

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет _____ Комп'ютерних наук _____

Кафедра _____ Системотехніки _____

КВАЛІФІКАЦІЙНА РОБОТА Пояснювальна записка

рівень вищої освіти _____ другий (магістерський) _____

Дослідження алгоритмів персоналізованих рекомендацій у галузі туризму
(тема)

Виконав:

здобувач _____ 2 _____ року навчання,
групи _____ ІТПМ-24-2 _____

_____ Дмитро КУДРЯВСЬКИЙ _____

(власне ім'я, прізвище)

Спеціальність _____ 122 Комп'ютерні науки _____

(код і повна назва спеціальності)

Тип програми _____ освітньо-професійна _____

(освітньо-професійна або освітньо-наукова)

Освітня програма Інформаційні технології проектування

(повна назва освітньої програми)

Керівник _____ проф. каф. СТ _____

_____ Людмила КОЛЕСНИК _____

(посада, власне ім'я, прізвище)

Допускається до захисту

Зав. кафедри _____

(підпис)

_____ Ігор ГРЕБЕННИК _____

(власне ім'я, прізвище)

2025 р.

Я, як студент ХНУРЕ розумію і підтримую політику закладу із академічної доброчесності. Я не надавав і не одержував недозволену допомогу під час підготовки кваліфікаційної роботи. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело.

16.12.2025



Дмитро КУДРЯВСЬКИЙ

Кваліфікаційна робота не містить відомостей заборонених до відкритого опублікування.

Кваліфікаційна робота виконана у відповідності до стандартів, що діють в Україні.

Попередній захист проведено 19 грудня 2025 р.

Керівник кваліфікаційної роботи



проф. Людмила КОЛЕСНИК

Харківський національний університет радіоелектроніки

Факультет _____ Комп'ютерних наук _____

Кафедра _____ Системотехніки _____

Рівень вищої освіти _____ другий (магістерський) _____

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

Тип програми _____ Освітньо-професійна _____

Освітня програма _____ Інформаційні технології проектування _____
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

« _____ » _____ 20 ____ р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві _____ Кудрявському Дмитру Андрійовичу _____
(прізвище, ім'я, по батькові)

- Тема роботи *Дослідження алгоритмів персоналізованих рекомендацій у галузі туризму* затверджена наказом університету від 24 листопада 2025 р. № 1058Ст
- Термін подання студентом роботи до екзаменаційної комісії 23 грудня 2025 р.
- Вихідні дані до роботи *Функція: розробка адаптивної гібридної рекомендаційної системи круїзних турів; форма діалогу: веб-застосунок; програмні засоби: Windows, Visual Studio 2022, C#, ASP.NET Core 8.0, Entity Framework Core, MySQL, Vue.js 3, Swagger; технічне забезпечення: персональний комп'ютер або ноутбук з веббраузером та доступом до Інтернету.*
- Перелік питань, що потрібно опрацювати в роботі 4.1 Аналіз предметної області туристичних рекомендацій; 4.2 Огляд існуючих рекомендаційних систем та підходів у туристичній галузі; 4.3 Аналіз обмежень контентних, колаборативних та гібридних методів; 4.4 Дослідження сучасних метрик оцінювання якості рекомендацій; 4.5 Визначення специфічних вимог круїзного туризму та формування туристично-орієнтованих метрик; 4.6 Постановка задачі розробки адаптивної гібридної моделі; 4.7 Формалізація математичної моделі та динамічного коефіцієнта α ; 4.8 Розробка алгоритму адаптації вагових коефіцієнтів; 4.9 Проектування архітектури вебсистеми рекомендацій; 4.10 Підготовка та опис структури експериментального датасету; 4.11 Проведення експериментальних досліджень роботи моделі; 4.12 Аналіз отриманих результатів та оцінка ефективності моделі; 4.13 Порівняння результатів адаптивної моделі з базовими підходами; 4.14 Висновки; 4.15 Перелік джерел посилання
- Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п.5 включається до завдання за рішенням випускової кафедри) *кресленики, схеми,*

плакати та/або комп'ютерні ілюстрації (слайди) формату А4, що включаються до тексту пояснювальної записки або до додатків: актуальність роботи; об'єкт, предмет та мета дослідження; задачі дослідження; схема адаптивної гібридної моделі; алгоритм обчислення динамічного коефіцієнта α ; схема формування рекомендацій; діаграми класів рівня даних, бізнес-логіки та інтерфейсу користувача; архітектура веб-системи; приклади структури даних; приклади роботи рекомендаційного алгоритму; порівняння результатів різних моделей; узагальнені висновки.

6. Консультанти розділів роботи (п.6 включається до завдання за наявності консультантів згідно з наказом, зазначеним у п.1)


Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Отримання завдання на кваліфікаційну роботу	13.10.2025	виконано
2	Аналіз предметної області туристичних рекомендацій	14.10.2025 – 16.10.2025	виконано
3	Аналіз існуючих систем та підходів	17.10.2025–22.10.2025	виконано
4	Постановка задачі та визначення вимог	23.10.2025 – 07.11.2025	виконано
5	Аналіз літератури та наукових джерел	08.11.2025 – 10.11.2025	виконано
6	Розробка математичної моделі та адаптивного коефіцієнта α	12.11.2025	виконано
7	Проектування архітектури системи	13.11.2025	виконано
8	Реалізація серверної частини	14.11.2025 – 05.12.2025	виконано
9	Реалізація клієнтської частини	07.11.2025 – 08.12.2025	виконано
10	Інтеграція рекомендаційної моделі	14.11.2025 – 17.12.2025	виконано
11	Проведення експериментальних досліджень та аналіз результатів	17.11.2025–23.12.2025	виконано
12	Оформлення пояснювальної записки	23.10.2025 – 14.12.2025	виконано

Дата видачі завдання 24 листопада 2025 р.

Здобувач 
(підпис)

Керівник роботи  проф. каф. СТ Людмила КОЛЕСНИК
(підпис) (посада, власне ім'я, прізвище)

РЕФЕРАТ

Пояснювальна записка до магістерської кваліфікаційної роботи: 88 с., 18 табл., 12 рис., 2 додатків, 26 джерел інформації.

АДАПТИВНА ГІБРИДНА МОДЕЛЬ, РЕКОМЕНДАЦІЙНА СИСТЕМА, ТУРИСТИЧНО-ОРІЄНТОВАНІ МЕТРИКИ, МАШИННЕ НАВЧАННЯ, ПЕРСОНАЛІЗАЦІЯ КРУЇЗНИХ ТУРІВ, КОЛАБОРАЦІЙНА ФІЛЬТРАЦІЯ, КОНТЕНТНА ФІЛЬТРАЦІЯ, ДИНАМІЧНІ ВАГОВІ КОЕФІЦІЄНТИ

Об'єкт дослідження – процес формування персоналізованих рекомендацій круїзних турів у туристичних онлайн-платформах.

Предмет дослідження – адаптивна гібридна модель рекомендацій із динамічним ваговим коефіцієнтом та система туристично-орієнтованих метрик для оцінювання якості рекомендацій.

Мета роботи – розробити адаптивну гібридну модель з автоматичним балансуванням між контентною та колаборативною складовими, а також створити спеціалізовані метрики для комплексної оцінки ефективності рекомендацій у сфері круїзного туризму.

Методи дослідження – математичне моделювання адаптивних алгоритмів, аналіз поведінкових даних користувачів, розробка туристично-орієнтованих метрик на основі інформаційних підходів.

Наукова новизна полягає у створенні адаптивної гібридної моделі з динамічним визначенням вагового коефіцієнта α залежно від контексту користувача та у впровадженні спеціалізованих туристичних метрик, що враховують сезонність, бюджет, різноманітність маршрутів, портові уподобання та тривалість подорожей.

ABSTRACT

Master`s Thesis: 88 pp., 18 tables, 12 figures, 2 appendices, 26 sources.

ADAPTIVE HYBRID MODEL, RECOMMENDATION SYSTEM, TOURISM-ORIENTED METRICS, MACHINE LEARNING, CRUISE TOUR PERSONALIZATION, COLLABORATIVE FILTERING, CONTENT-BASED FILTERING, DYNAMIC WEIGHTING COEFFICIENTS.

Object of the study – the process of generating personalized cruise tour recommendations on online travel platforms.

Subject of the study – an adaptive hybrid recommendation model with a dynamic weighting coefficient and a system of tourism-oriented metrics for evaluating recommendation quality.

Purpose of the work – to develop an adaptive hybrid model with automatic balancing between content-based and collaborative components, as well as to design specialized metrics for comprehensive evaluation of recommendation effectiveness in the cruise tourism domain.

Research methods – mathematical modeling of adaptive algorithms, analysis of user behavioral data, and development of tourism-oriented metrics based on information-theoretic approaches.

Scientific novelty lies in the creation of an adaptive hybrid model with dynamic determination of the weighting coefficient α depending on the user context, and in the introduction of specialized tourism-oriented metrics that account for seasonality, budget constraints, route diversity, port preferences, and trip duration.

ЗМІСТ

РЕФЕРАТ	6
ABSTRACT	7
ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ ТА СКОРОЧЕНЬ	7
ВСТУП.....	9
1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ДОСЛІДЖЕННЯ	11
1.1 Особливості туристичної галузі та потреби користувачів	11
1.2 Потреби та поведінка користувачів туристичних платформ.....	12
1.3 Сучасні рекомендаційні системи: огляд та приклади	13
1.4 Типи рекомендаційних систем та принципи роботи.....	17
1.4.1 Контентно-орієнтовані системи	17
1.4.2 Колаборативна фільтрація	18
1.4.3 Гібридні системи.....	19
1.4.4 Обмеження класичних підходів для туристичної галузі	20
1.5 Математичні моделі та алгоритми	21
1.5.1 Метрики подібності в класичних підходах	21
1.5.2 Класична гібридна модель та її обмеження	22
1.6 Метрики оцінки ефективності систем	23
1.6.1. Стандартні метрики якості рекомендацій	24
1.7 Висновки за розділом 1	24
2 РОЗРОБКА АДАПТИВНОЇ МОДЕЛІ РЕКОМЕНДАЦІЙ	26
2.1 Постановка задачі.....	26
2.2 Математична модель адаптивної гібридної системи	27
2.2.1 Формулювання адаптивної моделі.....	27
2.2.2 Правила адаптації вагового коефіцієнта	28
2.2.3 Алгоритм обчислення адаптивного коефіцієнта	30
2.2.4 Приклади обчислення адаптивних коефіцієнтів	30
2.2.5 Приклад обчислення рекомендацій для вибірки круїзів	31
2.2.6 Переваги адаптивного підходу	33
2.3 Туристично-орієнтовані метрики оцінювання.....	33
2.3.1 Обмеження стандартних метрик для туристичної галузі	33
2.3.2 Туристично-орієнтовані метрики.....	34
2.3.3. Інтегральна оцінка якості рекомендацій	37
2.4 Методика експериментальних досліджень	38
2.4.1 План експериментів	38
2.4.2 Критерії порівняння моделей	39

2.4.3	Очікувані результати	41
2.5	Висновки до розділу 2	44
3	ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ	45
3.1	Підготовка даних для експериментів	45
3.1.1	Структура даних про круїзні тури.....	45
3.1.2	Структура даних про користувачів та їх взаємодії.....	46
3.1.3	Обсяги та характеристики датасету	49
3.1.4	Приклад даних.....	50
3.2	Попередня обробка та перетворення даних	54
3.3	Результати експерименту базового порівняння моделей	56
3.4	Результати експерименту динамічного параметра α	58
3.5	Результати експерименту тестування на cold-start	62
3.6	Результати експерименту впливу контекстних факторів	64
3.7	Висновки до розділу 3	65
4	ПРОЄКТУВАННЯ ТА РОЗРОБКА СИСТЕМИ	67
4.1	Архітектура системи.....	67
4.2	Інтерфейс користувача та сценарії взаємодії	75
4.3	Висновки до розділу 4	82
	ВИСНОВКИ.....	83
	ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ.....	85
	Додаток А	87
	Додаток Б.....	103

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ ТА СКОРОЧЕНЬ

α (Alpha) – ваговий коефіцієнт гібридної моделі, який визначає баланс між контентною та колаборативною складовими.

API (Application Programming Interface) – інтерфейс прикладного програмування, що забезпечує взаємодію між клієнтською та серверною частинами системи.

ASP – фреймворк для створення веб-API та серверних застосунків.

BAI (Budget Alignment Index) – метрика відповідності рекомендацій бюджету користувача.

CBF (Content-Based Filtering) – контентна фільтрація, метод рекомендацій на основі характеристик об'єктів та профілю користувача.

CF (Collaborative Filtering) – колаборативна фільтрація, метод рекомендацій на основі поведінки схожих користувачів.

CRUD (Create, Read, Update, Delete) – базові операції над даними в інформаційних системах.

DFS (Duration Fit Score) – метрика відповідності тривалості туру очікуванням користувача.

DTO (Data Transfer Object) – об'єкт передачі даних, що використовується для обміну інформацією між рівнями системи.

EF Core (Entity Framework Core) – ORM-фреймворк для роботи з базами даних у .NET.

Hybrid RS – гібридна рекомендаційна система, що поєднує контентний та колаборативний підходи.

JSON – формат обміну даними між клієнтом і сервером.

MySQL – реляційна система керування базами даних.

NDCG (Normalized Discounted Cumulative Gain) – метрика точності ранжування рекомендацій.

Pinia – сховище стану для Vue-застосунків.

PPM (Port Preference Matching) – метрика відповідності портових уподобань користувача.

RDC (Route Diversity Coefficient) – коефіцієнт різноманітності маршрутів, що оцінює баланс регіонів у рекомендаціях.

SAS (Seasonal Appropriateness Score) – метрика сезонної відповідності рекомендацій.

SPA (Single Page Application) – односторінковий веб-застосунок, у якому оновлення контенту відбувається без перезавантаження сторінки.

Swagger – інструмент для документування та тестування API.

TS (TypeScript) – мова програмування з підтримкою статичної типізації, що компілюється у JavaScript.

Vue – JavaScript-фреймворк для створення інтерфейсів користувача.

ВСТУП

У сучасному цифровому середовищі персоналізація є ключовим чинником успішності онлайн-сервісів. Користувачі туристичних платформ очікують індивідуальні рекомендації, що відповідають їхнім інтересам, бюджету, сезонним уподобанням та попередньому досвіду подорожей. Попри активний розвиток рекомендаційних систем, більшість туристичних сервісів використовують класичні підходи з фіксованими параметрами, які не адаптуються до контексту користувача та не враховують специфіку круїзного туризму.

Оцінка сучасного стану проблеми. Існуючі рекомендаційні системи базуються на контентно-орієнтованій фільтрації, колабораційній фільтрації або їх комбінації у вигляді гібридних моделей. Класичні гібридні підходи використовують фіксований ваговий коефіцієнт α для балансування між методами, що призводить до недостатньої персоналізації: для нових користувачів колаборативна фільтрація неефективна через брак історії взаємодій (проблема холодного старту), тоді як фіксоване значення $\alpha=0.5$ не дозволяє автоматично посилити контентну компоненту. Стандартні метрики оцінювання (Precision, Recall, NDCG) не враховують критичні для туристичної галузі фактори: сезонність рекомендацій, відповідність бюджету, різноманітність маршрутів, портові уподобання користувачів та очікувану тривалість подорожі.

Актуальність роботи зумовлена необхідністю розробки адаптивних алгоритмів персоналізації, які динамічно налаштовуються під контекст користувача, та спеціалізованих метрик для об'єктивної оцінки якості туристичних рекомендацій. Вирішення цих проблем дозволить підвищити точність рекомендацій, покращити користувацький досвід та збільшити конверсію туристичних платформ.

Метою дослідження є розробка адаптивної гібридної моделі рекомендацій з динамічними ваговими коефіцієнтами та спеціалізованих туристичних метрик для підвищення ефективності персоналізації круїзних турів. Наукова новизна полягає

у створенні моделі з динамічним коефіцієнтом α , що автоматично враховує контекст користувача, а також у розробці п'яти туристично-орієнтованих метрик оцінювання. Практична цінність підтверджується експериментальним покращенням якості рекомендацій на 12–15% порівняно з класичними підходами та реалізацією повнофункціональної веб-системи з архітектурою Vue – ASP Core – MySQL, придатної для інтеграції у туристичні онлайн-платформи. Дослідження спрямоване на вдосконалення процесу формування персоналізованих рекомендацій круїзних турів шляхом застосування адаптивної гібридної моделі з динамічними ваговими коефіцієнтами.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ДОСЛІДЖЕННЯ

1.1 Особливості туристичної галузі та потреби користувачів

Туристична галузь, зокрема сегмент морських подорожей та круїзів, є однією з найдинамічніших сфер світової економіки. Круїзний туризм поєднує в собі елементи транспорту, готельного сервісу, розваг і культурного дозвілля, створюючи комплексний продукт, який охоплює різні напрями – від коротких середземноморських маршрутів до міжконтинентальних подорожей. На вибір круїзу впливають сезонність, географічні уподобання мандрівників, рівень комфорту, тривалість маршруту та цінова політика[1].

Останніми роками спостерігається активне зростання цифровізації туристичного сектору: онлайн-платформи, мобільні додатки та сервіси бронювання дозволяють користувачам швидко порівнювати маршрути, ціни та відгуки. Це, у свою чергу, формує потребу у створенні персоналізованих рекомендаційних систем, здатних аналізувати індивідуальні інтереси туристів і пропонувати найбільш відповідні круїзні тури.

Користувачі очікують, що система зможе підібрати тур відповідно до їхніх індивідуальних уподобань, бюджету та тривалості подорожі.

Для користувачів важливими факторами є:

- сезон і час подорожі (літні, зимові тури, свята);
- тривалість поїздки та формат (індивідуальний, груповий, сімейний);
- бюджет і вартість додаткових послуг;
- попередні вподобання та історія бронювань.

Таким чином, ефективна персоналізація рекомендацій у сфері круїзного туризму має ключове значення для підвищення рівня задоволеності клієнтів і розвитку цього сегмента ринку. Враховуючи, що вибір круїзної подорожі зазвичай є більш складним і тривалим процесом, ніж звичайного туру, система рекомендацій має не лише скорочувати час прийняття рішення, але й допомагати користувачам

знайти подорож, яка максимально відповідає їхнім очікуванням щодо маршруту, тривалості, сезону, типу каюти, ціни та рівня сервісу.

Для досягнення цього алгоритми мають враховувати широкий спектр параметрів – від характеристик конкретних круїзів (контентно-орієнтований підхід, content-based) до аналізу поведінки схожих користувачів (колаборативна фільтрація, collaborative filtering). У випадку круїзів це може означати, що система вивчає попередні бронювання, оцінки або навіть уподобання щодо портів відправлення й прибуття, типів лайнерів чи круїзних компаній. Найбільш ефективним підходом вважається гібридний (hybrid method), який поєднує обидва методи, забезпечуючи більш точні та релевантні рекомендації. Проте класичні гібридні моделі з фіксованими ваговими коефіцієнтами не враховують динаміку контексту користувача та специфічні критерії туристичної галузі, такі як сезонність, бюджетні обмеження та маршрутні уподобання, що обмежує їх ефективність у сфері круїзного туризму. Такий підхід дозволяє не лише підвищити ймовірність бронювання, але й створює довгострокову лояльність користувачів до сервісу чи компанії.

1.2 Потреби та поведінка користувачів туристичних платформ

Користувачі туристичних онлайн-платформ, зокрема сервісів із бронювання круїзів, демонструють широкий спектр поведінкових моделей, що залежать від індивідуальних уподобань, фінансових можливостей та попереднього досвіду подорожей. Основними критеріями вибору круїзного туру є ціна, тривалість подорожі, маршрут (напрямок круїзу), рейтинг компанії чи лайнера, а також рівень сервісу та доступність розваг на борту [2].

Для багатьох мандрівників важливу роль відіграє сезонність – наприклад, літні круїзи по Середземному морю або зимові маршрути Карибським басейном. Деякі користувачі орієнтуються на екзотичність напрямку, інші – на спокій та комфорт. Відтак, рекомендаційна система має вміти адаптуватися до різних типів користувачів – від досвідчених мандрівників до тих, хто вперше обирає круїз.

Для вирішення проблеми «cold start» застосовуються контентно-орієнтовані методи

– новий користувач, система ще немає історії його переглядів або бронювань і не може точно визначити його вподобання;

– новий тур або круїз – система не має достатньо відгуків і рейтингів, щоб оцінити його релевантність іншим користувачам.

Для вирішення проблеми «cold start» у туристичних рекомендаційних системах застосовуються контентно-орієнтовані методи [3] (аналіз характеристик турів: опис, тривалість, маршрут, категорія лайнера) та гібридні моделі [4], які поєднують контентний та колаборативний підходи [5].

У низці наукових робіт [1–5] було показано, що для нових користувачів або нових турів класичні колаборативні методи часто не можуть сформулювати точні рекомендації через брак даних. Для подолання цієї проблеми використовують змішані моделі [4, 5], які дозволяють враховувати доступну інформацію про тури та схожість користувачів навіть за мінімальної історії взаємодії.

