia. URL: https://en.wikipedia.org/wiki/Vector_space_model (дата звернення: 24.11.2025).
28. Yang C., Dong L., Wang J. та ін. Machine Learning and Digital Twin-Based Assistance System for Career Prediction // Journal of Electrical Systems. 2024. Vol. 20. № 3s. P. 1822–1832.
29. Yang C., Dong L. Intelligent talent recommendation algorithm for college students using deep learning // Journal of Electrical Systems. 2024. Vol. 20. № 3. P. 1822–1832.
30. Dossari A., Smith R., Brown T. та ін. Prediction of recommendations for employment utilizing machine learning procedures // Materials Today: Proceedings. 2021. Vol. 51. P. 2847–2853.
31. Сіренко В. В., Грушко С. С., Тіменко А. В. Інформаційна платформа профорієнтації з аналізом ринку праці та прогнозуванням затребуваності професій // Інформаційні технології і автоматизація – 2025 : XVIII міжнар. наук.-практ. конф., 30–31 жовтня 2025 р. : тези доповідей. Одеса, 2025.

ДОДАТОК А

СТРУКТУРА ПРОЄКТУ

Рисунок А.1 – Структура проєкту



ДОДАТОК В

ЛІСТИНГИ ПРОГРАМ

Лістинг В.1 – app.py

```

from fastapi import FastAPI, HTTPException
from fastapi.middleware.cors import CORSMiddleware
from pydantic import BaseModel
from typing import List, Dict, Any
import uvicorn

from models import TestQuestion, TestAnswer, RecommendationResult
from recommendation import calculate_recommendations, convert_answers_to_vector
from data import PROFESSIONS, TEST_QUESTIONS, ANALYTICS_DATA

app = FastAPI(title="CareerGuide API", version="1.0.0")

# CORS middleware для фронтенду
app.add_middleware(
    CORSMiddleware,
    allow_origins=["http://localhost:3000"],
    allow_credentials=True,
    allow_methods=["*"],
    allow_headers=["*"],
)

@app.get("/api/test/questions")
async def get_test_questions() -> List[TestQuestion]:
    """Повертає 5 питань тесту"""
    return TEST_QUESTIONS

@app.post("/api/test/submit")
async def submit_test_answers(answers: Dict[int, str]) -> RecommendationResult:
    """Приймає відповіді тесту та повертає рекомендації"""
    if len(answers) != 5:
        raise HTTPException(status_code=400, detail="Потрібно відповісти на всі 5 питань")

    # Перетворюємо відповіді в вектор користувача
    user_vector = convert_answers_to_vector(answers)

    # Розраховуємо рекомендації
    recommendations = calculate_recommendations(user_vector)

    # Обчислюємо профіль користувача для відображення
    user_profile = {

```

```

        "technical": int(user_vector[0] * 5),
        "creative": int(user_vector[1] * 5),
        "analytical": int(user_vector[2] * 5),
        "social": int(user_vector[3] * 5),
        "business": int(user_vector[4] * 5)
    }

    return RecommendationResult(
        user_profile=user_profile,
        recommendations=recommendations
    )

@app.get("/api/analytics/stats")
async def get_analytics_stats():
    """Статистика для головної сторінки"""
    return ANALYTICS_DATA["stats"]

@app.get("/api/analytics/popular")
async def get_popular_professions():
    """Популярні професії"""
    return ANALYTICS_DATA["popular_professions"]

@app.get("/api/analytics/salaries")
async def get_salary_data():
    """Дані про зарплати"""
    return ANALYTICS_DATA["salaries"]

@app.get("/api/analytics/skills")
async def get_skills_data():
    """Затребувані навички"""
    return ANALYTICS_DATA["skills"]

if __name__ == "__main__":
    uvicorn.run(app, host="0.0.0.0", port=8000)

```

Лістинг В.2 – models.py

```

from pydantic import BaseModel
from typing import List, Dict, Any

class TestQuestion(BaseModel):
    id: int
    question: str
    options: Dict[str, str]

class TestAnswer(BaseModel):
    question_id: int
    answer: str

class UserProfile(BaseModel):
    technical: int

```

```

    creative: int
    analytical: int
    social: int
    business: int

class Recommendation(BaseModel):
    id: int
    name: str
    category: str
    salary: int
    description: str
    similarity_score: float
    match_percentage: int

class RecommendationResult(BaseModel):
    user_profile: UserProfile
    recommendations: List[Recommendation]