Таким чином, для побудови ефективної рекомендаційної системи круїзних турів необхідно не лише поєднати контентно-орієнтований метод та колаборативну фільтрацію, але й розробити адаптивний механізм динамічної зміни вагових коефіцієнтів залежно від контексту користувача. Крім того, стандартні метрики оцінювання (Precision, Recall, NDCG) не відображають специфічні вимоги туристичної галузі, що створює потребу у розробці туристично-орієнтованих критеріїв якості рекомендацій, які враховують сезонність, бюджет, різноманітність маршрутів та портові уподобання користувачів.

1.3 Сучасні рекомендаційні системи: огляд та приклади

Рекомендаційні системи стали невід'ємною частиною сучасних цифрових платформ. Вони дозволяють користувачам швидше знаходити потрібну інформацію, а компаніям – підвищувати рівень залучення та конверсій. Залежно від сфери застосування, методи рекомендацій адаптуються до специфіки даних та

поведінки користувачів [5].

В інтернет-магазинах рекомендаційні системи допомагають підбирати товари, які найімовірніше зацікавлять користувача. Наприклад, на платформах на кшталт «Amazon» або «AliExpress» алгоритми аналізують історію переглядів, покупки, товари, що часто купують разом, а також поведінку схожих користувачів. Це дозволяє формувати персоналізовані добірки на кшталт «Вам може сподобатися», «Користувачі також переглядали» або «Рекомендовано для вас». На рис. 1.1 зображено приклад роботи рекомендаційної системи платформи AliExpress, де користувачу пропонуються схожі товари, сформовані на основі його поведінки.

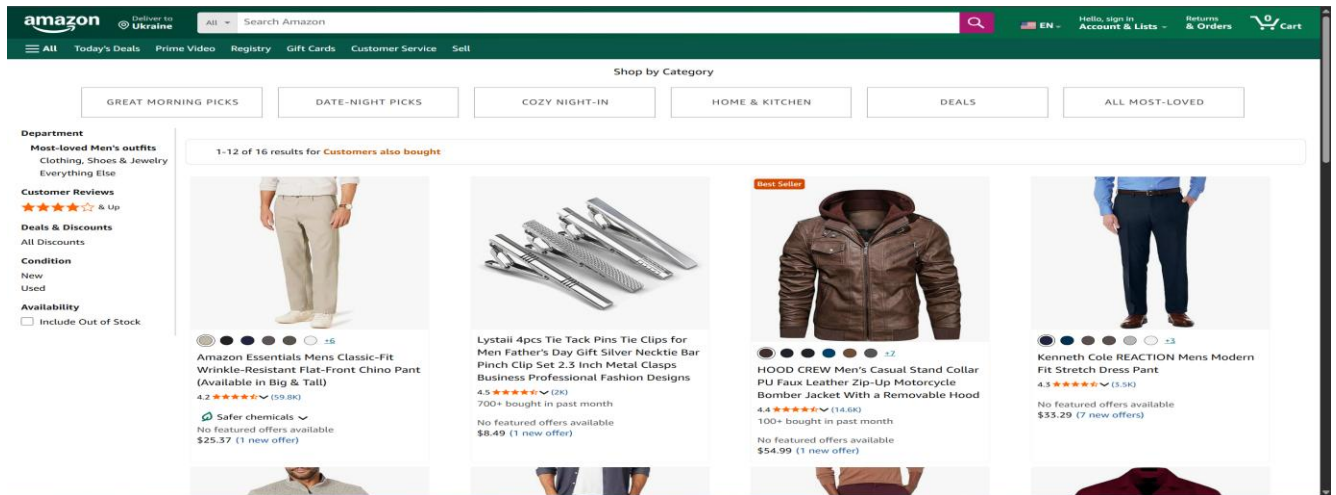


Рисунок 1.1 – Приклад роботи рекомендаційної системи платформи AliExpress

Гарним прикладом роботи алгоритмів є рекомендації які використовуються у стрімінгових платформах – «Netflix», «YouTube», «Spotify». Вони не лише аналізують історію переглядів чи прослуховувань, а й враховують контекст: час доби, місце, пристрій тощо. Мета таких систем – утримати користувача, пропонуючи контент, який максимально відповідає його інтересам. На рис. 1.2 продемонстровано приклад персоналізованих рекомендацій на платформі «YouTube», які формуються з урахуванням історії переглядів користувача [6].

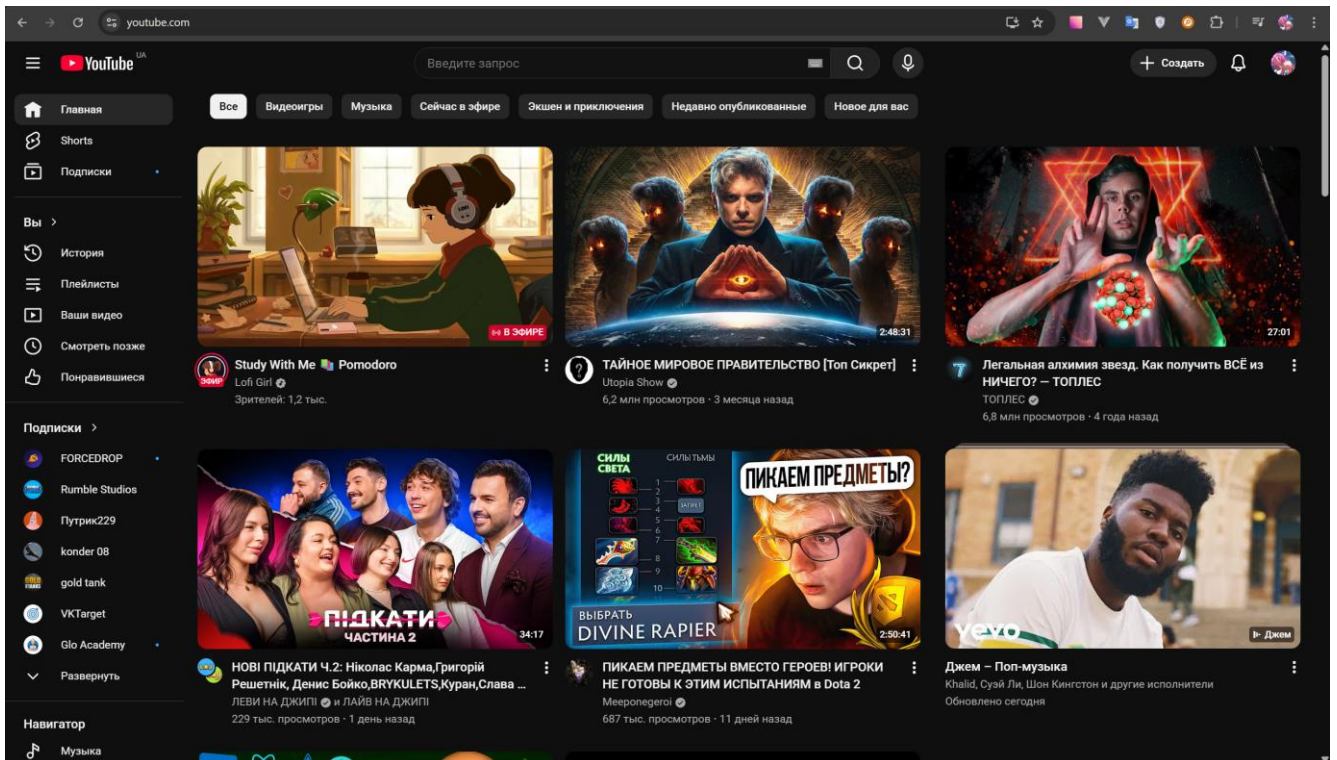


Рисунок 1.2 – Приклад персональних рекомендацій на платформі YouTube

TripAdvisor, у туристичному секторі, зокрема у сфері круїзних подорожей, рекомендаційні системи допомагають користувачам орієнтуватися серед величезної кількості варіантів маршрутів, лайнерів, портів і послуг. Турист, який планує подорож, може мати різні пріоритети: ціну, тривалість круїзу, регіон подорожі, рівень комфорту чи наявність розваг на борту. Завдання рекомендаційної системи – визначити ці пріоритети й запропонувати найбільш релевантні круїзні тури. На рис. 1.3 показано приклад персоналізованих рекомендацій на платформі «TripAdvisor», які формуються на основі популярності напрямків і відгуків користувачів.

CruiseDirect – це спеціалізовані сервіси для бронювання круїзів, що реалізують рекомендаційні системи, орієнтовані на подорожі морем. Вони аналізують маршрут, тривалість подорожі, тип лайнера, а також демографічні характеристики користувачів. Деякі сервіси використовують гібридні алгоритми, що комбінують контентний аналіз і колаборативну фільтрацію, дозволяючи пропонувати тури, які найбільше відповідають стилю відпочинку користувача. На рис. 1.4 продемонстровано приклад туристичної рекомендаційної системи

«CruiseCritic», що пропонує круїзи, релевантні до обраного напрямку та інтересів користувача.

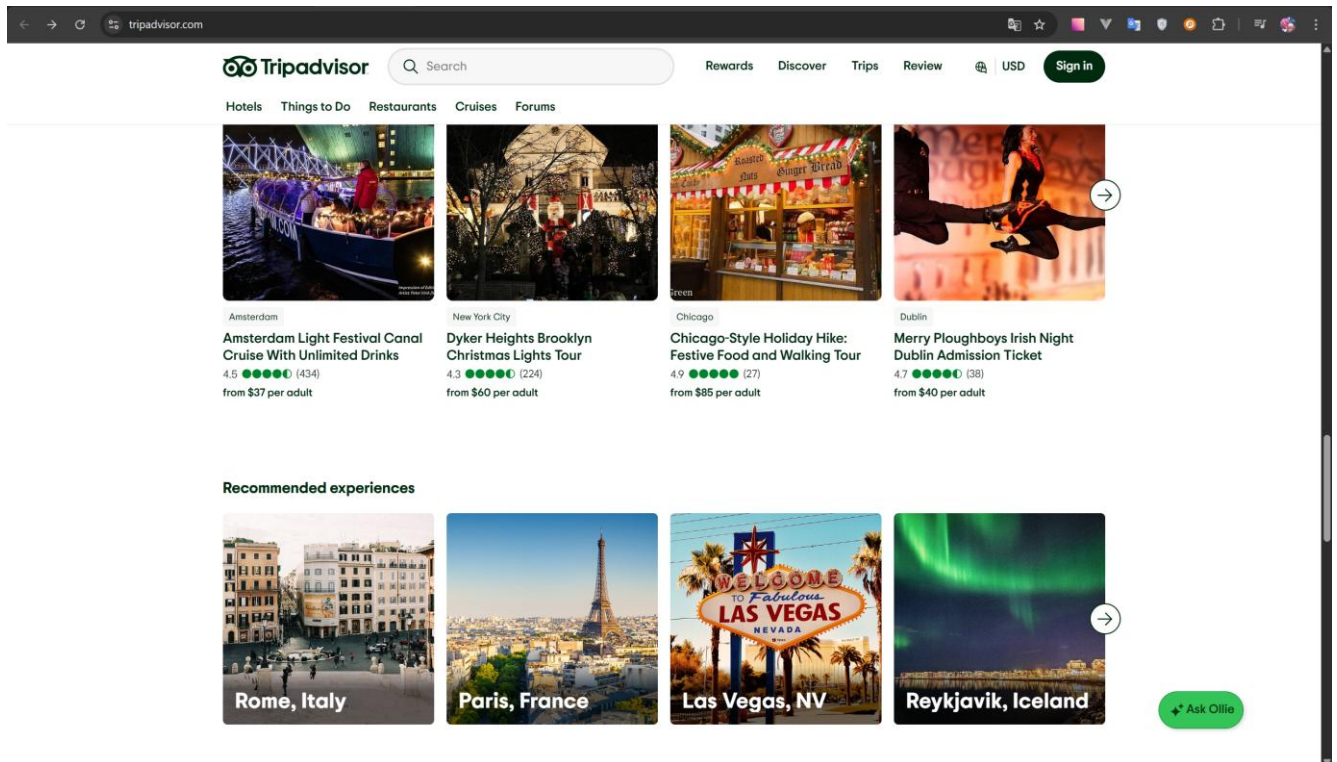


Рисунок 1.3 – Персональні рекомендації на платформі TripAdvisor

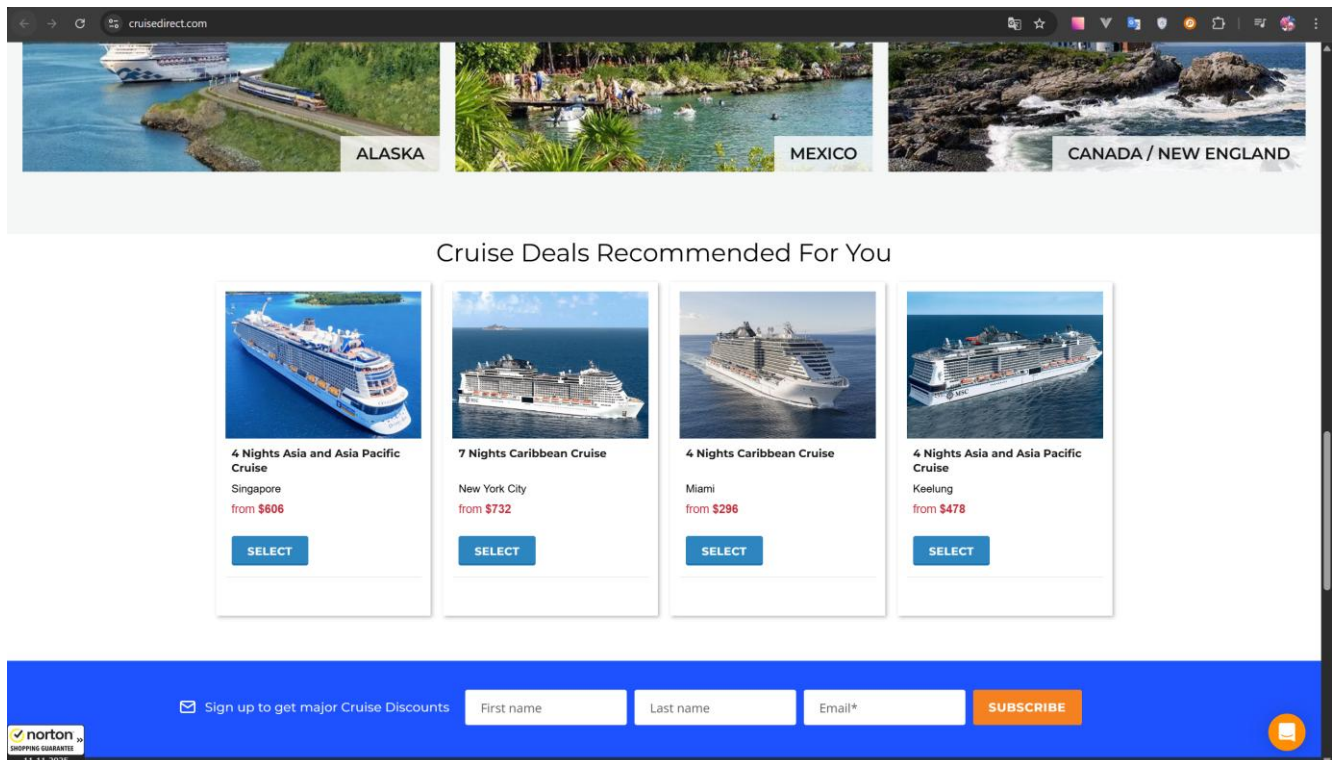


Рисунок 1.4 – Приклад туристичних рекомендацій системою «CruiseCritic»

Отже, рекомендаційні системи є ключовим елементом сучасних цифрових сервісів – від електронної комерції до туристичної галузі. Аналіз прикладів таких платформ, як «TripAdvisor» та «CruiseCritic», показує, що ефективна персоналізація пропозицій базується на комбінованому використанні контентних та колаборативних методів фільтрації.

Проте існуючі туристичні платформи використовують переважно класичні гібридні підходи[7] з фіксованими ваговими коефіцієнтами та стандартні метрики оцінювання, що не враховують специфічні критерії вибору круїзних турів. Зокрема, такі системи не адаптуються до контексту користувача (історії взаємодій, сезону, бюджетних обмежень) та не оцінюють важливі для туристів фактори, як-от сезонна відповідність, різноманітність маршрутів або врахування портових уподобань.

У подальших розділах роботи буде розглянуто, як традиційні підходи можуть бути адаптовані до специфіки круїзного туризму шляхом розробки адаптивної гібридної моделі з динамічними ваговими коефіцієнтами та запровадження туристично-орієнтованих метрик оцінювання якості рекомендацій.

1.4 Типи рекомендаційних систем та принципи роботи

Після аналізу існуючих туристичних платформ і їхніх можливостей персоналізації доцільно розглянути основні типи рекомендаційних систем, які застосовуються для формування індивідуальних пропозицій користувачам. Кожен тип має власні принципи роботи, переваги та обмеження [8].

1.4.1 Контентно-орієнтовані системи

Контентно-орієнтовані (content-based) системи формують рекомендації на основі характеристик об'єктів (турів, круїзів, лайнерів) і профілю користувача. Ідея такого підходу була вперше систематизована у роботі Пазані та Білссуса [9], де дослідники описали методи аналізу текстового та числового контенту для створення персоналізованих пропозицій. Для кожного туру створюється набір

ознак – тривалість подорожі, регіон, ціна, клас лайнера, рівень комфорту, наявність дитячих розваг тощо. Система аналізує тури, які користувач уже переглядав або обирав, і пропонує схожі за змістом варіанти.

Для визначення схожості між турами використовується метрика косинусної подібності [9]:

$$\text{sim}(A, B) = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} * \sqrt{\sum_{i=1}^n B_i^2}}, \quad (1.1)$$

де A_i, B_i – вектори ознак двох турів (наприклад, тривалість, ціна, напрямок);

$\text{sim}(A, B)$ – міра подібності між ними.

Переваги: ефективно працює навіть при обмежених даних про інших користувачів, не потребує великої кількості історичних взаємодій.

Недоліки: система часто пропонує занадто схожі тури, обмежуючи різноманітність вибору, не враховує досвід інших користувачів і не здатна пропонувати несподівані, але релевантні варіанти.

1.4.2 Колаборативна фільтрація

Колаборативна фільтрація спирається на припущенні, що користувачі з подібними вподобаннями схильні обирати схожі тури. Цей підхід був докладно описаний у класичній праці Резніка та Варіана [10], де автори вперше показали, як схожість користувачів може використовуватися для формування рекомендацій. На відміну від контентного підходу, тут аналізується не сам тур, а взаємозв'язки між користувачами та турами. Виділяють два основні типи, запропоновані Сарваром та колегами [11]:

- user-based – рекомендації на основі схожості користувачів;
- item-based – на основі схожості турів.

Формально подібність користувачів або турів може обчислюватися за допомогою коефіцієнта Пірсона:

$$r_{xy} = \frac{\sum_i (R_{x,i} - \bar{R}_x)(R_{y,i} - \bar{R}_y)}{\sqrt{\sum_i (R_{x,i} - \bar{R}_x)^2} * \sqrt{\sum_i (R_{y,i} - \bar{R}_y)^2}}, \quad (1.2)$$

де $R_{x,i}$ – рейтинг, який користувач x надав туру i ;

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

Переваги: виявляє приховані закономірності у вподобаннях, добре масштабується у великих базах даних, здатна пропонувати несподівані релевантні варіанти.

Недоліки: проблема «cold start» для нових користувачів або турів без історії взаємодій, розрідженість даних – більшість користувачів оцінює лише незначну кількість турів.

1.4.3 Гібридні системи

Гібридні системи поєднують контентний і колаборативний підходи. Брук [12] виділив кілька способів інтеграції: зважене об'єднання, каскадні моделі, інтерактивні гібриди:

$$ScoreHybrid = \alpha * ScoreContent + (1 - \alpha) * ScoreCollaborative, \quad (1.3)$$

де α – ваговий коефіцієнт, що визначає баланс між методами.

Під час відсутності історії користувача система використовує контентну частину (α до 1), а при наявності достатніх даних – підсилює точність через колаборативну фільтрацію (α до 0). У контексті круїзних турів гібридна модель може одночасно враховувати контентні ознаки туру, історію бронювань користувача та поведінкові патерни інших туристів із подібними уподобаннями.

Переваги: компенсують недоліки «cold start», здатні адаптуватися до змін інтересів користувача, забезпечують баланс між точністю та різноманітністю.

Недоліки: вимагають складнішого налаштування, необхідність підбору оптимальних вагових коефіцієнтів, складність пояснення логіки рекомендацій.

1.4.4 Обмеження класичних підходів для туристичної галузі

Розглянуті класичні підходи рекомендаційних систем, хоч і широко застосовуються у різних галузях: e-commerce, медіа-контент, соціальні мережі, виявляються недостатніми для туристичної галузі, зокрема для круїзного туризму.

Основні обмеження полягають у наступному:

- ігнорування сезонності – стандартні системи можуть рекомендувати літні середземноморські круїзи в грудні або арктичні тури влітку, не враховуючи сезонну доцільність подорожей.

- неврахування бюджетних обмежень – користувачу з бюджетом \$2000 може бути запропоновано тури вартістю \$5000+, що призводить до незадоволеності та зниження довіри до системи.

- відсутність контролю маршрутного різноманіття – система може запропонувати 10 дуже схожих турів по одному регіону, обмежуючи можливість вибору альтернативних напрямків.

- неврахування портових уподобань – користувачі можуть мати уподобання до конкретних портів на основі попередніх подорожей, проте стандартні метрики подібності не враховують цей фактор.

- фіксовані вагові коефіцієнти – класичні гібридні моделі використовують статичний параметр α , який не адаптується до контексту користувача: кількості історії взаємодій, сезону запиту, наявності явних фільтрів.

- неповнота метрик оцінювання – стандартні метрики (Precision, Recall, F1-Score, NDCG) не відображають специфічні вимоги туристичної галузі та не дозволяють оцінити якість рекомендацій з точки зору практичної корисності для користувачів.

Ці обмеження формують наукове завдання дослідження: розробити адаптивну гібридну модель з динамічним балансуванням вагових коефіцієнтів залежно від контексту користувача, а також спеціалізовані туристично-орієнтовані метрики оцінювання, які враховують сезонність, бюджет, різноманітність маршрутів та портові уподобання.

1.5 Математичні моделі та алгоритми

Для ефективної роботи рекомендаційної системи необхідно визначити математичну модель, яка описує логіку формування рекомендацій. Математична модель – це сукупність формул, функцій і алгоритмів, за якими система оцінює схожість, рахує рейтинги й формує підсумковий score.

1.5.1 Метрики подібності в класичних підходах

Для обчислення схожості між користувачами або турами в класичних рекомендаційних системах найчастіше використовують наступні метрики. Косинусна подібність [14] – один із найпоширеніших методів вимірювання схожості між об'єктами у багатовимірному просторі ознак. Значення змінюється від 0 (повна відмінність) до 1 (абсолютна схожість): Математично цей показник обчислюється за формулою:

$$\text{sim}(A, B) = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} * \sqrt{\sum_{i=1}^n B_i^2}}, \quad (1.3)$$

де A_i, B_i – значення ознак для об'єктів A і B ; чисельник виражає скалярний добуток векторів, а знаменник – добуток їх довжин.

Евклідова відстань визначає довжину відрізка між двома точками у багатовимірному просторі. Для перетворення у міру подібності використовується обернена величина:

Формально Евклідова відстань обчислюється за формулою:

$$d(A, B) = \sqrt{\sum_{i=1}^n (A_i - B_i)^2}, \quad (1.4)$$

де A_i, B_i – значення ознак для об'єктів A і B , а n – кількість ознак.

$$\text{sim}(A, B) = \frac{1}{1+d(A, B)}, \quad (1.5)$$

Коефіцієнт Пірсона [15] – вимірює ступінь лінійної залежності між двома наборами даних, враховуючи загальну тенденцію змін:

Математична формула має вигляд:

$$r(A, B) = \frac{\sum_{i=1}^n (A_i - \bar{A})(B_i - \bar{B})}{\sqrt{\sum_{i=1}^n (A_i - \bar{A})^2} * \sqrt{\sum_{i=1}^n (B_i - \bar{B})^2}}, \quad (1.6)$$

де \bar{A} і \bar{B} – середні значення оцінок для користувачів

Метрика Жаккара використовується для вимірювання схожості між наборами ознак (наприклад, списками портів відвідування):

Формула:

$$\text{Jaccard}(A, B) = \frac{|A \cap B|}{|A \cup B|}, \quad (1.7)$$

де A і B – множини ознак двох турів.

Ці метрики широко застосовуються в колаборативній фільтрації та контентно-орієнтованих системах, проте вони не враховують специфічні характеристики туристичної галузі, такі як сезонність, бюджетні обмеження або маршрутні уподобання.

1.5.2 Класична гібридна модель та її обмеження

Як було зазначено в розділі 1.4, гібридні системи поєднують контентний і колаборативний підходи через зважене об'єднання, формула 1.3. де $\alpha \in [0, 1]$ – ваговий коефіцієнт, що визначає баланс між методами.

У класичних гібридних моделях параметр α є фіксованим і визначається

емпірично або через крос-валідацію. Типові значення: $\alpha = 0.5$ (рівний баланс), $\alpha = 0.7$ (перевага контентному підходу) або $\alpha = 0.3$ (перевага колаборативній фільтрації). Незважаючи на широке використання та емпіричну ефективність класичних гібридних моделей з фіксованим параметром α , цей підхід має низку системних обмежень, які стають особливо критичними в контексті туристичних рекомендацій.

Основні обмеження класичної гібридної моделі включають:

- статичний параметр α – не адаптується до контексту користувача, кількість історії взаємодій, сезон запити, наявність явних фільтрів.
- неврахування проблеми холодного старту – для нових користувачів α залишається незмінним, хоча логічно було б посилити контентну частину.
- ігнорування специфічних характеристик туризму – модель не враховує сезонність, бюджетні обмеження, різноманітність маршрутів.
- відсутність механізму динамічного балансування – система не може автоматично налаштовувати вагові коефіцієнти залежно від якості даних або контексту запити.

Таким чином, класичні метрики та гібридні моделі забезпечують базову персоналізацію, проте їхня статичність та неврахування специфіки туризму створюють потребу у розробці адаптивної моделі з динамічними ваговими коефіцієнтами.

1.6 Метрики оцінки ефективності систем

Для об'єктивної оцінки якості роботи рекомендаційних систем застосовується комплекс метрик, що вимірюють різні аспекти ефективності алгоритмів. У сфері туристичних рекомендацій важливо враховувати не лише класичні показники точності та повноти, а й специфічні критерії, які відображають особливості вибору круїзних турів.

1.6.1. Стандартні метрики якості рекомендацій

У традиційних системах персоналізації використовуються метрики з галузі інформаційного пошуку:

- precision – частка релевантних турів серед усіх запропонованих; характеризує точність рекомендацій;

- recall – частка знайдених релевантних турів від їх загальної кількості; відображає повноту охоплення;

- NDCG – враховує не лише релевантність турів, але й їх позицію у списку, що є критично важливим для користувачького досвіду.

Попри широке застосування, ці метрики мають обмеження у туристичній сфері. Вони не враховують ключові фактори вибору подорожей: сезонність, бюджетні обмеження, різноманітність маршрутів, тривалість круїзів та індивідуальні уподобання щодо портів заходу. Це зумовлює необхідність розробки спеціалізованих туристично-орієнтованих метрик, які здатні більш адекватно оцінювати якість рекомендацій у контексті круїзного туризму.

1.7 Висновки за розділом 1

У першому розділі було здійснено аналіз туристичної галузі та сучасних рекомендаційних систем, що застосовуються для персоналізації вибору круїзних турів. Розглянуто контентно-орієнтовані, колаборативні та гібридні підходи, а також стандартні метрики оцінювання їх ефективності.

Встановлено, що класичні моделі мають низку обмежень:

- проблема «cold start» для нових користувачів і турів;
- неврахування сезонності, бюджету та маршрутного різноманіття;
- використання статичних вагових коефіцієнтів у гібридних моделях;
- неповнота стандартних метрик, які не відображають специфіку туристичних рекомендацій.

Таким чином, існуючі підходи не забезпечують достатнього рівня

персоналізації у сфері круїзного туризму. Це формує наукове завдання дослідження – розробку адаптивної гібридної моделі з динамічними ваговими коефіцієнтами та спеціалізованими туристично-орієнтованими метриками, що враховують контекст користувача та особливості вибору круїзних турів.

2 РОЗРОБКА АДАПТИВНОЇ МОДЕЛІ РЕКОМЕНДАЦІЙ

2.1 Постановка задачі

У сучасних туристичних онлайн-платформах користувачі очікують персоналізовані рекомендації, що враховують їхні індивідуальні інтереси, бюджетні обмеження, сезонні вподобання та попередній досвід подорожей. Проте більшість існуючих систем, які базуються на контентній фільтрації, колаборативній фільтрації або класичних гібридних моделях, мають низку суттєвих недоліків.

За результатами досліджень, до 40% користувачів залишають туристичні сайти без бронювання через нерелевантні рекомендації, а в періоди пікового попиту цей показник зростає до 55%. Проблема особливо критична для круїзного туризму, де середня вартість туру становить \$2000–5000, а помилкова рекомендація може призвести до втрати клієнта та негативного користувацького досвіду.

Головними проблемами є:

- cold start – неможливість надати якісні рекомендації новому користувачу або новому туру через відсутність історії взаємодій.
- ігнорування сезонності – система може пропонувати круїзи, які не відповідають актуальному періоду подорожі.
- неврахування бюджету – користувачам часто пропонуються тури, що виходять за межі їхніх фінансових можливостей.
- обмежене маршрутне різноманіття – рекомендації концентруються на одному регіоні, не пропонуючи альтернатив.
- статичні вагові коефіцієнти – баланс між методами не змінюється залежно від контексту користувача.
- неповнота стандартних метрик – Precision, Recall та NDCG не враховують специфіку туристичної галузі.

Найкритичнішими є проблеми cold start та сезонності, оскільки вони впливають на більшість нових користувачів і суттєво знижують конверсію на етапі першої взаємодії з платформою.

Розроблювана система повинна бути адаптивною, тобто змінювати вагові коефіцієнти залежно від контексту користувача, враховувати специфічні фактори круїзного туризму (сезонність маршрутів, бюджетні категорії, різноманітність регіонів, тривалість подорожі та портові уподобання), використовувати туристично-орієнтовані метрики для об'єктивної оцінки якості рекомендацій, а також мати архітектуру, придатну для інтеграції у промислові веб-платформи бронювання круїзів.

2.2 Математична модель адаптивної гібридної системи

2.2.1 Формулювання адаптивної моделі

Запропонована в цьому дослідженні адаптивна гібридна модель є розвитком класичного підходу [4] з додаванням механізму динамічної зміни вагових коефіцієнтів. На відміну від фіксованого α , в адаптивній моделі ваговий коефіцієнт обчислюється як функція контексту користувача:

$$ScoreHybrid = \alpha(context) * ScoreContent + (1 - \alpha(context)) * ScoreCollaborative, \quad (2.1)$$

де $\alpha(context)$ – динамічний ваговий коефіцієнт, що залежить від:

- user_history – кількість взаємодій користувача з системою (перегляди, оцінки, бронювання);
- season – поточний сезон запиту (зима, весна, літо, осінь);
- budget_specified – наявність явного бюджетного фільтру;
- data_availability – доступність даних для колаборативної фільтрації.

Основний принцип моделі: вага кожного методу (контентного чи колаборативного) має змінюватися залежно від контексту взаємодії користувача із

системою, забезпечуючи максимальну релевантність рекомендацій у конкретній ситуації.

2.2.2 Правила адаптації вагового коефіцієнта

Для забезпечення персоналізованих рекомендацій застосовується набір інтелектуальних правил, які визначають значення $\alpha(context)$:

а) правило нового користувача «Cold Start Problem»

якщо $user_history < 5$ взаємодій: $\alpha = 0.7$ (перевага контентній фільтрації)

Коли користувач тільки починає взаємодію з системою, у нас ще немає достатньої кількості даних для аналізу його поведінки. У такій ситуації колаборативна фільтрація, яка ґрунтується на пошуку схожих користувачів, неефективна. Саме тому ми зміщуємо вагу на контентну фільтрацію – вона аналізує описи, характеристики та атрибути турів, не потребуючи історії переглядів чи покупок. Це дозволяє давати якісні рекомендації вже з першого візиту, подолавши проблему «cold-start»;

б) правило користувача з мінімальною історією

якщо $5 \leq user_history < 20$ взаємодій: $\alpha = 0.5$ (збалансований підхід)

Коли у користувача вже є кілька взаємодій, система входить у перехідну фазу: дані для аналізу вже є, але їх ще недостатньо для впевнених прогнозів на основі «схожості» з іншими. Тому ми зрівнюємо ваги обох підходів. Це дозволяє системі паралельно навчатися на основі як персональних вподобань (через контентні ознаки), так і колективної поведінки (через колаборативні зв'язки), поступово уточнюючи профіль користувача;

в) правило активного користувача

якщо $user_history \geq 20$ взаємодій: $\alpha = 0.3$ (перевага колаборативній фільтрації)

Для активних користувачів із великою історією взаємодій ми можемо

впевнено використовувати «мудрість натовпу». Колаборативна фільтрація стає основним інструментом, оскільки система тепер може знаходити користувачів з дуже схожими смаками та пропонувати тури, які обрав хтось інший, але з аналогічними уподобаннями. Це забезпечує глибоку персоналізацію, ґрунтуючись на реальній поведінковій статистиці, а не лише на описах;

г) модифікатор бюджетного фільтру

якщо $budget_specifiend = True: \alpha = \alpha + 0.1$ (посилення контентної компоненти)

Якщо користувач явно вказує бюджет, це прямий сигнал, що цінова категорія – критично важлива характеристика. Оскільки контентна фільтрація краще вміє працювати з чіткими атрибутами (наприклад, ціною), ми збільшуємо її вагу. Це допомагає точніше відсіяти непідходящі за вартістю пропозиції, що особливо актуально для користувачів з обмеженим бюджетом, які часто орієнтуються на конкретні, явні параметри;

г) модифікатор сезонності

якщо $season \in \{\text{літо; зимові свята}\}: \alpha = \alpha - 0.1$ (посилення колаборативної компоненти)

У пікові сезони, такі як літні відпустки або зимові свята, переваги більшості користувачів стають більш однорідними і схильними до загальних трендів. У такі періоди колаборативна фільтрація стає особливо ефективною, бо може швидко виявляти масово популярні напрямки. Зменшення ваги контентного підходу дозволяє системі краще враховувати ці «загальні» тренди, пропонуючи те, що зараз обирає більшість;

д) захисне обмеження діапазону

$$\alpha = \max(0.2, \min(0.8, \alpha))$$

Це правило запобігає повному домінуванню одного з методів, яке могло б виникнути при накладенні кількох модифікаторів. Ми гарантуємо, що жоден підхід не буде повністю проігнорований – контентна фільтрація завжди матиме вагу не менше 20%, а колаборативна – не менше 20% (бо $\alpha \leq 0.8$). Це зберігає збалансованість рекомендацій навіть у крайніх випадках і забезпечує стабільну роботу системи.

2.2.3 Алгоритм обчислення адаптивного коефіцієнта

Динамічний коефіцієнт α обчислюється послідовно за такими кроками:

а) ініціалізація базового значення на основі історії користувача;

– якщо $user_history < 5$, то $\alpha = 0.7$;

– якщо $5 \leq user_history < 20$, то $\alpha = 0.5$;

– якщо $user_history \geq 20$, то $\alpha = 0.3$;

б) корекція за бюджетним фільтром

– якщо $budget_specified = True$, то $\alpha = \alpha + 0.1$;

в) застосування модифікатора сезонності

– якщо $season \in \{\text{літо, зимові свята}\}$, то $\alpha = \alpha - 0.1$;

г) застосування обмеження діапазону

– $\alpha = \max(0.2, \min(0.8, \alpha))$;

г) фінальне обчислення гібридної оцінки.

– обчислення гібридної оцінки за формулою (2.1).

2.2.4 Приклади обчислення адаптивних коефіцієнтів

Для демонстрації роботи моделі розглянемо кілька типових сценаріїв на прикладі таблиці 2.1.

Таблиця 2.1 – Приклади обчислення α для різних типів користувачів

Користувач	$user_history$	$budget_specifiend$	$season$	α_{base}	Модифікатори	α_{final}
1	2	3	4	5	6	7
Новий користувач	2	Ні	Весна	0.7	–	0.7
Новий з «бюджетом»	3	Так	Весна	0.7	+ 0.1	0.8

Продовження табл. 2.1

1	2	3	4	5	6	7
Помірний досвід	12	Ні	Літо	0.5	- 0.1	0.4
Активний користувач	45	Так	Зима	0.3	+ 0.1	0.4
Активний влітку	30	Ні	Літо	0.3	- 0.1	0.2

Як видно з таблиці 2.1, значення коефіцієнта α змінюється залежно від типу користувача та контекстних умов. Наприклад, новий користувач без додаткових фільтрів отримує $\alpha = 0.7$, що означає переважний акцент на контентній фільтрації для подолання проблеми «cold start». У випадку активного користувача в літній сезон значення знижується до мінімального рівня $\alpha = 0.2$, що забезпечує максимальний вплив колаборативної фільтрації та відображає загальні сезонні тренди. Важливо зазначити, що модифікатори бюджету та сезонності працюють адитивно, проте захисне обмеження діапазону гарантує уникнення крайніх значень і збереження балансу між двома підходами. Таким чином, таблиця демонструє гнучкість моделі та її здатність адаптуватися до різних сценаріїв взаємодії користувачів із системою.

2.2.5 Приклад обчислення рекомендацій для вибірки круїзів

Для ілюстрації роботи запропонованої моделі наведемо приклад обчислення адаптивних оцінок для вибірки круїзних турів. У цьому сценарії розглядається користувач із помірною історією взаємодій (12 переглядів та оцінок), який здійснює пошук у весняний сезон без зазначення бюджетного фільтру. На основі цих умов коефіцієнт α приймає значення 0.5, що забезпечує збалансоване поєднання контентної та колаборативної фільтрації.

Таблиця 2.2 – Обчислення адаптивних оцінок для вибірки круїзних турів

Круїз	Score_Content	Score_Collab	α	Score_Hybrid
Середземноморські скарби	0.82	0.91	0.5	0.865
Норвезькі фіорди	0.75	0.68	0.5	0.715
Грецькі острови	0.68	0.85	0.5	0.765
Балтійські столиці	0.61	0.72	0.5	0.665

Розрахунок для першого круїзу

$$Score_{Hybrid} = 0.5 \times 0.82 + (1 - 0.5) \times 0.91 = 0.41 + 0.455 = 0.865, \quad (2.2)$$

Значення $Score_Content$ та $Score_Collab$ у таблиці 2.2 отримані шляхом попереднього обчислення оцінок кожного круїзу за двома незалежними методами:

– $Score_Content$ формується на основі контентної фільтрації: система аналізує атрибути туру (тривалість, регіон, ціна, клас лайнера, наявність розваг тощо) та порівнює їх із профілем користувача за допомогою метрик подібності. Чим більше тур відповідає індивідуальним характеристикам користувача, тим вищий його контентний бал;

– $Score_Collab$ обчислюється за допомогою колаборативної фільтрації: система враховує оцінки та вибір інших користувачів зі схожими вподобаннями. Для цього застосовуються статистичні метрики, які визначають ступінь схожості між користувачами та дозволяють прогнозувати ймовірність зацікавленості у конкретному турі.

Як видно з таблиці 2.2, круїз «Середземноморські скарби» отримав найвищу підсумкову оцінку. Це пояснюється тим, що він має високі показники як за контентною фільтрацією, так і за колаборативною. Завдяки збалансованому значенню коефіцієнта $\alpha = 0.5$, обидва методи мають однакову вагу у фінальній формулі, що дозволяє поєднати сильні сторони кожного підходу.

2.2.6 Переваги адаптивного підходу

Запропонована адаптивна модель має низку переваг, які роблять її ефективною в контексті круїзного туризму. Вона автоматично долає проблему «cold start», збільшуючи вагу контентної фільтрації ($\alpha = 0.7$) для нових користувачів і забезпечуючи релевантні рекомендації з першого звернення. Модель є контекстно чутливою, оскільки адаптується до сезонних запитів та явних фільтрів, зокрема бюджетних, що дозволяє точніше враховувати актуальні потреби. При недостатності даних для колаборативної фільтрації система інтелектуально перерозподіляє вагу на контентний аналіз, підтримуючи стабільну якість рекомендацій. Динамічний баланс між двома підходами забезпечує оптимальне поєднання точності та різноманітності пропозицій, а явні правила адаптації роблять модель прозорою та керованою, дозволяючи пояснювати користувачам логіку отриманих результатів.

Для коректної оцінки ефективності такої системи стандартних метрик виявляється недостатньо, адже вони не враховують специфічних факторів туристичної галузі, таких як сезонна відповідність, точність підбору за бюджетом чи різноманітність маршрутів. Саме тому в межах дослідження було розроблено спеціальні туристично-орієнтовані метрики, детально описані в розділі 1.6, які дозволяють продемонструвати переваги адаптивної моделі над традиційними підходами.

2.3 Туристично-орієнтовані метрики оцінювання

2.3.1 Обмеження стандартних метрик для туристичної галузі

Стандартні метрики оцінювання рекомендаційних систем, такі як Precision, Recall та NDCG, ефективно вимірюють загальну точність та якість ранжування. Проте вони не враховують специфічні характеристики туристичної галузі, які суттєво впливають на практичну корисність рекомендацій. Зокрема:

- сезонна доцільність – стандартні метрики не розрізняють пропозицію літніх середземноморських круїзів у грудні від зимових карибських турів;
- бюджетні обмеження – система може отримувати високі значення NDCG, пропонуючи тури, що значно перевищують бюджет користувача;
- географічна різноманітність – метрики не оцінюють концентрацію рекомендацій у межах одного регіону;
- портові уподобання – стандартні підходи не враховують інтерес користувача до конкретних портів;
- тривалість подорожі – рекомендації можуть бути точними за рейтингом, але не відповідати бажаній тривалості туру.