```

Лістинг В.3 – recommendation.py

```

import numpy as np
from typing import List, Dict
from data import PROFESSIONS

def convert_answers_to_vector(answers: Dict[int, str]) ->
np.array:
    """
    Перетворює відповіді з 5 питань тесту в вектор користувача
    answers: {1: "option_a", 2: "option_c", ...}
    Returns: [technical, creative, analytical, social, business]
    """
    # Стартовий вектор
    user_vector = np.array([0.0, 0.0, 0.0, 0.0, 0.0])

    # Питання 1: Що вас більше приваблює в роботі?
    if answers[1] == "tech":          # Розробка нових технологій
        user_vector += [0.3, 0.0, 0.2, 0.0, 0.0]
    elif answers[1] == "people":      # Спілкування з людьми
        user_vector += [0.0, 0.0, 0.0, 0.3, 0.1]
    elif answers[1] == "creative":    # Створення креативного
контенту
        user_vector += [0.0, 0.3, 0.0, 0.0, 0.1]
    elif answers[1] == "management":  # Управління процесами
        user_vector += [0.0, 0.0, 0.1, 0.1, 0.3]

    # Питання 2: Який стиль роботи підходить більше?
    if answers[2] == "independent":   # Самостійна робота
        user_vector += [0.2, 0.2, 0.2, 0.0, 0.0]
    elif answers[2] == "team":        # Командна робота
        user_vector += [0.0, 0.1, 0.0, 0.3, 0.1]
    elif answers[2] == "creative_team": # Творча співпраця

```

```

        user_vector += [0.0, 0.3, 0.0, 0.2, 0.0]
    elif answers[2] == "leadership": # Керування командами
        user_vector += [0.0, 0.0, 0.1, 0.2, 0.3]

    # Питання 3: Яке робоче середовище ближче?
    if answers[3] == "tech_office": # Офіс з сучасними тех-
нологіями
        user_vector += [0.3, 0.0, 0.1, 0.0, 0.0]
    elif answers[3] == "open_space": # Відкритий простір для
спілкування
        user_vector += [0.0, 0.0, 0.0, 0.3, 0.1]
    elif answers[3] == "creative_studio": # Творча студія
        user_vector += [0.0, 0.3, 0.0, 0.1, 0.0]
    elif answers[3] == "negotiations": # Переговорна кімната
        user_vector += [0.0, 0.0, 0.1, 0.1, 0.3]

    # Питання 4: Які завдання мотивують найбільше?
    if answers[4] == "technical_problems": # Вирішення складних
технічних проблем
        user_vector += [0.3, 0.0, 0.3, 0.0, 0.0]
    elif answers[4] == "helping_people": # Допомога людям
        user_vector += [0.0, 0.0, 0.0, 0.3, 0.1]
    elif answers[4] == "creating_new": # Створення нового та
оригінального
        user_vector += [0.1, 0.3, 0.1, 0.0, 0.0]
    elif answers[4] == "business_goals": # Досягнення бізнес-
цілей
        user_vector += [0.0, 0.0, 0.1, 0.1, 0.3]

    # Питання 5: Ставлення до навчання нових навичок?
    if answers[5] == "tech_tools": # Люблю вивчати технології
        user_vector += [0.2, 0.0, 0.1, 0.0, 0.0]
    elif answers[5] == "psychology": # Цікавлюсь психологією
та комунікаціями
        user_vector += [0.0, 0.0, 0.0, 0.2, 0.1]
    elif answers[5] == "creative_skills": # Розвиваю творчі зді-
бності
        user_vector += [0.0, 0.2, 0.0, 0.0, 0.1]
    elif answers[5] == "business": # Вивчаю бізнес та мене-
джмент
        user_vector += [0.0, 0.0, 0.0, 0.1, 0.2]

    # Нормалізація вектора до діапазону [0, 1]
    return np.clip(user_vector, 0, 1)

def calculate_recommendations(user_vector: np.array) ->
List[Dict]:
    """
    Розраховує топ-5 професій за Content-Based алгоритмом
    """
    recommendations = []

    for prof_name, prof_data in PROFESSIONS.items():