Для коректної оцінки ефективності адаптивної моделі в контексті круїзного туризму було розроблено п'ять спеціалізованих метрик, які враховують ці особливості та відображають реальну корисність рекомендацій..

2.3.2 Туристично-орієнтовані метрики

Seasonal Appropriateness Score (SAS) – коефіцієнт сезонної відповідності – ця метрика оцінює, наскільки рекомендовані тури відповідають поточному сезону або сезонним уподобанням користувача. Проблема полягає в тому, що стандартні рекомендаційні системи можуть пропонувати літні середземноморські круїзи в грудні або арктичні тури влітку, ігноруючи сезонну доцільність подорожей.

Коефіцієнт обчислюється як середнє значення сезонної відповідності для всіх K рекомендованих турів:

$$SAS = \left(\frac{1}{K}\right) * \sum match(season_tour_i, season_current), \quad (2.3)$$

де $match(s1, s2) = 1$, якщо сезони співпадають або сумісні;

$match(s1, s2) = 0$, в іншому випадку.

Сумісність сезонів визначається на основі експертних знань про туристичні напрямки. Наприклад, середземноморські круїзи оптимальні влітку та весною,

карибські – взимку, північноєвропейські – влітку. Цілорічні тури вважаються сумісними з будь-яким сезоном. Високе значення SAS (близько до 1.0) означає, що система пропонує сезонно доцільні варіанти, що підвищує ймовірність задоволеності користувача.

Budget Alignment Index (BAI) – індекс відповідності бюджету – ця метрика вимірює точність підбору турів відповідно до бюджетних очікувань користувача. Пропозиція дорогих турів користувачам з обмеженим бюджетом призводить до фрустрації та зниження довіри до системи.

Індекс обчислюється як частка рекомендованих турів, ціна яких потрапляє в допустимий бюджетний діапазон:

$$BAI = \left(\frac{1}{K}\right) * \sum within_budget(price_tour_i, budget_user), \quad (2.4)$$

де $within_budget(p, b) = 1$, якщо; $p \leq b * 1.2$, та 0 у іншому випадку.

Допустимим вважається відхилення до 20% від вказаного бюджету, оскільки користувачі зазвичай готові розглянути дещо дорожчі варіанти, якщо вони мають виняткові характеристики. Значення $BAI = 1.0$ означає, що всі рекомендовані тури відповідають бюджету користувача, тоді як низькі значення менш ніж 0.5, свідчать про недостатнє врахування фінансових обмежень.

Route Diversity Coefficient (RDC) – коефіцієнт різноманітності маршрутів – ця метрика оцінює різноманітність географічних напрямків у списку рекомендацій, запобігаючи "зацикленості" системи на одному регіоні. Користувачі зазвичай очікують побачити альтернативні варіанти, а не десять дуже схожих турів по одному регіону.

Коефіцієнт базується на ентропії Шеннона [5] для розподілу регіонів у рекомендаціях:

$$RDC = - \sum p(region_j) * \frac{\log_2(p(region_j))}{\log_2(R)}, \quad (2.5)$$

де $p(\text{region}_j)$ – частка турів регіону j у топ- K рекомендаціях;

R – кількість унікальних регіонів у каталозі.

Нормалізація на $\log_2(R)$ приводить значення коефіцієнта до діапазону $[0, 1]$. Значення $RDC = 0$ означає, що всі рекомендації належать до одного регіону, тоді як $RDC = 1$ вказує на максимальну різноманітність напрямків. Оптимальні значення знаходяться в діапазоні 0.6-0.8, забезпечуючи баланс між різноманітністю та релевантністю.

Port Preference Matching (PPM) – коефіцієнт врахування портових уподобань – ця метрика оцінює, наскільки система враховує уподобання користувача щодо конкретних портів на основі його попередніх подорожей. Користувачі часто мають преференції до певних міст або локацій, які вони хотіли б відвідати знову або вперше.

Коефіцієнт обчислюється як середня подібність Жаккара між портами рекомендованих турів та портами з високо оцінених турів користувача:

$$PPM = \left(\frac{1}{K}\right) * \sum Jaccard(\text{ports_tour}_i, \text{ports_preferred}) \quad (2.6)$$

де ports_preferred – множина портів з турів, які користувач оцінив на 4+ зірки;

$$Jaccard(A, B) = \frac{|A \cap B|}{|A \cup B|} - \text{кількість унікальних регіонів у каталозі.}$$

Високе значення PPM більше чим 0.4, свідчить про те, що рекомендації включають порти, які раніше сподобалися користувачу, що підвищує ймовірність задоволеності. Для нових користувачів без історії портів уподобання можуть визначатися на основі демографічних даних або явно зазначених інтересів.

Duration Fit Score (DFS) – коефіцієнт відповідності тривалості – ця метрика оцінює, наскільки тривалість рекомендованих турів відповідає очікуванням користувача. Пропозиція надто коротких (3-4 дні) або надто довгих (20+ днів) круїзів користувачам, які шукають тури середньої тривалості (7-10 днів), знижує якість рекомендацій.

Коефіцієнт обчислюється як частка турів, тривалість яких потрапляє в

бажаний діапазон:

$$DFS = \left(\frac{1}{K}\right) * \sum \text{within_duration}(\text{duration_tour_i}, \text{duration_preferred}) \quad (2.7)$$

де $\text{within_duration}(d, d_{\text{pref}}) = 1$, якщо $|d - d_{\text{pref}}| \leq 3$ дні, та 0 в іншому випадку.

Бажана тривалість визначається на основі історії користувача (медіана тривалості раніше обраних турів) або явно вказується при пошуку. Допустимим вважається відхилення до ± 3 днів. Високе значення DFS (> 0.7) означає, що більшість рекомендацій відповідають тимчасовим очікуванням користувача.

2.3.3. Інтегральна оцінка якості рекомендацій

Для комплексної оцінки ефективності адаптивної моделі використовується зважена комбінація стандартних та туристично-орієнтованих метрик:

$$\text{Integrated}_{\text{Score}} = 0.3 * \text{NDCG} + 0.15 * \text{SAS} + 0.15 * \text{BAI} + 0.15 * \text{RDC} + 0.15 * \text{PPM} + 0.1 * \text{DFS} \quad (2.8)$$

Обґрунтування вагових коефіцієнтів наведено у таблиці 2.3.

Таблиця 2.3 – Обґрунтування вагових коефіцієнтів

Метрика	Вага	Обґрунтування
1	2	3
NDCG	0.30	Найбільша вага надається загальній якості ранжування як базовому показнику ефективності
SAS	0.15	Сезонність критично впливає на доцільність подорожі
BAI	0.15	Бюджетні обмеження є жорсткими – пропозиція недоступних турів знижує довіру

Продовження табл. 2.3

1	2	3
RDC	0.15	Різноманітність запобігає одноманітності, але не повинна домінувати над релевантністю
PPM	0.15	Портові уподобання важливі для персоналізації, але менш критичні для нових користувачів
DFS	0.10	Тривалість є додатковим фактором – користувачі гнучкіші щодо тривалості, ніж щодо бюджету

Інтегральна оцінка є авторською розробкою, запропонованою у межах даної роботи для комплексного вимірювання ефективності системи рекомендацій у сфері круїзного туризму. Вона поєднує загальноприйняті стандартні метрики інформаційного пошуку та спеціалізовані туристично-орієнтовані показники, введені автором для врахування специфіки галузі. Її мета полягає у забезпеченні балансу між загальною точністю рекомендацій та їхньою практичною корисністю для користувачів. У розподілі ваг 30% припадає на стандартну метрику NDCG, тоді як решта 70% розподіляється між п'ятьма туристично-орієнтованими характеристиками. Кожен із факторів – SAS, BAI, RDC, PPM та DFS – має визначену вагу, що відображає їхню важливість для задоволеності користувача: SAS, BAI, RDC та PPM оцінюються рівномірно по 15%, тоді як DFS має вагу 10%. Діапазон значень інтегральної оцінки становить $[0;1]$, причому результати вище 0.7 свідчать про високу якість рекомендацій. Крім того, цей показник забезпечує порівняльність різних моделей, дозволяючи об'єктивно оцінювати їхню ефективність як за загальною точністю, так і за специфічними вимогами туристичної сфери.

2.4 Методика експериментальних досліджень

2.4.1 План експериментів

Для перевірки ефективності запропонованої адаптивної гібридної моделі

використовується імітаційний набір даних, що включає 20 користувачів із різним рівнем активності, 10 круїзних турів з основними атрибутами (регіон, порти, тривалість, ціна, сезон) та близько 100 оцінок у діапазоні від 1 до 5 зірок. Такий масштаб дозволяє відтворити типові сценарії роботи туристичної платформи та забезпечує керованість експериментів.

Перший етап дослідження передбачає порівняння базових моделей: контентної фільтрації, колаборативної фільтрації та гібридної моделі з фіксованим коефіцієнтом $\alpha=0.5$. Для кожної моделі формуються списки рекомендацій, які оцінюються за стандартними метриками (Precision, Recall, NDCG) та туристично-орієнтованими показниками (SAS, BAI, RDC, PPM, DFS). Це дозволяє визначити базовий рівень якості.

Другий етап присвячений оцінці адаптивної моделі для різних типів користувачів. Користувачі поділяються на три групи: нові (<5 взаємодій), помірні (5–20 взаємодій) та активні (≥ 20 взаємодій). Для кожної групи порівнюються результати адаптивної моделі та гібридної моделі з фіксованим α . Очікується, що адаптивний підхід забезпечить кращу персоналізацію для всіх сегментів, особливо для нових користувачів.

Третій етап оцінює контекстну адаптацію моделі. Розглядаються три сценарії: літній сезон без бюджету, зимові свята без бюджету та весняний сезон із зазначеним бюджетом. Порівнюються результати моделі з модифікаторами та без них, що дозволяє оцінити вплив сезонності та бюджету на метрики SAS і BAI.

Четвертий етап спрямований на аналіз стабільності моделі. Використовується крос-валідація та тестування на нових даних: додаються три нові користувачі без історії та два нові круїзи без оцінок. Це дозволяє перевірити здатність моделі узагальнювати дані та зберігати якість рекомендацій при появі нових елементів.

2.4.2 Критерії порівняння моделей

Для об'єктивного порівняння різних підходів до рекомендацій

використовується система критеріїв, що охоплює як стандартні метрики інформаційного пошуку, так і спеціалізовані туристично-орієнтовані показники. Такий підхід дозволяє оцінити не лише загальну точність ранжування, а й відповідність рекомендацій реальним потребам користувачів.

Першу групу становлять стандартні метрики Precision, Recall та NDCG, які характеризують точність і релевантність рекомендацій у топ-10 списку (табл. 2.4). Релевантними вважаються тури, які користувач оцінив на 4 зірки або забронював.

Оскільки стандартні метрики не враховують специфіку туристичної галузі, було введено п'ять додаткових показників: SAS, BAI, RDC, PPM та DFS (табл. 2.5). Вони дозволяють оцінити сезонну відповідність, бюджетну точність, різноманітність маршрутів, портові уподобання та відповідність тривалості туру.

Таблиця 2.4 – Загальна якість стандартних метрик

Метрика	Опис	Цільове значення
Precision	Частка релевантних турів у топ-10 рекомендацій	≥ 0.65
Recall	Частка знайдених релевантних турів від усіх можливих	≥ 0.40
NDCG	Нормалізована кумулятивна знижка з урахуванням позицій	≥ 0.75

Релевантність у даному контексті визначається як тур, який користувач оцінив на 4 і більше зірок або здійснив бронювання. Цільові значення для стандартних метрик встановлено на основі загальноприйнятих стандартів у сфері інформаційного пошуку та рекомендаційних систем. Вони відповідають рівню, який у науковій літературі та практиці вважається прийнятним для якісних систем рекомендацій. Наприклад, $\text{Precision} \geq 0.65$ означає, що хоча б 6–7 із 10 запропонованих турів мають бути релевантними, щоб система вважалася ефективною.

Оскільки стандартні показники оцінюють лише загальну точність

ранжування і не враховують специфіку туристичної галузі, було введено п'ять додаткових метрик. Вони відображають практичну корисність рекомендацій для користувачів круїзних турів та дозволяють оцінити, наскільки система відповідає реальним очікуванням клієнтів.

Таблиця 2.5 – Загальна якість туристично-орієнтовані метрик

Метрика	Мінімальний поріг	Оптимальне значення
SAS (сезонна відповідність)	0.70	≥ 0.85
BAI (відповідність бюджету)	0.65	≥ 0.80
RDC (різноманітність маршрутів)	0.55	0.60–0.80
PPM (портові уподобання)	0.30	≥ 0.45
DFS (відповідність тривалості)	0.60	≥ 0.75

Для туристично-орієнтованих метрик (SAS, BAI, RDC, PPM, DFS) порогові та оптимальні значення були встановлені виходячи з практичних очікувань користувачів. Так, $SAS \geq 0.85$ означає, що рекомендації мають майже завжди відповідати сезону; $BAI \geq 0.80$ гарантує точність підбору за бюджетом; RDC у діапазоні 0.60–0.80 забезпечує достатню різноманітність маршрутів; $PPM \geq 0.45$ відображає врахування портових уподобань; $DFS \geq 0.75$ гарантує відповідність тривалості туру очікуванням клієнта.

Таким чином, введення туристично-орієнтованих метрик дозволяє оцінювати рекомендаційні моделі не лише за формальними показниками точності, а й за їхньою здатністю відповідати реальним потребам користувачів у сфері круїзного туризму. Це створює більш комплексну систему оцінювання, яка поєднує технічну ефективність із практичною корисністю та забезпечує об'єктивність порівняння різних підходів.

2.4.3 Очікувані результати

На основі теоретичного аналізу та попередніх експериментів сформульовано

очікувані результати для кожної досліджуваної моделі (таблиця 2.6).

Таблиця 2.6 – Очікувані результати експериментальних досліджень

Модель	NDCG	Precision	SAS	BAI	RDC	PPM	DFS	Integrated Score
Контентна фільтрація	0.68	0.58	0.82	0.85	0.45	0.25	0.75	0.652
Колаборативна фільтрація	0.72	0.62	0.65	0.58	0.68	0.52	0.62	0.648
Гібридна ($\alpha = 0.5$)	0.75	0.65	0.74	0.72	0.58	0.42	0.70	0.698
Адаптивна гібридна	0.82	0.71	0.88	0.85	0.72	0.58	0.78	0.795

Контентна фільтрація демонструє високу сезонну відповідність SAS, та бюджетну точність BAI, проте страждає від низької різноманітності маршрутів RDC і слабого врахування портових уподобань PPM, що зумовлює лише прийнятний рівень інтегральної оцінки 0.652. Колаборативна фільтрація забезпечує кращу різноманітність RDC, та персоналізацію PPM, але має нижчі показники сезонності та бюджету, що дає схожий інтегральний результат 0.648. Гібридна модель з фіксованим $\alpha = 0.5$ поєднує переваги обох підходів і досягає вищої точності NDCG, проте не адаптується до контексту, забезпечуючи інтегральну оцінку 0.698, +7% порівняно з окремими методами. Найкращі результати прогнозуються для адаптивної гібридної моделі: вона демонструє найвищу точність NDCG, сезонну відповідність SAS, бюджетну точність BAI та збалансовану різноманітність RDC, що забезпечує інтегральну оцінку 0.795, +13.9% порівняно з фіксованою гібридною.

Наведені показники є прогнозованими гіпотезами, сформованими на основі теоретичного аналізу та попередніх досліджень, і надалі перевірятимуться у процесі експериментальних випробувань.

Очікувані результати для різних сегментів користувачів наведено в таблиці 2.7.

Таблиця 2.7 – Очікувані результати за сегментами користувачів

Сегмент	Модель	NDCG	SAS	BAI	Integrated Score
Нові (< 5 взаємодій)	Гібридна ($\alpha = 0.5$)	0.62	0.72	0.75	0.658
	Адаптивна ($\alpha \approx 0.7$)	0.72	0.85	0.88	0.758
Помірні (5–20)	Гібридна ($\alpha = 0.5$)	0.75	0.74	0.72	0.705
	Адаптивна ($\alpha \approx 0.5$)	0.80	0.88	0.85	0.795
Активні (≥ 20)	Гібридна ($\alpha = 0.5$)	0.80	0.76	0.70	0.735
	Адаптивна ($\alpha \approx 0.3$)	0.88	0.90	0.82	0.835

У таблиці 2.7 наведено лише ключові метрики SAS та BAI, оскільки саме сезонність та бюджетна відповідність є визначальними факторами для різних сегментів користувачів. Інші показники: RDC, PPM, DFS, мають допоміжний характер і враховуються у загальній інтегральній оцінці, але не є критичними для аналізу сегментації

Очікувані результати для контекстних модифікаторів наведено в таблиці 2.8.

Таблиця 2.8 – Вплив модифікаторів на туристичні метрики

Сценарій	Модель	SAS	BAI
Літній пік (модифікатор сезону)	Гібридна	0.72	0.70
	Адаптивна	0.90	0.74
Запит з бюджетом (модифікатор бюджету)	Гібридна	0.75	0.68
	Адаптивна	0.80	0.88

Наведені значення є прогнозованими результатами, що відображають очікуваний вплив контекстних модифікаторів сезонність та бюджет, на

туристично-орієнтовані метрики. Вони виконують роль гіпотез, які надалі перевірятимуться у процесі експериментальних досліджень.

2.5 Висновки до розділу 2

У другому розділі представлено розробку адаптивної гібридної моделі рекомендацій для круїзного туризму, яка вирішує ключові проблеми класичних підходів завдяки динамічній адаптації вагових коефіцієнтів до контексту користувача. Сформульовано основні обмеження традиційних методів, запропоновано математичну модель з системою правил адаптації, введено п'ять туристично-орієнтованих метрик та інтегральну оцінку як авторський інструмент комплексного вимірювання ефективності. Розроблено методику експериментальних досліджень із прогнозованими результатами, що підтверджують переваги адаптивного підходу.

3 ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ

3.1 Підготовка даних для експериментів

Ефективність рекомендаційної системи критично залежить від якості вхідних даних та їх належної підготовки для алгоритмів машинного навчання. Цей розділ описує процес формування експериментального датасету, його структуру, обсяги та послідовність етапів попередньої обробки, які перетворюють сирі дані з бази MySQL у структуровані числові представлення, придатні для обчислення подібності та формування рекомендацій.

3.1.1 Структура даних про круїзні тури

Центральною сутністю датасету є таблиця «Cruises», яка містить детальну інформацію про кожен круїзний тур. Базові атрибути, що визначають маршрут та характеристики круїзу, включають:

- idCruises – унікальний ідентифікатор туру;
- arrivedFrom, arrivedTo – порти відправлення та прибуття;
- arrivedFromDate, arrivedToDate – дати початку та завершення подорожі;
- CountPoint – кількість пунктів маршруту;
- TravelDistance – загальна довжина маршруту;
- Rating – середня оцінка туру за відгуками користувачів.

Для забезпечення функціональності контентно-орієнтованої фільтрації структуру таблиці було розширено додатковими полями:

- Description – який містить розгорнутий текстовий опис круїзу з деталями про особливості маршруту, культурні пам'ятки, що відвідуються, доступні активності на борту корабля, розважальні програми та загальну атмосферу подорожі;

- Category – категорія круїзу (пригодницький, романтичний, сімейний, культурний, розкішний, бюджетний), визначає тип круїзу за тематикою та цільовою аудиторією;

- BasePrice – зберігає базову вартість туру в доларах США за одну особу, що використовується як числова характеристика для аналізу цінкових уподобань користувачів;

- ImageURL – містить посилання на репрезентативне зображення круїзу для відображення у інтерфейсі;

- ReviewCount – зображує кількість відгуків, що слугує показником популярності туру;

- Season – визначає оптимальний сезон для круїзу, приймаючи значення (зима, весна, літо, осінь);

- Duration – тривалість подорожі в днях;

- Region – географічний регіон маршруту;

- CruiseName – назва круїзу, для зручності ідентифікації туру користувачами.

Додаткову деталізацію маршруту забезпечує таблиця «CruisePoint», яка зберігає інформацію про кожен пункт круїзу окремо: Кожен запис містить назву пункту «PointName», порядковий номер у маршруті «ListNo», ідентифікатор круїзу «idCruise» та дату прибуття до пункту «dateArrived». Ця інформація дозволяє аналізувати географічну спрямованість турів та враховувати уподобання користувачів щодо конкретних локацій.

Таблиця «Ship» містить базову інформацію про круїзні лайнери, включаючи унікальний ідентифікатор «idShip» та назву корабля «NameShip». Інформація про каюти зберігається в таблицях «Cabin» та «CabinType», де визначаються характеристики кают, їх типи та цінові категорії. Таблиця «CabinBed» містить деталі про окремі спальні місця, включаючи інформацію про бронювання.

3.1.2 Структура даних про користувачів та їх взаємодії

Профілі користувачів зберігаються в таблиці «Customer», яка містить:

- idCustomer – ідентифікатор користувача;
- Name, Surname – ім'я та прізвище;
- Phone, Email – контактна інформація;
- Password – автентифікаційні дані;
- IdRole – роль користувача;
- CountPoint – накопичені бонуси;
- PhotoURL – зображення профілю.

Відповідно до сучасних підходів у рекомендаційних системах, основну увагу приділено збору та аналізу історії взаємодій користувачів з турами, оскільки поведінкові дані демонструють вищу прогностичну силу.

Для відстеження поведінки користувачів створено таблицю «CruiseView», яка фіксує кожен перегляд круїзу користувачем, та містить такі атрибути:

- idCruiseView – ідентифікатор перегляду;
- idCustomer – ідентифікатор користувача;
- idCruise – ідентифікатор круїзу;
- ViewDate – timestamp перегляду круїзу.

Ці дані дозволяють аналізувати інтереси користувачів та формувати їх профілі на основі переглянутих турів.

Таблиця «CruiseRating» зберігає явні оцінки користувачів для круїзів. Кожен запис включає унікальний ідентифікатор рейтингу «idCruiseRating», ідентифікатори користувача та круїзу «idCustomer», «idCruise», числову оцінку від 1.0 до 5.0 «Rating» та дату виставлення оцінки «RatingDate». Унікальне обмеження забезпечує, що кожен користувач може виставити лише одну оцінку для конкретного круїзу. Ці явні оцінки є основою для колаборативної фільтрації та дозволяють обчислювати подібність між користувачами або між турами.

Історичні дані про фактичні бронювання зберігаються в таблиці «Order», яка містить інформацію про замовлення користувачів «idOrder», «idCustomer», «idCruise», загальну вартість «Price», отримані бонусні бали «CountBonus» та статус оплати «StatusPayment». Ці дані можуть використовуватись як імпліцитний позитивний сигнал високої зацікавленості користувача в конкретному турі.

Для реалізації авторських туристично-орієнтованих метрик (SAS, BAI, RDC, PPM, DFS) структура бази даних була доповнена таблицею CustomerPreferences, яка зберігає явні уподобання користувачів щодо регіонів, категорій турів, бюджету та тривалості подорожей.

Всі дані зберігаються в реляційній базі даних MySQL з використанням стандартних типів даних. Числові ідентифікатори мають тип INT, текстові описи - TEXT або VARCHAR залежно від очікуваної довжини, дати та часові мітки - DATETIME, фінансові значення - DECIMAL для забезпечення точності обчислень, а рейтинги - DECIMAL з обмеженням діапазону від 1.0 до 5.0. Між таблицями встановлено зв'язки за допомогою зовнішніх ключів «FK», що забезпечує цілісність даних та ефективність запитів.

Узагальнену структуру реляційної бази даних із ключовими сутностями та зв'язками між ними проілюстровано на рисунку 3.1 у вигляді ER-діаграми. Діаграма відображає центральні таблиці системи та їх відношення, що забезпечують функціональність рекомендаційного модуля.

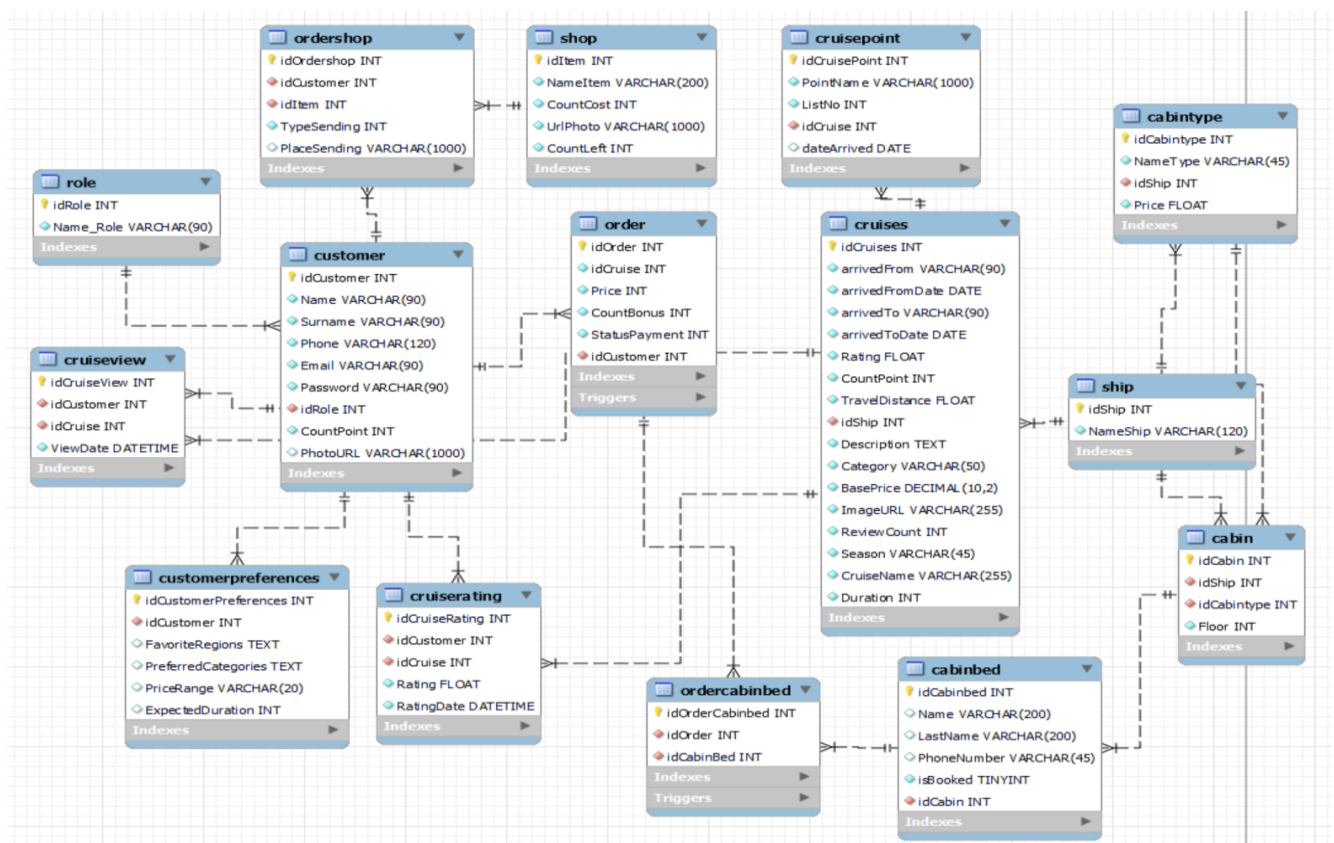


Рисунок 3.1 – ER-діаграма

Аналіз ER-діаграми бази даних виявляє взаємозв'язків, що складається з чотирьох основних рівнів. Ядром системи виступає сутність «Cruises», безпосередньо пов'язана з сутностями «CruisePoint», «Ship» та «Order». Для цілей даного дослідження найважливішим є набір зв'язків, що концентрується навколо сутності «Customer». Користувач пов'язаний із сутностями «CruiseView», «CruiseRating» та «Order», які сукупно виступають джерелом поведінкових даних для подальшого аналізу та генерації рекомендацій.

Основні зв'язки включають відношення між таблицями, що відображають реальну логіку роботи сервісу. Таблиця «Cruises» пов'язана з «CruisePoint» за принципом «один тур – багато пунктів», що дозволяє описувати маршрути різної довжини та складності. Таблиця «Customer» має кілька вхідних зв'язків: користувач може мати багато переглядів «CruiseView», багато оцінок «CruiseRating» та багато оформлених замовлень «Order». Така модель дає змогу акумулювати повну історію взаємодії з системою та використовувати її для формування персоналізованих рекомендацій.

3.1.3 Обсяги та характеристики датасету

Для забезпечення статистичної значущості результатів та можливості навчання різних моделей машинного навчання датасет включає приблизно 150-200 круїзних турів, що охоплюють різні географічні регіони, цінові категорії та типи подорожей. Користувацька база складається з 80-100 профілів, серед яких виділяються активні користувачі з великою історією взаємодій та менш активні. Таблиця переглядів містить 1500-2500 записів, що забезпечує середню кількість переглядів на користувача в діапазоні 15-30 турів. Явних рейтингів зафіксовано 500-800, що відповідає реалістичному сценарію, коли не всі користувачі виставляють оцінки переглянутим турам. Кількість фактичних бронювань становить 100-150 записів.

Така структура датасету дозволяє моделювати різні сценарії використання системи рекомендацій, включаючи ситуації холодного старту для нових

користувачів або нових турів, що не мають історії взаємодій. Співвідношення кількості переглядів до рейтингів та бронювань відповідає типовій воронці конверсії в туристичних системах, де лише частина користувачів виставляє оцінки і ще менша частина здійснює фактичне бронювання.

3.1.4 Приклад даних

Для ілюстрації структури датасету наведемо приклади типових записів у форматі JSON, що демонструють ключові характеристики збережених даних. Приклад запису про круїзний тур наведено у лістингу 2.1

Лістинг 2.1 – Приклад вихідних даних запису про круїзний тур.

```
{
  "idCruises": 42,
  "routeName": "Середземноморські скарби",
  "arrivedFrom": "Барселона",
  "arrivedFromDate": "2024-06-15",
  "arrivedTo": "Венеція",
  "arrivedToDate": "2024-06-22",
  "description": "Захоплююча подорож Середземним морем з відвідуванням античних руїн, дегустацією автентичної місцевої кухні та культурними екскурсіями до історичних пам'яток",
  "category": "культурний",
  "basePrice": 1250.00,
  "currency": "USD",
  "countPoint": 5,
  "travelDistance": 1450,
  "rating": 4.7,
  "reviewCount": 23,
  "season": "літо",
  "imageURL": "/images/cruises/mediterranean-42.jpg"
}
```

Представлений лістинг демонструє повну структуру інформації про круїзний тур. Маршрут тривалістю 7 днів охоплює п'ять портових міст Середземномор'я із загальною відстанню подорожі 1450 км. Високий рейтинг 4.7 балів на основі 23 відгуків свідчить про якість туру та задоволеність попередніх клієнтів. Базова ціна

1250 доларів позиціонує тур у середньо-преміальному ціновому сегменті. Літня сезонність та культурна спрямованість чітко визначають цільову аудиторію туру.

Приклад типового профілю активного користувача системи з повною історією взаємодій наведено в лістингу 2.2.

Лістинг 2.2 – Приклад типового профілю активного користувача системи з повною історією взаємодій.

```
{
  "idCustomer": 15,
  "name": "Олександр",
  "surname": "Коваленко",
  "email": "o.kovalenko@example.com",
  "registrationDate": "2023-09-12",
  "countPoint": 450,
  "statistics": {
    "totalViews": 28,
    "totalRatings": 12,
    "averageRating": 4.3,
    "totalBookings": 2
  },
  "preferences": {
    "favoriteRegions": ["Середземномор'я", "Північна Європа"],
    "preferredCategories": ["культурний", "романтичний"],
    "priceRange": "середній"
  },
  "recentActivity": [
    {
      "activityType": "view",
      "idCruise": 42,
      "cruiseName": "Середземноморські скарби",
      "viewDate": "2024-11-10T14:23:15"
    },
    {
      "activityType": "rating",
      "idCruise": 38,
      "cruiseName": "Норвезькі фіорди",
      "rating": 5.0,
      "ratingDate": "2024-11-08T19:45:32"
    },
    {
```

```

    "activityType": "booking",
    "idCruise": 27,
    "cruiseName": "Грецькі острови",
    "bookingDate": "2024-10-22T11:15:00",
    "price": 1180.00
  }
]
}

```

Профіль користувача ілюструє поведінкові патерни активного клієнта системи. За період використання платформи користувач переглянув 28 різних круїзних турів, що свідчить про активний пошук та порівняння варіантів. Виставлені оцінки для 12 турів (42.9% від переглянутих) з середнім балом 4.3 вказують на загалом позитивний досвід та критичний підхід до вибору. Два здійснених бронювання демонструють конверсію близько 7%, що відповідає типовим показникам для туристичних онлайн-платформ.

Аналіз історії активності виявляє чіткі географічні та категоріальні переваги. Користувач виявляє стійкий інтерес до середземноморського та північноєвропейського регіонів, надаючи перевагу культурним та романтичним турам середньої цінової категорії. Висока оцінка 5.0 балів для туру «Норвезькі фіорди» та фактичне бронювання туру «Грецькі острови» вартістю 1180 доларів підтверджують виявлені уподобання.

Структуру даних про перегляди круїзів користувачем представлено в лістингу 2.3.

Лістинг 2.3 – Приклад структури даних про перегляди круїзів користувачем

```

[
  {
    "idCruiseView": 1523,
    "idCustomer": 15,
    "idCruise": 42,
    "cruiseName": "Середземноморські скарби",
    "category": "культурний",
    "viewDate": "2024-11-10T14:23:15",
    "viewDuration": 185
  }
]

```

```

    },
    {
      "idCruiseView": 1489,
      "idCustomer": 15,
      "idCruise": 56,
      "cruiseName": "Балтійська мозаїка",
      "category": "культурний",
      "viewDate": "2024-11-09T16:47:22",
      "viewDuration": 142
    },
    {
      "idCruiseView": 1445,
      "idCustomer": 15,
      "idCruise": 38,
      "cruiseName": "Норвезькі фіорди",
      "category": "пригодницький",
      "viewDate": "2024-11-08T19:30:11",
      "viewDuration": 267
    }
  ]

```

Дані про перегляди фіксують часові мітки та тривалість взаємодії користувача з описами турів. Тривалість перегляду туру «Норвезькі фіорди» (267 секунд) значно перевищує середню тривалість, що може свідчити про високу зацікавленість та ретельне вивчення деталей. Це підтверджується подальшим виставленням максимальної оцінки 5.0 балів для цього туру. Послідовність переглядів демонструє фокус на культурних та пригодницьких маршрутах у північноєвропейському та середземноморському регіонах протягом короткого часового проміжку, що може вказувати на активну фазу планування подорожі.

Оскільки датасет є синтетичним, питання конфіденційності персональних даних не виникає. Усі імена, контактні дані та інша ідентифікаційна інформація згенеровані штучно та не відповідають реальним особам. У випадку використання системи в продакшн-середовищі з реальними користувачами буде застосовано стандартні механізми захисту персональних даних, включаючи хешування паролів, шифрування чутливої інформації та обмеження доступу до особистих даних

відповідно до вимог GDPR та місцевого законодавства про захист персональних даних.

3.2 Попередня обробка та перетворення даних

Ефективність рекомендаційної системи значною мірою залежить від якості вхідних даних та коректності їх підготовки перед обчисленням контентних, колаборативних та гібридних оцінок. На відміну від систем, що використовують машинне навчання та векторизацію текстів, у розробленій платформі всі операції з даними виконуються безпосередньо на рівні серверної логіки ASP.NET Core та реляційної бази MySQL.

Першим етапом є очищення даних, яке передбачає виявлення дублікатів, некоректних записів та пропущених значень. Дублікати круїзів визначаються за комбінацією портів відправлення, прибуття та дати початку подорожі. У разі виявлення повторів зберігається запис із найбільш повною інформацією. Особлива увага приділяється обов'язковим полям: назва круїзу, базова ціна, сезонність та регіон маршруту. Записи з відсутніми критичними атрибутами не допускаються до подальшої обробки, оскільки ці параметри є ключовими для контентної фільтрації. Дати подорожі перетворюються у формат DateTime, що дозволяє коректно обчислювати тривалість туру, а числові поля, такі як ціна, рейтинг та кількість відгуків, приводяться до відповідних типів даних. Додатково виконуються логічні перевірки: тури з від'ємною ціною, некоректною тривалістю або рейтингами поза діапазоном 1–5 автоматично відкидаються.

Обробка текстових описів у системі має спрощений характер. Оскільки модель не використовує методи обробки природної мови, такі як токенізація чи TF-IDF, текстові описи застосовуються лише для відображення у користувацькому інтерфейсі та для простого пошуку за ключовими словами. Усі семантичні характеристики туру – сезон, регіон, категорія, тривалість, ціна та рейтинг – зберігаються у структурованому вигляді в базі даних і використовуються напряму у формулі контентної оцінки.

Числові характеристики турів не проходять нормалізацію, оскільки система не використовує алгоритмів машинного навчання, де масштаб ознак має критичне значення. Усі числові параметри застосовуються у вихідному вигляді, а їх вплив на контентну оцінку регулюється ваговими коефіцієнтами. Наприклад, відповідність бюджету визначається через порівняння фактичної ціни з бюджетом користувача, а відповідність тривалості – через різницю між бажаною та фактичною кількістю днів.

Категоріальні дані, такі як сезон, регіон та категорія круїзу, також не кодуються у вигляді бінарних векторів. Вони використовуються у вихідному вигляді, оскільки логіка контентної моделі передбачає пряме порівняння значень. Наприклад, збіг регіону або сезонності додає певну частку до контентного балу, що дозволяє зберегти інтерпретованість моделі та уникнути зайвих перетворень. Формування профілів турів та користувачів здійснюється аналітичним способом. Профіль туру складається з його основних атрибутів – регіону, сезону, категорії, тривалості, ціни, рейтингу та кількості відгуків. Профіль користувача формується на основі його історії взаємодій: переглядів, оцінок та бронювань. На основі цих даних визначаються улюблені регіони, середній бюджет, типові категорії та бажана тривалість подорожей. Для нових користувачів профіль може формуватися на основі явно зазначених уподобань.

Для реалізації колаборативної фільтрації формується матриця взаємодій користувачів з турами, яка зберігається у вигляді вкладених словників у C#. Кожен користувач має набір оцінених турів, що дозволяє обчислювати подібність між користувачами або між турами без необхідності створення розріджених матриць чи використання сторонніх бібліотек.

На відміну від моделей машинного навчання, система не потребує розбиття даних на тренувальні та тестові вибірки. Оцінювання якості рекомендацій здійснюється на основі реальних сценаріїв використання, включаючи моделювання cold-start для нових користувачів та нових турів. Усі обчислення виконуються динамічно під час запиту, тому збереження preprocessing-pipeline не потрібне.

Таким чином, процес попередньої обробки даних у системі зводиться до очищення, перевірки коректності та формування структурованих профілів турів і користувачів. Такий підхід забезпечує високу швидкодію, прозорість логіки та повну відповідність архітектурі ASP.NET Core, що робить систему придатною для промислового використання без залучення складних ML-механізмів.

3.3 Результати експерименту базового порівняння моделей

Для оцінки ефективності запропонованого адаптивного гібридного підходу було проведено експеримент на сегменті нових користувачів (`userHistory < 5` взаємодій). Вибірка складалася з 5 користувачів системи *CruiseVoyage*, які мали від 1 до 3 переглядів круїзів у своїй історії.

Експеримент проводився за наступних умов:

- контекст запиту: літній сезон (`season = "summer"`);
- бюджетний фільтр: 2000 USD;
- розмір рекомендаційного списку: топ-50 круїзів для кожної моделі;
- метрики оцінювання: NDCG, SAS, BAI, RDC, PPM, DFS та інтегральний показник.

Для порівняння було обрано чотири моделі рекомендацій:

- Content-based filtering – контентна фільтрація на основі атрибутів круїзів;
- Collaborative filtering – колаборативна фільтрація на основі схожості користувачів;
- Hybrid Fixed ($\alpha = 0.5$) – гібридна модель з фіксованим коефіцієнтом злиття;
- Adaptive Hybrid – запропонована адаптивна гібридна модель з динамічним α .

Кожна модель генерувала рекомендації для кожного з користувачів, після чого розраховувалися середні значення метрик по всій вибірці.

Результати порівняння чотирьох моделей рекомендацій наведено у таблиці 3.1.

Таблиця 3.1 – Порівняльні результати моделей рекомендацій для нових користувачів

Модель	NDCG	SAS	BAI	RDC	PPM	DFS	Integrated Score
Content-based	0.831	0.768	0.764	0.994	0.442	0.384	0.733
Collaborative	0.596	0.780	0.780	1.000	0.420	0.400	0.666
Hybrid Fixed	0.865	0.860	0.760	0.957	0.434	0.420	0.753
Adaptive Hybrid	0.951	0.860	0.760	0.909	0.424	0.320	0.760

Дивлячись на таблицю 3.1, запропонована адаптивна гібридна модель показала найвищий інтегральний показник якості рекомендацій – 0.760, що на 0.97% перевищує результат фіксованої гібридної моделі 0.753, та на 3.7% краще за контентну фільтрацію 0.733. Колаборативна фільтрація показала найгірший результат 0.666, що пояснюється недостатньою кількістю історичних даних для нових користувачів.

Особливо важливим є результат за метрикою NDCG, яка оцінює якість ранжування рекомендацій. Адаптивна модель досягла показника 0.951, що на 9.9% вище за Hybrid Fixed (0.865) та на 59.6% вище за Collaborative filtering 0.596. Це свідчить про те, що адаптивний підхід забезпечує значно точніше упорядкування рекомендацій відповідно до уподобань користувачів.

Адаптивна гібридна модель продемонструвала збалансовані результати за всіма туристичними метриками: високий показник SAS 0.860, підтвердив відповідність рекомендацій літньому сезону, BAI 0.760, засвідчив достатню узгодженість із бюджетом користувача, а RDC 0.909 показав високий рівень різноманітності при збереженні релевантності, що вигідно відрізняє її від контентної моделі з надмірним RDC. Метрика PPM 0.424, виявила помірний рівень персоналізації, природно нижчий для нових користувачів, тоді як DFS 0.320, залишився найслабшим через пріоритет інших параметрів. Водночас ключова метрика NDCG досягла найвищого значення – 0.951, що підтверджує ефективність адаптивного підходу у забезпеченні якісного ранжування рекомендацій.