```

```

# Створюємо вектор професії
prof_vector = np.array([
    prof_data["technical"],
    prof_data["creative"],
    prof_data["analytical"],
    prof_data["social"],
    prof_data["business"]
])

# Косинусна подібність
dot_product = np.dot(user_vector, prof_vector)
magnitude = np.linalg.norm(user_vector) *
np.linalg.norm(prof_vector)
similarity = dot_product / magnitude if magnitude > 0
else 0

# Додаємо до результатів з відсотком відповідності
recommendations.append({
    **prof_data,
    "similarity_score": similarity,
    "match_percentage": int(similarity * 100)
})

# Повертаємо топ-5 за similarity score
return sorted(recommendations, key=lambda x:
x["similarity_score"], reverse=True)[:5]

```

Лістинг В.4 – data.py

```

from typing import List, Dict

# Статичні дані про професії з Content-Based характеристиками
PROFESSIONS = {
    "Розробник програмного забезпечення": {
        "id": 1,
        "name": "Розробник програмного забезпечення",
        "category": "IT",
        "salary": 45000,
        "description": "Створення та підтримка програмних дода-
тків та систем",
        "skills": ["програмування", "алгоритми", "бази даних",
"тестування"],
        # Content-Based характеристики
        "technical": 0.9,      # Технічні навички
        "creative": 0.3,      # Творчість
        "analytical": 0.8,    # Аналітичне мислення
        "social": 0.2,        # Соціальна взаємодія
        "business": 0.1       # Бізнес-орієнтованість
    },
    "UX/UI дизайнер": {
        "id": 2,
        "name": "UX/UI дизайнер",

```

```

        "category": "Дизайн",
        "salary": 35000,
        "description": "Розробка інтерфейсів та користувацького
досвіду",
        "skills": ["дизайн", "прототипування", "психологія ко-
ристувача", "графічні редактори"],
        "technical": 0.6,
        "creative": 0.9,
        "analytical": 0.6,
        "social": 0.4,
        "business": 0.3
    },
    "Маркетинг-менеджер": {
        "id": 3,
        "name": "Маркетинг-менеджер",
        "category": "Маркетинг",
        "salary": 32000,
        "description": "Розробка та впровадження маркетингових
стратегій",
        "skills": ["аналітика", "реклама", "соціальні мережі",
"планування"],
        "technical": 0.2,
        "creative": 0.7,
        "analytical": 0.8,
        "social": 0.6,
        "business": 0.8
    },
    "Проект-менеджер": {
        "id": 4,
        "name": "Проект-менеджер",
        "category": "Менеджмент",
        "salary": 40000,
        "description": "Управління проектами та координація ко-
манд",
        "skills": ["планування", "комунікації", "управління ри-
зиками", "agile"],
        "technical": 0.3,
        "creative": 0.4,
        "analytical": 0.7,
        "social": 0.8,
        "business": 0.9
    },
    "Графічний дизайнер": {
        "id": 5,
        "name": "Графічний дизайнер",
        "category": "Дизайн",
        "salary": 25000,
        "description": "Створення візуального контенту та гра-
фіки",
        "skills": ["графічний дизайн", "Adobe Creative Suite",
"типографіка", "брендинг"],
        "technical": 0.4,
        "creative": 0.95,

```

```

        "analytical": 0.3,
        "social": 0.3,
        "business": 0.2
    },
    "Python розробник": {
        "id": 6,
        "name": "Python розробник",
        "category": "IT",
        "salary": 42000,
        "description": "Розробка веб-додатків та систем на
Python",
        "skills": ["Python", "Django", "PostgreSQL", "API"],
        "technical": 0.95,
        "creative": 0.2,
        "analytical": 0.9,
        "social": 0.1,
        "business": 0.1
    },
    "Frontend розробник": {
        "id": 7,
        "name": "Frontend розробник",
        "category": "IT",
        "salary": 38000,
        "description": "Розробка користувацького інтерфейсу веб-
додатків",
        "skills": ["JavaScript", "React", "CSS", "HTML"],
        "technical": 0.8,
        "creative": 0.6,
        "analytical": 0.7,
        "social": 0.2,
        "business": 0.1
    },
    "Test QA Engineer": {
        "id": 8,
        "name": "Test QA Engineer",
        "category": "IT",
        "salary": 35000,
        "description": "Тестування програмного забезпечення та
забезпечення якості",
        "skills": ["тестування", "автоматизація", "Selenium",
"баг-трекінг"],
        "technical": 0.7,
        "creative": 0.3,
        "analytical": 0.9,
        "social": 0.3,
        "business": 0.2
    }
}

# Питання тесту
TEST_QUESTIONS = [
    {
        "id": 1,

```

```

"question": "Що вас більше приваблює у роботі?",
"options": {
  "tech": "Розробка нових технологій та програмування",
  "people": "Спілкування з людьми та допомога їм",
  "creative": "Створення креативного контенту",
  "management": "Управління процесами та командами"
},
{
  "id": 2,
  "question": "Який стиль роботи вам підходить більше?",
  "options": {
    "independent": "Самостійна робота з мінімальним
контролем",
    "team": "Командна робота над спільними проектами",
    "creative_team": "Творча співпраця з іншими",
    "leadership": "Керування та координація інших"
  },
  {
    "id": 3,
    "question": "Яке робоче середовище вам ближче?",
    "options": {
      "tech_office": "Офіс з сучасними технологіями",
      "open_space": "Відкритий простір для спілкування",
      "creative_studio": "Творча студія або майстерня",
      "negotiations": "Переговорна кімната"
    },
    {
      "id": 4,
      "question": "Які завдання вас мотивують найбільше?",
      "options": {
        "technical_problems": "Вирішення складних технічних
проблем",
        "helping_people": "Допомога людям вирішити їхні
проблеми",
        "creating_new": "Створення чогось нового та оригі-
нального",
        "business_goals": "Досягнення бізнес-цілей"
      },
      {
        "id": 5,
        "question": "Як ви ставитесь до навчання нових навичок?",
        "options": {
          "tech_tools": "Люблю вивчати нові технології та
інструменти",
          "psychology": "Цікавлюсь психологією та комунікаці-
ями",
          "creative_skills": "Розвиваю творчі здібності",
          "business": "Вивчаю бізнес та менеджмент"
        }
      }
    }
  }
}

```

```

    }
]

# Дані для аналітики
ANALYTICS_DATA = {
    "stats": {
        "vacancies_analyzed": 8,
        "professions_in_db": 6,
        "tests_completed": 3,
        "accuracy_percentage": 95
    },
    "popular_professions": [
        {"name": "Python розробник", "vacancies": 3},
        {"name": "Frontend розробник", "vacancies": 3},
        {"name": "Test QA Engineer", "vacancies": 2}
    ],
    "salaries": {
        "average": 34375,
        "minimum": 25000,
        "maximum": 40000
    },
    "skills": [
        {"name": "Python", "mentions": 3},
        {"name": "Django", "mentions": 3},
        {"name": "PostgreSQL", "mentions": 3},
        {"name": "React", "mentions": 3},
        {"name": "CSS", "mentions": 3},
        {"name": "JavaScript", "mentions": 3}
    ]
}

```

ВЕБПЛАТФОРМА ПРОФОРІЄНТАЦІЇ З АНАЛІЗОМ РИНКУ ПРАЦІ ТА ПРОГНОЗУВАННЯМ ЗАТРЕБУВАНOSTІ ПРОФЕСІЙ

Володимир СІРЕНКО

КНТ-614

Актуальність дослідження

23%

випускників працюють
за спеціальністю

+340%

зростання попиту
на ІТ-фахівців

<3%

ІТ-професій у традиційних
методиках

Проблема

Традиційні методи профорієнтації (Клімов, Голланд, Голомшток) не враховують сучасні професії в галузі ІТ та не інтегруються з актуальними даними ринку праці

Мета та завдання

МЕТА РОБОТИ

Розробка веб-платформи професійної орієнтації з використанням алгоритму Content-Based Filtering для генерації персоналізованих рекомендацій

ЗАВДАННЯ

- 01 Аналіз предметної області та існуючих систем
- 02 Дослідження та обґрунтування вибору алгоритму
- 03 Розробка математичної моделі рекомендацій
- 04 Проектування та реалізація системи CareerGuide
- 05 Визначення технічних вимог до системи

Обґрунтування вибору алгоритму

КОЛАБОРАТИВНА ФІЛЬТРАЦІЯ

Аналіз схожих користувачів

✗ Проблема холодного старту

CONTENT-BASED FILTERING

Аналіз характеристик

✓ Оптимальний вибір

ГІБРИДНІ МЕТОДИ

Комбінація підходів

✗ Надлишкова складність

ПЕРЕВАГИ CONTENT-BASED FILTERING

→ Незалежність від історії взаємодій

→ Прозорість рекомендацій

→ Детермінованість результатів

→ Відсутність проблеми холодного старту

Математична модель

П'ЯТИВИМІРНИЙ ВЕКТОРНИЙ ПРОСТІР

$$P = (T, C, S, A, B)$$

T Технічна складність

C Креативність

S Соціальна взаємодія

A Аналітичні здібності

B Бізнес-орієнтованість

КОСИНУСНА ПОДІБНІСТЬ

$$\text{sim}(U, P) = (U \cdot P) / (||U|| \times ||P||)$$

Міра відповідності профілю та професії

ФІНАЛЬНА ОЦІНКА

$$S(u, p) = \alpha \times \text{sim} + \beta \times M(p) + \gamma \times A(u, p)$$