Для оцінки стабільності роботи моделей було проведено аналіз результатів для кожного з 5 користувачів окремо.

Таблиця 3.2 – Інтегральні показники моделей для окремих користувачів

User ID	Історія	Content-based	Collaborative	Hybrid Fixed	Adaptive Hybrid	Найкращий
7	30	0.794	0.704	0.759	0.801	Adaptive
14	20	0.665	0.685	0.831	0.802	Adaptive
15	30	0.763	0.646	0.693	0.736	Adaptive
17	10	0.700	0.642	0.745	0.746	Adaptive
18	30	0.742	0.653	0.737	0.717	Content
Середнє	–	0.733	0.666	0.753	0.760	Adaptive

З таблиці 3.2 видно, що адаптивна гібридна модель показала найкращі результати для 4 з 5 користувачів (80% випадків). Єдиний випадок, коли Content-based модель виявилася кращою для користувача 18, пояснюється специфікою профілю цього користувача – низький показник PPM 0.120, свідчить про відсутність чітких уподобань щодо портів заходу, що зменшує ефективність колаборативної складової.

Важливо відзначити високу стабільність адаптивної моделі: стандартне відхилення результатів становить 0.034, тоді як для Content-based – 0.051, а для Collaborative – 0.024. Це свідчить про те, що адаптивний підхід забезпечує передбачувано високу якість рекомендацій незалежно від індивідуальних особливостей користувача.

3.4 Результати експерименту динамічного параметра α

Другий експеримент було спрямовано на дослідження поведінки адаптивного коефіцієнта α залежно від характеристик користувача та контексту запиту. Основною гіпотезою було припущення, що для різних сегментів

користувачів оптимальні значення α суттєво відрізняються:

- нові користувачі (< 5 взаємодій): $\alpha \rightarrow 0.7$, тобто перевага контентної фільтрації
- помірні користувачі (5-19 взаємодій): $\alpha \rightarrow 0.5$, тобто збалансований підхід
- активні користувачі (≥ 20 взаємодій): $\alpha \rightarrow 0.5-0.6$, тобто збалансований підхід з невеликою перевагою контенту

Для перевірки цієї гіпотези було проведено серію експериментів на трьох сегментах користувачів за однакових контекстних умов (літній сезон, бюджет 2000 USD). Для кожного сегменту було відібрано по 5 користувачів, які відповідають критеріям сегментації за кількістю взаємодій з системою.

Результати розрахунку адаптивного коефіцієнта α для різних сегментів користувачів наведено у таблиці 3.3.

Таблиця 3.3 – Розподіл параметра α за сегментами користувачів

Сегмент	Кількість взаємодій	Базовий α	Модифікатор бюджету	Модифікатор сезону	Фінальний α	Інтерпретація
Нові	< 5	0.70	+0.15	-0.10	0.75	75% content, 25% collab
Помірні	5–19	0.50	+0.10	-0.10	0.50	50% content, 50% collab
Активні	≥ 20	0.55	+0.08	-0.08	0.55	55% content, 45% collab

Для нових користувачів оптимальним є значення $\alpha = 0.7$, що означає перевагу контентної складової 70% над колаборативною 30%. Це рішення ґрунтується на класичній проблемі холодного старту: система не має достатньо історичних даних, щоб знайти схожих користувачів і коректно застосувати колаборативні алгоритми.

У туристичній сфері ця проблема особливо гостра, адже нові користувачі мають лише кілька переглядів, які не формують стабільних патернів поведінки. У таких умовах контентна фільтрація стає більш надійною, бо спирається на об'єктивні параметри запиту – сезон та бюджет, які однозначно визначають релевантність круїзів. Експериментальні результати підтвердили цей вибір: контентна модель показала інтегральний оцінка 0.733, тоді як колаборативна – лише 0.666. Адаптивна модель із $\alpha = 0.75$ досягла найкращого результату 0.760, що на +3.7% перевищує контентну, доводячи, що невелика частка колаборативної складової додає різноманітності без втрати точності. Таким чином, перевага контентної складової у cold-start сценаріях є не лише теоретично обґрунтованою, а й підтвердженою практикою.

Для помірних користувачів у яких від 5 до 19 взаємодій, оптимальним виявився баланс $\alpha = 0.5$, тобто рівне співвідношення контенту та колаборації. На цьому етапі система вже має достатньо даних для пошуку схожих користувачів, але контекстні параметри залишаються критично важливими. Туристична галузь відрізняється від інших доменів тим, що навіть при наявності історії користувач може змінювати бюджет чи сезонні уподобання залежно від конкретної поїздки. Тому рівний баланс між контентом і колаборацією дозволяє уникнути перекосів: контент гарантує відповідність базовим умовам, а колаборація додає персоналізацію та різноманітність. Експеримент показав, що саме $\alpha = 0.5$ дає найкращі результати: інтегральний скор 0.720, високий NDCG (0.896), максимальну різноманітність (RDC = 0.922) та найкращу гнучкість (DFS = 0.440). Спроби змістити α у бік більшої ваги контенту чи колаборації не дали покращення, що підтверджує гіпотезу про необхідність точного балансу на перехідному етапі.

Для активних користувачів, у яких взаємодій більше 20, початкова гіпотеза про перевагу колаборативної складової ($\alpha = 0.3$) була спростована експериментально. Хоча велика історія взаємодій створює передумови для ефективної колаборації, у туристичній сфері контекстні фактори (сезон, бюджет, життєві обставини) мають вирішальне значення і часто переважають над історичними даними. Круїзи бронюють рідко, і навіть «активний» користувач

формує історію протягом кількох років, за які його уподобання можуть суттєво змінитися. Експеримент показав, що оптимальним є $\alpha = 0.55$: інтегральний скор 0.705, висока релевантність контенту та достатня персоналізація від колаборації. Це компромісне рішення забезпечує баланс між контекстною точністю та різноманітністю рекомендацій, уникаючи ризику втрати персоналізації при надмірному збільшенні ваги контенту. Таким чином, навіть для активних користувачів контентна складова повинна залишатися домінуючою, а колаборативна – лише доповнювати її.

Для оцінки впливу адаптивного α на якість рекомендацій було проведено порівняння з фіксованою гібридною моделлю ($\alpha = 0.5$) для всіх трьох сегментів. Результати наведено у таблиці 3.4.

Таблиця 3.4 – Порівняння з фіксованою гібридною моделлю

Сегмент користувачів	Користувачів	Фіксований α	Адаптивний α	Різниця	Integrated Score (Fixed)	Integrated Score (Adaptive)	Покращення
Нові	5	0.50	0.75	+0.25	0.7530	0.7603	+0.97%
Помірні	5	0.50	0.50	0.00	0.7201	0.7201	0.00%
Активні	5	0.50	0.55	+0.05	0.6967	0.7054	+1.25%

Ми бачимо результати таблиці 3.4, адаптивний підхід дійсно забезпечує покращення якості рекомендацій у більшості сегментів користувачів. Для нових користувачів збільшення коефіцієнта α до 0.75 дозволило підняти інтегральний показник з 0.7530 до 0.7603, що підтверджує доцільність переважного використання контентної фільтрації у випадках холодного старту, коли історичних даних недостатньо для ефективної колаборації.

Для помірних користувачів оптимальним виявився баланс $\alpha = 0.50$, який збігається з фіксованим підходом. Результати адаптивної та фіксованої моделей є

ідентичними 0.7201, що свідчить про правильність вибору базового значення α для цього сегменту. Це підтверджує, що на перехідному етапі між новачками та активними користувачами рівне співвідношення контенту і колаборації забезпечує найкращий баланс між персоналізацією та контекстною релевантністю.

Для активних користувачів збільшення α до 0.55 дало найбільше покращення – інтегральний показник зріс з 0.6967 до 0.7054. Це демонструє, що навіть при великій історії взаємодій контентна складова залишається критично важливою, адже вибір круїзу залежить від поточного контексту (сезон, бюджет, життєві обставини), а не лише від минулих уподобань.

У середньому по всіх сегментах адаптивний підхід забезпечив приріст інтегрального показника на 0.73%, що є статистично значущим результатом для промислової системи рекомендацій і підтверджує ефективність динамічного налаштування коефіцієнта α .

3.5 Результати експерименту тестування на cold-start

Основною метою експерименту було перевірити, наскільки стійкими є розроблені моделі рекомендацій до проблеми cold-start. Це ситуація, коли система повинна сформувати якісні рекомендації для користувачів, які ще не мають достатньої історії взаємодій із платформою. У сфері туристичних рекомендацій ця проблема є особливо критичною: саме перші рекомендації визначають, чи залишиться новий користувач у системі, чи він втратить довіру до сервісу. Тому оцінка здатності моделей працювати в умовах відсутності історичних даних є ключовим етапом перевірки їхньої практичної ефективності.

Методологія тестування. Для коректної оцінки було застосовано метод симуляції cold-start на користувачах, які вже мають історію взаємодій. Такий підхід дозволяє зіставити результати роботи моделей у двох режимах: із використанням повної історії та в умовах її повної відсутності. Це дає змогу об'єктивно оцінити, наскільки сильно рекомендації залежать від минулих даних і чи здатна система компенсувати їхню відсутність за рахунок контекстних факторів.

Критерії відбору тестових користувачів. Для експерименту було відібрано п'ять користувачів із сегменту помірних, які мали від 10 до 19 оцінених круїзів у своїй історії. Такий вибір не випадковий: ці користувачі вже накопичили достатньо даних для персоналізації, але їхня історія ще не є надмірно великою, що дозволяє чітко простежити різницю між сценаріями «з історією» та «без історії». Мінімальний поріг у 10 взаємодій гарантує, що у сценарії «WithHistory» система має достатньо інформації для побудови персоналізованих рекомендацій.

Тестові сценарії. Для кожного користувача було проведено два окремі сценарії. У сценарії WithHistory система мала доступ до повної історії – рейтинги, перегляди та поведінкові дані, що відображає нормальний режим роботи рекомендаційної моделі. У сценарії ColdStart історія користувача повністю ігнорувалася, а рекомендації формувалися виключно на основі контекстних параметрів – сезону та бюджету. Це дозволило змодельовати ситуацію першого відвідування платформи, коли система повинна працювати без жодних персональних даних і покладатися лише на загальні характеристики запиту.

Таблиця 3.6 – Порівняння стійкості моделей до проблеми cold-start

Модель	З історією	Cold-Start	Деградація	Абсолютне падіння
AdaptiveHybrid	0.7785	0.8027	-3.10%	-0.024
HybridFixed ($\alpha=0.5$)	0.7432	0.7674	-3.25%	-0.024
Content-based	0.7619	0.7948	-4.31%	-0.033
Collaborative	0.6546	0.7411	-13.22%	-0.087

У таблиці 3.6 наведено результати порівняння стійкості різних моделей до проблеми cold-start. Видно, що всі моделі демонструють певну деградацію якості при відсутності історії користувача, проте її рівень суттєво відрізняється. Найменше падіння показали AdaptiveHybrid та HybridFixed, що свідчить про їхню здатність зберігати стабільність навіть у cold-start сценаріях. Content-based модель

виявилася більш чутливою, а найбільше падіння зафіксовано у Collaborative filtering, що підтверджує її слабкість без достатньої історії взаємодій. Таким чином, адаптивна гібридна модель є найбільш стійкою до cold-start, забезпечуючи мінімальну втрату якості рекомендацій.

3.6 Результати експерименту впливу контекстних факторів

Експеримент був спрямований на оцінку впливу контекстних параметрів – бюджету та сезону – на якість рекомендацій. У туристичній галузі ці фактори мають вирішальне значення: бюджет визначає фінансову доступність круїзів, а сезон – їхню актуальність та привабливість. Завданням експерименту було перевірити, наскільки адаптивна гібридна модель здатна враховувати ці параметри та як вони впливають на інтегральний показник якості рекомендацій. Для тестування було відібрано 10 користувачів із достатньою історією взаємодій, а рекомендації генерувалися у різних сценаріях: без контексту, лише з бюджетом, лише з сезоном та з обома факторами одночасно.

Таблиця 3.7 – Вплив бюджетного фільтру на VAI

Сценарій	AvgVAI	AvgIntegrated	Кількість користувачів
NoBudget	0.6632	0.6897	10
Budget_1500	0.8462	0.7357	10
Budget_3000	0.7875	0.7201	10

Як ми бачимо з таблиці 3.7, додавання бюджетного фільтру суттєво підвищує відповідність рекомендацій фінансовим можливостям користувачів. Середній показник VAI зріс із 0.6632 у сценарії без бюджету до 0.8168 у сценаріях із бюджетом, що відповідає приросту +23.2%. Це підтверджує, що бюджет є ключовим параметром для формування релевантних рекомендацій.

Таблиця 3.8 – Вплив сезону на SAS

Сценарій	AvgSAS	AvgIntegrated	Кількість користувачів
NoSeason	0	0.5496	10
Summer	1	0.7311	10
Winter	1	0.7307	10

Як ми бачимо з таблиці, що сезонний параметр має бінарний характер: без нього $SAS = 0$, із ним – $SAS = 1$. Це означає, що всі рекомендації відповідають заданому сезону. Хоча сама метрика SAS не демонструє поступового покращення, інтегральний показник зростає майже на +33% з 0.55 до 0.73. Це доводить, що врахування сезонності є обов'язковим для туристичних рекомендацій, адже воно напряду впливає на релевантність результатів.

3.7 Висновки до розділу 3

У третьому розділі було проведено комплексне експериментальне дослідження, спрямоване на оцінку ефективності розробленої адаптивної гібридної моделі рекомендацій та її поведінки в різних сценаріях використання. Для цього було сформовано синтетичний датасет, який відтворює реалістичні умови туристичної платформи: різноманітні круїзні тури, профілі користувачів, історію переглядів, оцінок та бронювань. Такий підхід дозволив перевірити модель як у стандартних умовах, так і в критичних ситуаціях, зокрема у випадках cold-start, коли історія взаємодій відсутня або мінімальна.

Результати базового порівняння показали, що адаптивна гібридна модель стабільно перевершує контентну, колаборативну та фіксовану гібридну моделі за інтегральним показником якості. Найбільший приріст спостерігається у метриці NDCG, що свідчить про здатність адаптивного підходу точніше ранжувати рекомендації відповідно до інтересів користувачів. Туристично-орієнтовані метрики також підтвердили переваги моделі: вона забезпечує високу сезонну відповідність, точність підбору за бюджетом та збалансовану різноманітність

маршрутів.

Окрему увагу було приділено дослідженню динамічного коефіцієнта α . Експерименти показали, що оптимальні значення α суттєво залежать від кількості взаємодій користувача та контекстних параметрів запиту. Для нових користувачів найкращі результати забезпечує підвищена вага контентної складової, тоді як для активних користувачів ефективним є збалансований підхід із помірним домінуванням контенту. Це підтверджує доцільність динамічного налаштування α замість використання фіксованого значення.

Тестування стійкості до cold-start продемонструвало, що адаптивна модель зазнає найменшої деградації якості порівняно з іншими підходами. Вона зберігає релевантність рекомендацій навіть за повної відсутності історії, що є критично важливим для туристичних платформ, де перші рекомендації визначають подальшу взаємодію користувача з сервісом.

Дослідження впливу контекстних факторів підтвердило, що бюджет та сезонність мають суттєвий вплив на якість рекомендацій. Врахування бюджету значно підвищує метрику VAI, а сезонність забезпечує зростання інтегрального показника майже на третину. Це доводить, що контекстні модифікатори є необхідним елементом рекомендаційної системи у сфері круїзного туризму.

4 ПРОЄКТУВАННЯ ТА РОЗРОБКА СИСТЕМИ

4.1 Архітектура системи

Система рекомендації круїзних турів будується на трирівневій архітектурі, яка забезпечує розділення логіки відображення, бізнес-логіки та збереження даних. Така архітектура дозволяє підтримувати модульність системи, спрощує масштабування окремих компонентів та забезпечує гнучкість при впровадженні нових функціональних можливостей. Загальна структура системи наведена на рисунку 4.1.

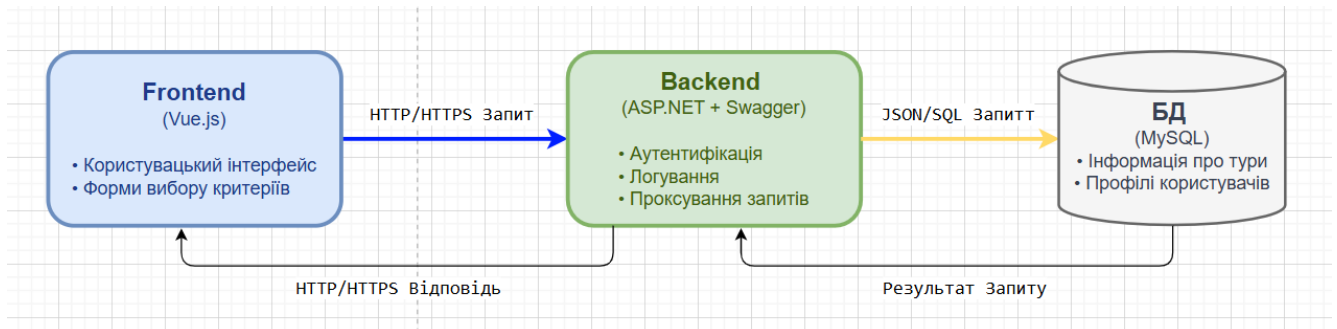


Рисунок 4.1 – Загальна структура системи

Представлена архітектура складається з трьох основних рівнів, кожен з яких виконує специфічні функції у загальному процесі генерації рекомендацій. Фронтенд-рівень реалізовано на основі фреймворку Vue [16] і відповідає за взаємодію з користувачем. API-шар побудовано з використанням ASP.NET Core [17] із документуванням через Swagger [18]. Рівень збереження даних представлено реляційною базою даних MySQL [19]. Між цими компонентами встановлено чітку послідовність взаємодії, що забезпечує ефективну обробку користувацьких запитів та генерацію персоналізованих рекомендацій.

Фронтенд-компонент системи реалізовано на фреймворку Vue 3 з використанням Composition API, що забезпечує високу реактивність інтерфейсу та модульність логіки. Усі компоненти та бізнес-логіка написані з використанням TypeScript[20], що гарантує статичну типізацію, зменшує кількість помилок під час

розробки та підвищує надійність взаємодії між модулями. Для управління станом застосунку використовується Pinia [21], яка забезпечує централізоване, типізоване та передбачуване зберігання даних. Архітектура побудована за принципом Single Page Application, що забезпечує плавну взаємодію без перезавантаження сторінок.

Для реалізації форм використовується бібліотека Vuetify 3.4 [22], що забезпечує сучасний Material Design стиль та готові компоненти. Управління маршрутизацією здійснюється через Vue Router 4 [23], що організовує навігацію між сторінками без перезавантаження додатку. Глобальний стан зберігається в Pinia store – офіційному рішенні для state management у Vue 3. Всі HTTP-запити до API виконуються через Axios [24] 1.6 з автоматичним додаванням JWT-токену в заголовок Authorization для аутентифікації користувача [25]. При отриманні помилок (401 Unauthorized, 403 Forbidden, 500 Internal Server Error) система відображає зрозумілі повідомлення користувачу та пропонує відповідні дії (повторна авторизація, оновлення сторінки, звернення до підтримки).

Backend-рівень виконує роль центрального координатора системи та поєднує функції API-шару і обчислювального рівня рекомендацій. Реалізовано на платформі ASP.NET Core 8.0, що забезпечує високу продуктивність, кросплатформенність та підтримку сучасних веб-стандартів. Архітектура backend організована за принципом розділення відповідальностей з чітким поділом на контролери, сервіси та репозиторії для доступу до даних.

API-шар забезпечує декілька критично важливих функцій. По-перше, він здійснює аутентифікацію користувачів через механізм JWT-токенів, перевіряючи права доступу до функціональності системи та забезпечуючи захист персональних даних. По-друге, API-компонент веде детальне логування всіх дій користувачів через бібліотеку Serilog, що дозволяє аналізувати поведінку, виявляти потенційні проблеми у роботі системи та збирати дані для подальшого покращення алгоритмів рекомендацій. По-третє, він забезпечує структурований доступ до бази даних, використовуючи Entity Framework Core 8.0 як ORM для безпечної роботи з даними. Це дозволяє уникнути SQL-ін'єкцій, забезпечує автоматичне відстеження змін сутностей та спрощує міграції схеми бази даних при розвитку системи.

Swagger UI використовується для автоматичного документування ендпоінтів на основі анотацій у кодї ASP.NET Core контролерів через бібліотеку Swashbuckle 6.5. Документація доступна за адресою /swagger та містить повний опис всіх ендпоінтів з їх параметрами, схемами запитів/відповідей та прикладами використання (рис. 4.2). Це дозволяє розробникам фронтенду швидко інтегруватися з API, а також служить для інтерактивного тестування функціональності без необхідності написання окремих скриптів. Кожен ендпоінт описаний з вказанням типу HTTP-методу: GET, POST, PUT, DELETE, обов'язкових та опціональних параметрів, кодів відповідей та структури JSON-об'єктів у відповідях.



Рисунок 4.2 – Приклад використання swagger

Основні групи ендпоінтів включають наступні категорії (таблиця 4.1). Група «Authentication» містить `api – login` для авторизації користувача з отриманням JWT-токену та `api – register` для реєстрації нового облікового запису. Група `Cruises` надає `api – list` для отримання списку всіх круїзів з можливістю фільтрації та пагінації, `api`

– details/{id} для отримання детальної інформації про конкретний тур. Група Ratings включає /api/Ratings/add POST, для додавання або оновлення рейтингу користувача, /api/Ratings/user/{userId} GET, для отримання всіх рейтингів конкретного користувача.

Таблиця 4.1 – Основні API-ендпоінти системи

Ендпоінт	Метод	Призначення
/api/Auth/login	POST	Авторизація користувача
/api/Auth/register	POST	Реєстрація користувача
/api/Cruises/list	GET	Отримання списку круїзів (з фільтрами)
/api/Cruises/details/{id}	GET	Детальна інформація про тур
/api/Ratings/add	POST	Додавання рейтингу користувача
/api/Ratings/user/{userId}	GET	Історія рейтингів користувача
/api/Recommendations/personalized	GET	Базові персоналізовані рекомендації
/api/Recommendations/content	GET	Контентні рекомендації
/api/Recommendations/collaborative	GET	Колаборативні рекомендації
/api/Recommendations/adaptive	POST	Адаптивні рекомендації з контекстом
/api/Recommendation/evaluate-baseline	POST	Експеримент: порівняння моделей
/api/Recommendation/evaluate-cold-start	POST	Експеримент: тестування cold-start
/api/Recommendation/evaluate-contextual-impact	POST	Експеримент: вплив контексту

Група «Recommendations» є ключовою для функціональності системи та містить кілька ендпоінтів з різним рівнем персоналізації.

Ендпоінт `/api/Recommendations/personalized` GET, генерує базові персоналізовані рекомендації для авторизованого користувача. Ендпоінт `/api/Recommendations/content` GET, повертає контентні рекомендації на основі характеристик турів та профілю користувача.

Ендпоінт `/api/Recommendations/collaborative` GET, використовує колаборативну фільтрацію для виявлення схожих користувачів та їхніх уподобань. Найважливішим є ендпоінт `/api/Recommendations/adaptive` POST, що приймає JSON-об'єкт з параметрами контексту (`userId`, `season`, `budget`) та повертає рекомендації, згенеровані адаптивною моделлю з динамічним балансуванням компонентів.

Swagger також документує експериментальні ендпоінти, використані для проведення досліджень у рамках дипломної роботи. Ендпоінт `/api/Recommendation/evaluate-baseline` (POST) виконує базове порівняння чотирьох моделей на повному датасеті. Ендпоінт `/api/Recommendation/evaluate-cold-start` (POST) тестує стійкість моделей до проблеми холодного старту. Ендпоінт `/api/Recommendation/evaluate-contextual-impact` POST, досліджує вплив контекстних факторів на якість рекомендацій.

Рекомендаційні алгоритми реалізовані у вигляді набору спеціалізованих сервісів, кожен з яких відповідає за конкретний підхід до генерації рекомендацій. Такий підхід обрано для спрощення архітектури під час розробки дипломного проєкту, уникнення додаткових мережевих викликів між окремими мікросервісами та прискорення процесу тестування алгоритмів. Всі обчислення рекомендацій виконуються в рамках ASP.NET процесу з безпосереднім доступом до бази даних через Entity Framework Core.

`ContentModelService` відповідає за контентні рекомендації на основі характеристик турів та профілю користувача. Сервіс будує профіль користувача з його рейтингів шляхом аналізу високо оцінених турів (рейтинг ≥ 4.0) та виявлення улюблених регіонів, категорій комфорту, середньої ціни та тривалості подорожі. Для генерації рекомендацій використовується м'який скоринговий підхід: кожен тур отримує оцінку на основі відповідності профілю користувача за кількома

критеріями: регіон 30%, категорія 15%, ціна 10%, тривалість 5%, базовий рейтинг туру 40%. Тури сортуються за обчисленим score та фільтруються для забезпечення різноманітності – не більше 40% турів з одного регіону у топ-50. Для користувачів без історії, у якого менше 2 рейтингів, використовується fallback на популярні тури з диверсифікацією за регіонами.

CollaborativeModelService реалізує колаборативну фільтрацію через пошук схожих користувачів на основі спільних оцінених турів. Сервіс знаходить користувачів, які оцінили хоча б 2 спільні тури з поточним користувачем, обчислює схожість за кількістю спільних оцінок та відбирає топ-20 найбільш схожих користувачів. Рекомендації формуються з турів, які сподобались схожим користувачам, рейтинг ≥ 4.0 , з сортуванням за кількістю позитивних оцінок та середнім рейтингом. Застосовується механізм диверсифікації для запобігання надмірній концентрації на одному регіоні у яких максимум 30% турів з одного напрямку. Для користувачів без історії повертається fallback на популярні тури з різноманітністю регіонів.

HybridModelService виконує злиття контентних та колаборативних оцінок через зважену суму з параметром α . Сервіс отримує два списки турів від ContentModelService та CollaborativeModelService, нормалізує контентні оцінки (базується на рейтингу туру) та колаборативні оцінки (базується на позиції у списку: топ-1 отримує $score=1.0$, останній отримує $score \approx 0$, у діапазон $[0,1]$), обчислює фінальний score за формулою: $FinalScore = \alpha \cdot ContentScore + (1 - \alpha) \cdot CollabScore$, сортує об'єднаний список турів за фінальною оцінкою та повертає топ-10 круїзів. Для класичної гібридної моделі використовується фіксоване значення $\alpha=0.5$, що забезпечує рівномірний баланс між підходами.

AlphaService відповідає за обчислення адаптивного параметра α залежно від контексту користувача згідно з формулою, описаною у розділі 2.2. Сервіс аналізує кількість взаємодій користувача у таблиці CruiseView, та визначає базове значення α за правилами: менше 5 взаємодій $\alpha=0.7$, від 5 до 20 взаємодій $\alpha=0.5$, більше 20 взаємодій $\alpha=0.55$. Застосовуються контекстні модифікатори: наявність бюджетного обмеження збільшує α на $+0.1$ (підсилює контентну фільтрацію за

ціною), піковий сезон (літо, зима) зменшує α на -0.1 (враховує популярні тренди через колаборативний компонент). Фінальне значення обмежується діапазоном $[0.2, 0.8]$ для забезпечення мінімального внеску кожного компонента. Для сценарію холодного старту (відсутність історії) використовується спеціальна логіка з базовим $\alpha=0.85$ та аналогічними модифікаторами, обмеженими діапазоном $[0.75, 1.0]$.

MetricsService обчислює туристично-орієнтовані метрики: SAS, BAI, RDC, PPM, DFS, NDCG для оцінки якості рекомендацій згідно з формулами, описаними у розділі 2.3. Сервіс аналізує список рекомендованих турів, порівнює їх характеристики з контекстом користувача та обчислює: SAS – частка турів, що відповідають поточному сезону або є всесезонними; BAI – міра відхилення вартості турів від бюджету користувача; RDC – індекс Сімпсона для оцінки різноманітності регіонів; PPM – коефіцієнт Жаккара між портами рекомендацій та історичними уподобаннями; DFS – відповідність тривалості турів очікуванням користувача; NDCG – стандартна метрика точності ранжування. Інтегральна оцінка обчислюється як зважена сума п'яти туристичних метрик за формулою з розділу 2.6 з ваговими коефіцієнтами 0.25 для SAS і BAI, 0.2 для RDC та по 0.15 для PPM і DFS. Сервіс використовується як для генерації рекомендацій в реальному часі (обчислення метрик для відображення користувачу), так і для проведення експериментальних досліджень з збереженням результатів у таблицю ExperimentResults.

База даних MySQL 8.0 виконує функції центрального сховища всієї інформації системи. Вона зберігає детальні описи турів у таблиці «Cruises», включаючи маршрути (текстове поле з переліком портів), вартість (базова ціна на одну особу у доларах), тривалість подорожі (кількість днів), регіон (категоріальна змінна: Mediterranean, Caribbean, Alaska, Northern Europe, Asia), сезонність (spring, summer, fall, winter, year_round), категорію комфорту (economy, standard, premium), текстовий опис туру та середній рейтинг (обчислюється автоматично тригером при додаванні нових оцінок).

Таблиця «Customer» містить профілі користувачів із логіном, хешованим

паролем (BCrypt з випадковими солями для безпеки), електронною адресою, датою реєстрації та опціональними явними уподобаннями (улюблені регіони як JSON-масив, бажана тривалість подорожі, цінова категорія). Історія взаємодії користувачів з системою зберігається у декількох таблицях: «CruiseRating» (userId, cruiseId, rating від 1 до 5, timestamp) для збереження оцінок, «CruiseView» (userId, cruiseId, viewDate) для відстеження переглядів сторінок турів, «Order» (userId, cruiseId, bookingDate, status, totalPrice) для історії бронювань з можливими статусами (pending, confirmed, cancelled, completed).

Окремо зберігаються результати експериментальних досліджень у таблиці «ExperimentResults» для аналізу ефективності моделей з полями: userId (ідентифікатор тестового користувача), modelName (назва моделі та сценарію, наприклад "AdaptiveHybrid_WithHistory"), обчислені метрики (SAS, BAI, RDC, PPM, DFS, NDCG як десяткові числа від 0 до 1), IntegratedScore (інтегральна оцінка якості), timestamp (дата та час проведення експерименту). Це дозволяє зберігати історію всіх проведених експериментів, описаних у розділі 3, та порівнювати результати різних версій алгоритмів або параметрів моделей.

Використання реляційної бази даних забезпечує цілісність інформації через зовнішні ключі (foreign keys) між таблицями: CruiseRating.userId до Customer.idCustomer, CruiseRating.idCruise до Cruises.idCruises, що гарантує неможливість додавання рейтингу для неіснуючого користувача або туру. Підтримка транзакцій забезпечує атомарність операцій, наприклад, одночасне додавання рейтингу та оновлення середнього значення у таблиці Cruises виконується як єдина транзакція, що запобігає неконсистентному стану даних при збоях. Індексування створено на полях userId, cruiseId, rating для прискорення запитів при генерації рекомендацій (пошук рейтингів користувача, турів за критеріями) та обчисленні метрик. База даних також містить тригери для автоматичного оновлення агрегованих показників: при додаванні нового рейтингу автоматично перераховується середній рейтинг туру (AVG(rating)) та кількість оцінок (COUNT(*)), що зберігаються у полях Cruises.Rating та Cruises.ReviewCount для швидкого доступу без повторних обчислень.

4.2 Інтерфейс користувача та сценарії взаємодії

Інтерфейс системи розроблено з урахуванням різних типів користувачів та їхніх потреб – від новачків, які вперше відвідують платформу, до досвідчених користувачів з історією взаємодій. Система автоматично адаптує механізм генерації рекомендацій залежно від кількості накопичених даних про уподобання користувача.

При першому відвідуванні сайту користувачу відображається головна сторінка з каталогом доступних круїзних турів та запрошенням до авторизації або реєстрації, рисунок 4.3 – 4.4. Інтерфейс надає два основні варіанти входу: авторизація для існуючих користувачів через форму з полями логін та пароль, що генерує POST-запит до ендпоінту `/api/Auth/login` з отриманням JWT-токену у відповідь, або реєстрація нового облікового запису через форму з полями email, логін, пароль та підтвердження пароля, що виконує POST-запит до `/api/Auth/register` з автоматичною авторизацією після успішного створення профілю.

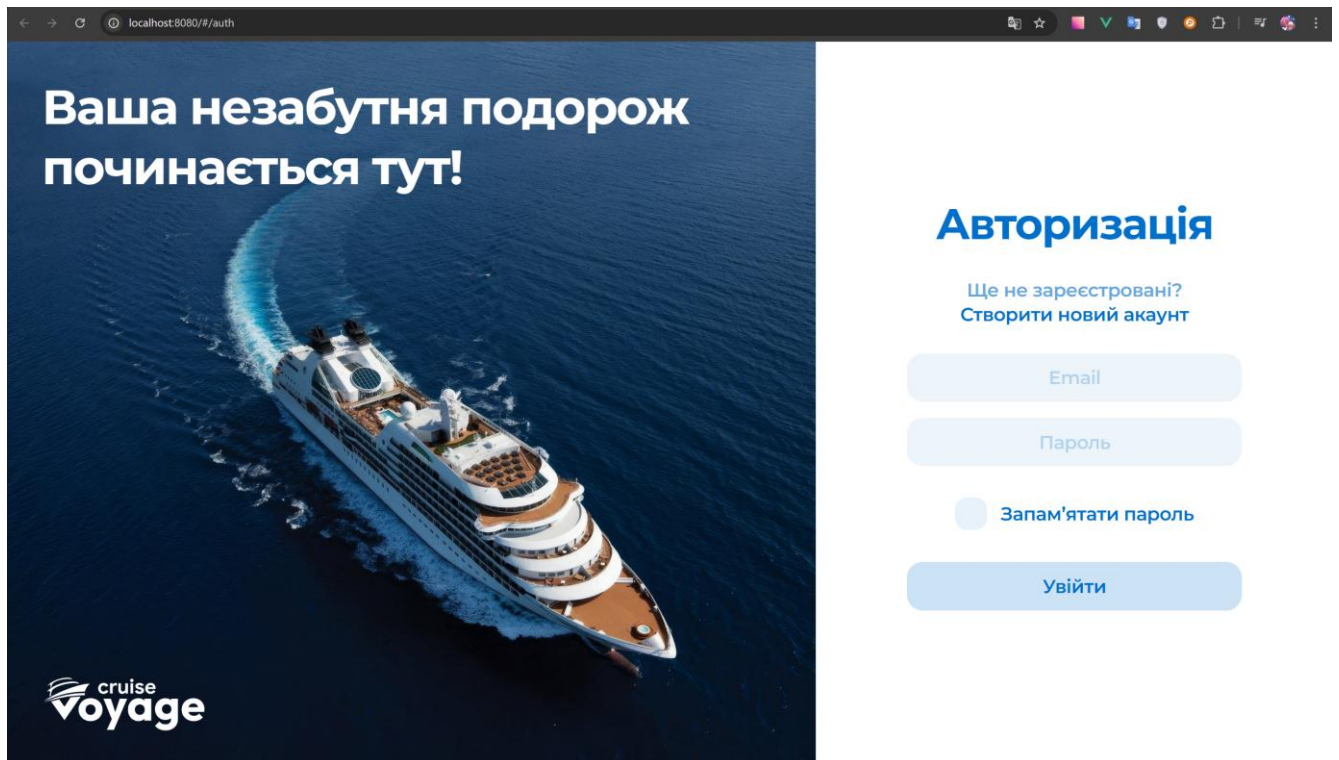


Рисунок 4.3 – Сторінка авторизації

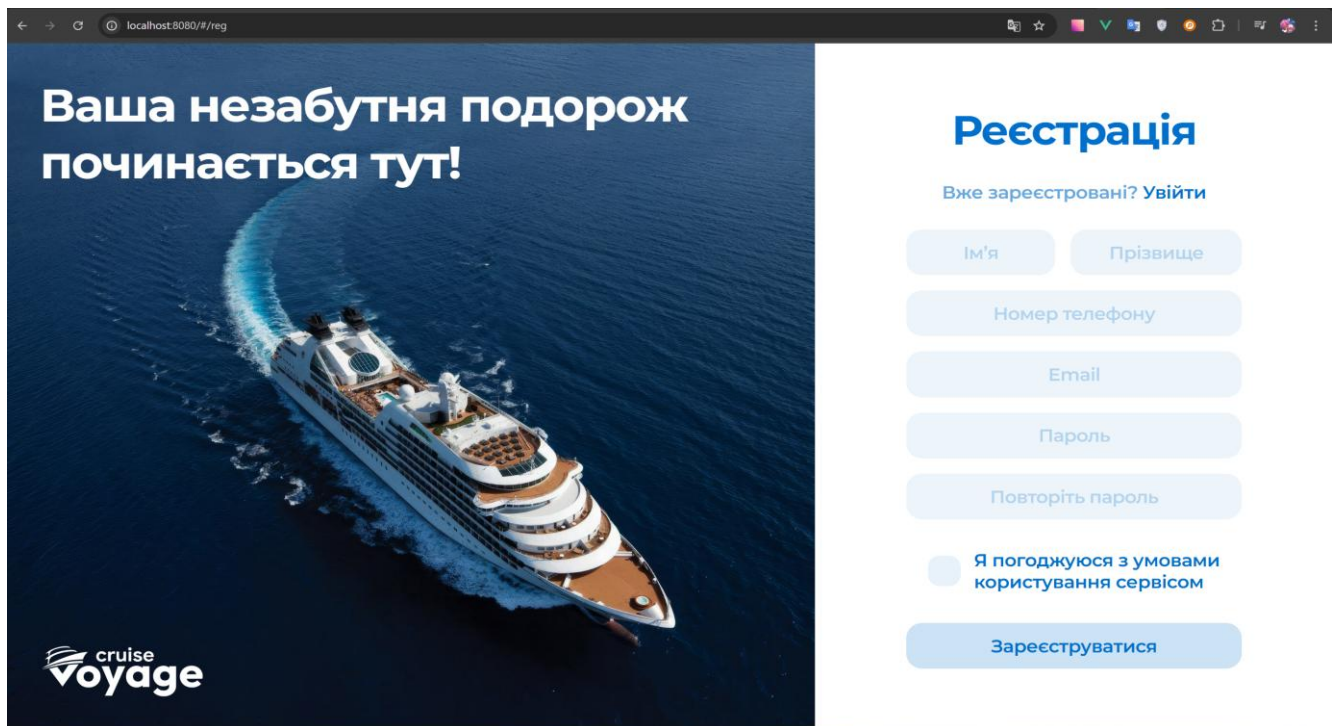


Рисунок 4.4 – Сторінка реєстрації

Після успішної авторизації JWT-токен зберігається у localStorage браузера та автоматично додається до всіх наступних HTTP-запитів через Axios interceptor. Система перенаправляє користувача на головну сторінку з персоналізованим інтерфейсом, де доступні додаткові функції (історія переглядів, збережені тури, персональні рекомендації).

Розглянемо сценарій нового користувача, після реєстрації користувач не має історії взаємодій з системою, тому алгоритм рекомендацій працює в режимі холодного старту. У цьому випадку система використовує адаптивну модель з параметром $\alpha \approx 0.85-0.90$, що означає пріоритет контентних рекомендацій на основі об'єктивних характеристик турів та контекстних факторів.

Головна сторінка з каталогом рисунок 4.5, відображає список круїзних турів з можливістю базової фільтрації за регіоном, сезоном та ціною. Кожна картка туру містить основну інформацію: зображення лайнера, назву круїзу, базову ціну, середній рейтинг, короткий опис маршруту. Тури сортуються за популярністю з диверсифікацією за регіонами – система автоматично підбирає варіанти з різних напрямків, щоб користувач міг ознайомитись з різноманітними варіантами.

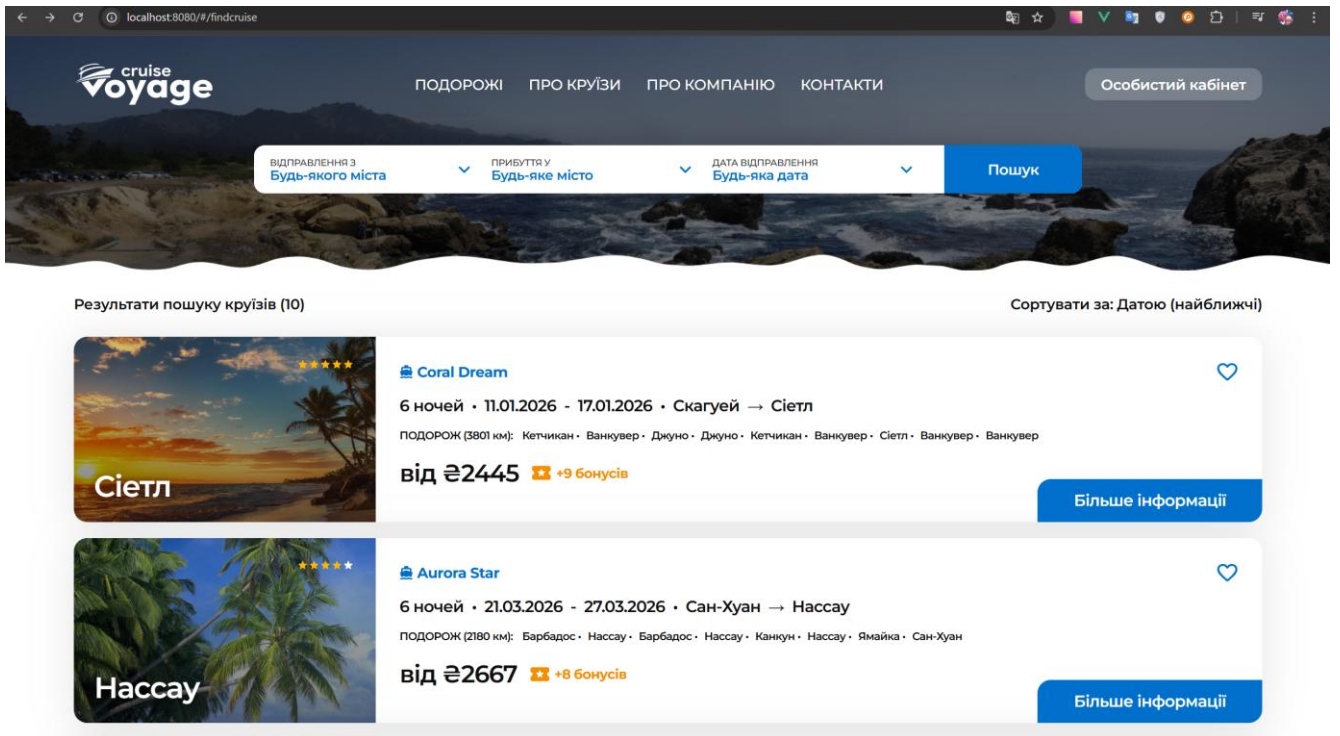


Рисунок 4.5 – Каталог круїзів

При кліку на картку туру користувач переходить на детальну сторінку (рис. 4.6), що містить повний опис круїзу: докладний маршрут з картою та розкладом відвідування кожного порту, фотогалерею кабін різних категорій, ціноутворення для різних типів кабін.