sim — косинусна подібність

M(p) — ринкова привабливість

A(u,p) — доступність для користувача

Архітектура системи CareerGuide

REACT 18

TypeScript

FASTAPI

Python

MONGODB

NoSQL

Презентаційний рівень

Бізнес-логіка

Рівень даних

FRONTEND

Тестування користувачів
Візуалізація результатів
Адаптивний дизайн

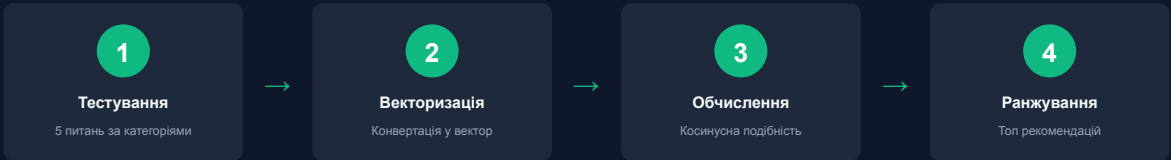
BACKEND

Обробка тестування
Двигун рекомендацій
REST API

DATABASE

Профілі користувачів
База професій
Векторні дані

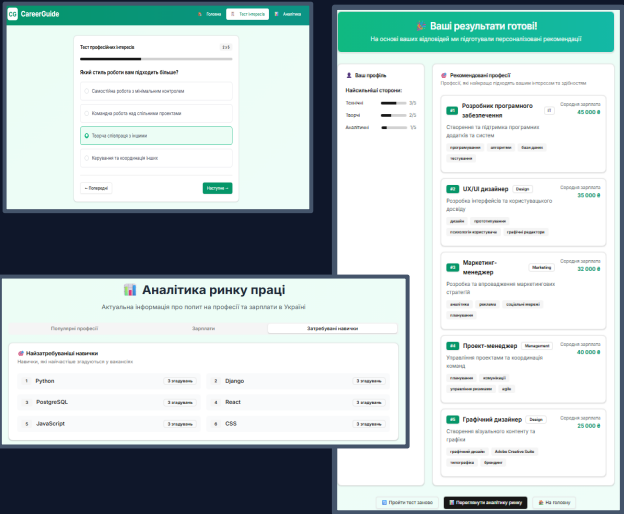
Реалізація алгоритму рекомендацій



ОПТИМІЗАЦІЯ ПРОДУКТИВНОСТІ

- NumPy — векторизовані операції
- Кешування — багаторівневий кеш
- Async — асинхронна обробка

Користувацький інтерфейс



ФУНКЦІОНАЛЬНІСТЬ

- Інтуїтивне тестування
5 питань з візуальним прогресом
- Персоналізовані результати
Топ професій з % відповідності
- Аналітичні дашборди
Статистика ринку праці
- Адаптивний дизайн
Підтримка мобільних пристроїв

Технічні характеристики

МІНІМАЛЬНА

4 ядра CPU

8 GB RAM

100 GB SSD

до 1000 користувачів

РЕКОМЕНДОВАНА

8 ядер CPU

16 GB RAM

500 GB SSD

до 5000 користувачів

КЛАСТЕРНА

16+ ядер CPU

32+ GB RAM

1 TB SSD

необмежено

UPTIME

99.9%

ЛАТЕНТНІСТЬ

<150ms

CDN

✓

DDoS ЗАХИСТ

✓

Висновки

01

Проведено аналіз та виявлено недоліки традиційних систем профорієнтації

02

Обґрунтовано вибір Content-Based Filtering як оптимального підходу

03

Розроблено математичну модель на основі п'ятивимірного векторного простору

04

Реалізовано систему CareerGuide на базі FARM Stack

ПРАКТИЧНА ЦІННІСТЬ

Поєднання наукової обґрунтованості психометричних підходів з адаптивністю та персоналізацією сучасних рекомендаційних алгоритмів для покращення якості професійного самовизначення молоді

**Дякую за
увагу!**

Готовий відповісти на ваші запитання