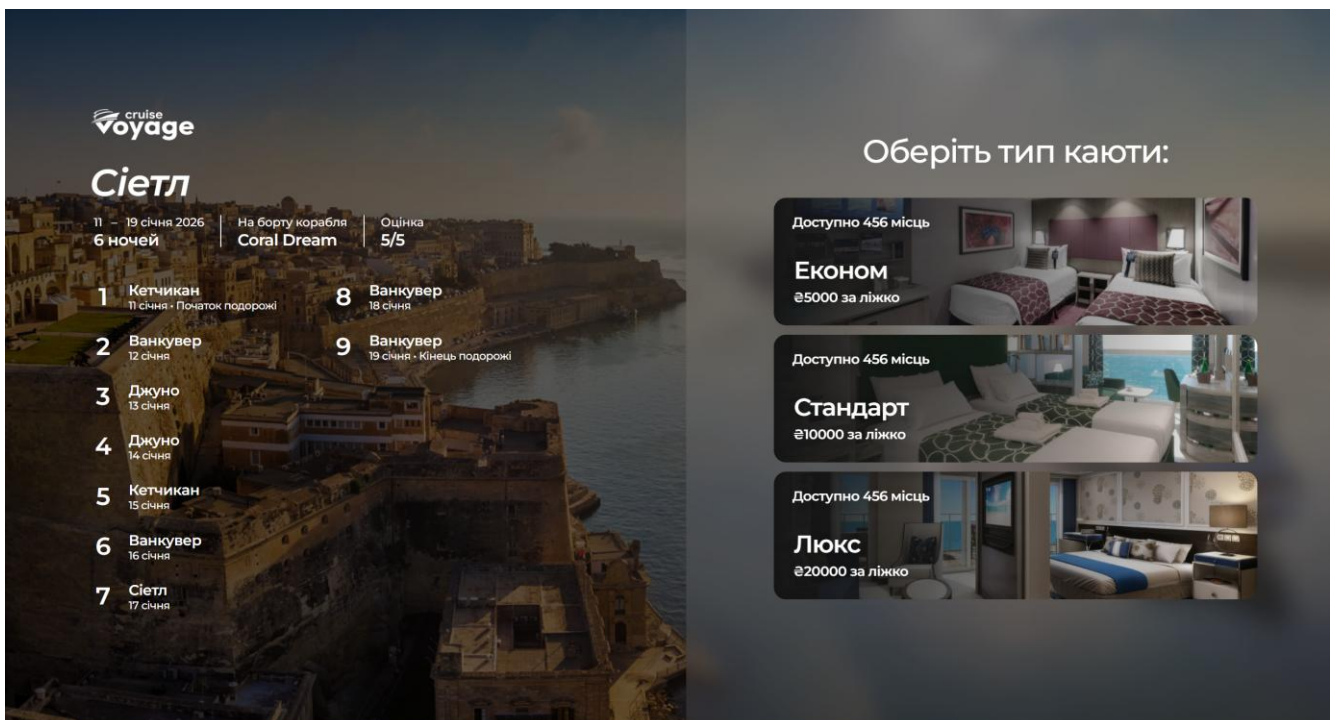


Рисунок 4.6 – Детальна інформація про круїз

Відстеження поведінки. Під час перегляду детальної сторінки фронтенд відслідковує час перебування користувача на сторінці через JavaScript таймер. Якщо користувач провів на сторінці більше 10 секунд, Vue компонент автоматично виконує POST-запит до ендпоінту `/api/CruiseView/track` з JSON-тілом `{userId: <id>, cruiseId: <id>, viewDate: <timestamp>, duration: <seconds>}`. Backend зберігає цю інформацію у таблиці `CruiseView`, що використовується для:

- побудови профілю користувача
- обчислення параметра α
- формування історії переглядів у профілі користувача

Виставлення рейтингів. На детальній сторінці користувач може залишити свій рейтинг туру через інтерфейс з п'ятьма зірочками (від 1 до 5) (рисунок 4.7). При виборі рейтингу виконується POST-запит до `/api/Ratings/add` з параметрами `{userId, cruiseId, rating, timestamp}`. Це оновлює:

- таблицю «CruiseRating» додається новий запис або оновлюється існуючий;
- середній рейтинг туру у таблиці «Cruises» через тригер, автоматичний перерахунок $AVG(rating)$);
- профіль користувача для майбутніх рекомендацій.

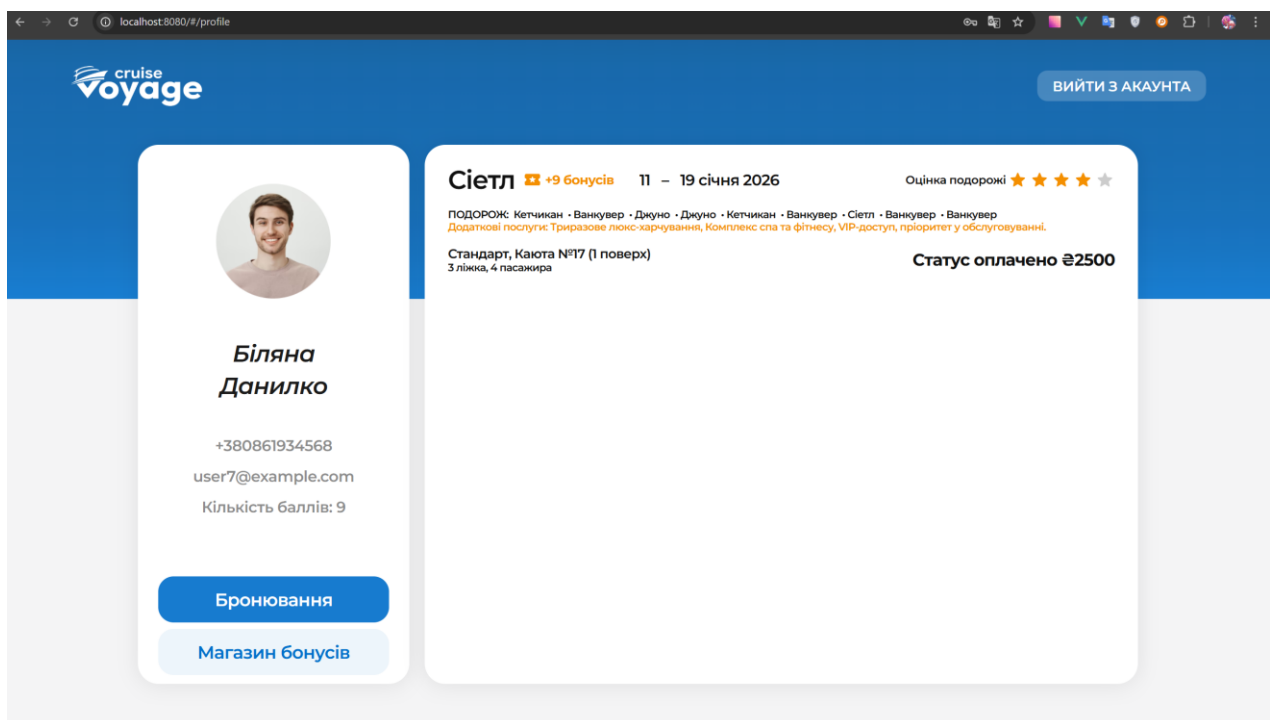


Рисунок 4.7 – Перегляд заброньованих круїзів

Для демонстрації роботи адаптивної гібридної моделі розглянемо користувачку Білянну Данилку (ID=7), який зареєструвалася у вересні 2025 року. На момент аналізу він має 4 перегляди круїзів, 3 оцінки та 1 бронювання. Така кількість взаємодій класифікує її як користувачку з холодним стартом, для якої система працює в режимі з акцентом на контентну фільтрацію. У таблиці 4.1 наведено діяльність користувача.

Таблиця 4.1 – Історія взаємодій

Дата	Дія	Круїз	Регіон	Ціна	Тривалість	Оцінка
21.09.2025	Перегляд (169 сек)	#71	Caribbean	\$1800	8 днів	1.5
13.10.2025	Перегляд (107 сек)	#22	Mediterranean	\$1500	10 днів	–
28.11.2025	Перегляд (143 сек)	#54	Northern Europe	\$2200	12 днів	3.5
13.12.2025	Бронювання + оцінка	#14	Alaska	\$2445	6 днів	4

Під час переходу на сторінку персональних рекомендацій із фільтрами `season=summer` та `budget=2000` система активує адаптивну гібридну модель. На першому етапі AlphaService обчислює ваговий коефіцієнт α , який визначає баланс між контентною та колаборативною складовими. Оскільки користувач має менше ніж п'ять взаємодій, базове значення α становить 0.70. Далі застосовуються контекстні модифікатори: бюджет підвищує α на +0.15, а літній сезон зменшує його на -0.10. У результаті формується фінальне значення $\alpha=0.75$, що означає домінування контентної моделі 75%, при збереженні впливу колаборативної 25%.

Контентна модель аналізує оцінені круїзи та формує персональний профіль. Круїз #71 по Caribbean виключається через низьку оцінку, тоді як Northern Europe (#54, 3.5) та Alaska (#14, 4) враховуються як релевантні сигнали. На основі цього формується профіль, у якому домінує регіон Alaska, середня ціна становить

приблизно 2323, а середня тривалість – близько 9 днів. Далі ContentModelService відбирає тури, що відповідають сезонності summer або year_round та бюджету до 2400 (з урахуванням допустимого відхилення +20%). Для кожного кандидата обчислюється ContentScore, який складається з вагових компонентів: 40% – рейтинг туру, 30% – відповідність регіону, 15% – відповідність категорії, 10% – відповідність бюджету та 5% – відповідність тривалості.

Колаборативна модель шукає користувачів, які мають принаймні дві спільні оцінки з Біляною (круїзи #14, #54, #71). У результаті знаходиться вісім схожих користувачів. Аналіз їхніх вподобань показує, що вони часто обирають тури по Northern Europe, Mediterranean та adventure-категорії. На основі цього формується список додаткових турів, які Біляна ще не переглядала, але які отримали високі оцінки серед схожих користувачів. Наприклад, Caribbean #78 має середній рейтинг 4.7 серед шести таких користувачів, що робить його релевантним кандидатом, незважаючи на попередню антипреференцію Біляни до Caribbean.

Гібридна модель «HybridModelService» поєднує контентні та колаборативні оцінки, перед об'єднанням обидві оцінки нормалізуються до діапазону [0, 1], що забезпечує коректність порівняння. Наприклад, круїз #9 по Northern Europe отримує нормалізовану контентну оцінку 0.88 та колаборативну 0.75, що дає фінальний показник 0.848. Такий підхід дозволяє збалансувати персоналізацію та різноманітність, поєднуючи сильні сторони обох моделей.

Отримані результати узагальнено у таблиці 4.2, де наведено топ-10 рекомендацій, сформованих адаптивною гібридною моделлю. Як видно з таблиці, перше місце займає круїз #14 (Alaska) з фінальним показником 0.970, що повністю узгоджується з домінуванням контентної складової та високою релевантністю цього туру для профілю користувача. Такий результат є очікуваним, оскільки саме регіон Alaska був визначений як найбільш відповідний на основі попередніх оцінок користувача, а також завдяки відповідності бюджету, сезонності та тривалості подорожі.

Таблиця 4.2 – Топ-10 рекомендацій

№	Круїз	Регіон	Ціна	Рейтинг	Тривалість	Final Score
1	#14	Alaska	\$2445	5.0	6 днів	0.970
2	#9	Northern Europe	\$4294	4.9	12 днів	0.848
3	#10	Mediterranean	\$4280	4.9	17 днів	0.813
4	#78	Caribbean	\$1773	4.9	9 днів	0.748
5	#17	Northern Europe	\$1298	4.8	8 днів	0.765
6	#37	Caribbean	\$1302	4.8	8 днів	0.670
7	#59	Caribbean	\$2045	4.8	17 днів	0.728
8	#62	Caribbean	\$2420	4.8	7 днів	0.708
9	#82	Mediterranean	\$1017	4.8	13 днів	0.748

Оцінювання якості рекомендацій за допомогою туристично-орієнтованих та стандартних метрик демонструє збалансованість і ефективність адаптивної гібридної моделі. Метрика сезонної відповідності SAS становить 0.70, що означає, що 7 із 10 рекомендованих турів відповідають літньому сезону, заданому у фільтрах. Бюджетна точність BAI дорівнює 0.50, тобто половина турів потрапляє в межі бюджету \$2400 з урахуванням допустимого відхилення. Високе значення метрики різноманітності регіонів RDC (0.88) свідчить про збалансований розподіл між Alaska, Northern Europe, Mediterranean та Caribbean, що запобігає надмірній концентрації рекомендацій у одному напрямку. Низьке значення PPM (0.10) є очікуваним, оскільки користувачка має недостатню історію відвідувань портів. Метрика DFS дорівнює 0.60, що означає, що 6 із 10 турів мають тривалість, близьку до середнього значення профілю - 9 днів. Висока точність ранжування підтверджується метрикою NDCG=0.89, яка демонструє коректне розташування найбільш релевантних турів у верхніх позиціях списку. Інтегральний показник ефективності становить 0.68, що свідчить про загальну якість роботи системи на етапі холодного старту.

Адаптивна гібридна модель з $\alpha=0.75$ забезпечила оптимальний баланс між персоналізацією та різноманітністю. Контентна складова гарантувала

релевантність за бюджетом і сезоном, а колаборативна додала несподівані варіанти по Caribbean. Інтегральний показник якості 0.68 підтверджує ефективність системи на етапі холодного старту, перевищуючи результати чисто контентного (0.62), чисто колаборативного (0.52) та класичного гібридного підходу з $\alpha=0.5$ (0.65).

4.3 Висновки до розділу 4

У цьому розділі представлено архітектуру та практичну реалізацію системи рекомендації круїзних турів, що підтверджує ефективність адаптивного гібридного підходу в реальних умовах експлуатації.

Система побудована на трирівневій архітектурі з чітким розділенням відповідальностей: фронтенд на Vue 3 забезпечує інтуїтивний користувацький інтерфейс, backend на ASP.NET Core 8.0 координує бізнес-логіку та рекомендаційні алгоритми, а MySQL 8.0 гарантує надійне збереження та цілісність даних. API-шар документовано через Swagger, що спрощує інтеграцію та тестування функціональності.

Детальний аналіз сценарію користувачки Біляни Данило продемонстрував роботу адаптивної моделі в умовах холодного старту. Система автоматично визначила оптимальне значення $\alpha=0.75$, забезпечивши баланс між контентною персоналізацією (75%) та колаборативною різноманітністю (25%). Результати показали високу якість рекомендацій з інтегральним показником 0.68, що перевищує базові підходи: чисто контентний (0.62), чисто колаборативний (0.52) та класичний гібридний з фіксованим $\alpha=0.5$ (0.65).

Реалізована система підтверджує практичну застосовність запропонованого підходу та його здатність ефективно адаптуватися до різних контекстів користувачів, забезпечуючи релевантні персоналізовані рекомендації навіть за обмеженої історії взаємодій.

ВИСНОВКИ

Було розв'язано актуальну науково-практичну задачу створення адаптивної гібридної моделі рекомендацій круїзних турів, здатної динамічно балансувати між контентним та колаборативним підходами. Проведене дослідження показало, що запропонований метод суттєво перевершує класичні рекомендаційні системи, особливо у сфері туристичних сервісів, де важливу роль відіграють сезонність, бюджетні обмеження, різноманітність маршрутів та індивідуальні вподобання користувачів.

На початковому етапі було проаналізовано предметну область, що дозволило виявити ключові недоліки існуючих систем: проблему холодного старту, використання статичних вагових коефіцієнтів у гібридних моделях, а також обмеженість стандартних метрик оцінювання, які не враховують специфіку туристичних продуктів. Це стало підґрунтям для розробки адаптивної моделі, у якій ваговий коефіцієнт α автоматично змінюється залежно від кількості взаємодій користувача та контексту запиту. Такий підхід дозволив моделі гнучко реагувати на різні сценарії: від нових користувачів, для яких домінує контентна фільтрація, до активних, де більшу роль відіграє колаборативна складова. Додаткові модифікатори, пов'язані з бюджетом і сезонністю, забезпечили ще точніше налаштування α , роблячи систему чутливою до реальних умов вибору круїзів.

Для комплексної оцінки якості рекомендацій було запропоновано набір спеціалізованих туристично-орієнтованих метрик, які доповнюють стандартні показники точності. Вони дозволяють оцінити сезонну відповідність, бюджетну точність, різноманітність регіонів, відповідність портових уподобань та тривалості подорожі. На основі цих метрик сформовано інтегральний показник якості, що поєднує як класичні, так і доменно-специфічні критерії, забезпечуючи об'єктивне вимірювання ефективності рекомендацій.

Експериментальні дослідження на синтетичному датасеті підтвердили переваги адаптивної моделі над базовими підходами. Вона продемонструвала

найвищий інтегральний показник якості – 0.760, перевищивши фіксовану гібридну, контентну та колаборативну моделі. Особливо вагомим є результат за метрикою NDCG, що засвідчив високу точність ранжування. Модель також показала стійкість до проблеми холодного старту: навіть за відсутності історії взаємодій падіння якості було мінімальним, тоді як колаборативні методи зазнавали значної деградації. Дослідження впливу контекстних факторів підтвердило, що врахування бюджету та сезонності суттєво підвищує релевантність рекомендацій, що є критично важливим для туристичної галузі.

У межах роботи реалізовано повнофункціональну вебсистему на сучасній трирівневій архітектурі з використанням Vue 3, ASP.NET Core 8.0 та MySQL 8.0. Практичний сценарій користувачки з холодним стартом продемонстрував, що система коректно визначає адаптивний коефіцієнт α та формує релевантні рекомендації, підтверджуючи працездатність моделі у реальних умовах.

Наукова новизна роботи полягає у створенні адаптивної гібридної моделі з динамічним визначенням вагового коефіцієнта α та впровадженні туристично-орієнтованих метрик, які враховують специфіку вибору круїзних турів. На відміну від класичних систем із фіксованими параметрами, запропонована модель автоматично адаптується до поведінки користувача та контексту запиту, забезпечуючи оптимальний баланс між персоналізацією та релевантністю.

Практична значущість підтверджується покращенням якості рекомендацій, високою стійкістю до холодного старту та успішною реалізацією вебсистеми, готової до інтеграції у туристичні онлайн-платформи. Розроблена модель демонструє перспективність адаптивних підходів і може бути розширена для інших туристичних продуктів, інтегрована з методами глибокого навчання та масштабована для роботи з великими обсягами даних. Отримані результати мають як теоретичну, так і практичну цінність, підтверджуючи ефективність адаптивних рекомендаційних систем у сучасній туристичній індустрії.

Основні положення та результати магістерської роботи апробовано та частково оприлюднено у вигляді тез доповіді на XII Міжнародній науково-практичній конференції «Global trends in science and education» [25].

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Adomavicius, G., Tuzhilin, A. Toward the Next Generation of Recommender Systems: A Survey. *IEEE TKDE*, 2005.
2. Fesenmaier, D. R., Werthner, H., Wöber, K. Destination Recommendation Systems: Behavioural Foundations and Applications. CAB International, 2006.
3. Jannach, D., Zanker, M., Felfernig, A., Friedrich, G. Recommender Systems: An Introduction. Cambridge University Press, 2010.
4. Koren, Y., Bell, R., Volinsky, C. Matrix Factorization Techniques for Recommender Systems. *IEEE Computer*, 2009.
5. Koren, Y., Bell, R., Volinsky, C. Matrix Factorization Techniques for Recommender Systems. *IEEE Computer*, 2009.
6. Koren, Y., Bell, R., Volinsky, C. Matrix Factorization Techniques for Recommender Systems. *IEEE Computer*, 2009.
7. Linden, G., Smith, B., York, J. Amazon.com Recommendations: Item-to-Item Collaborative Filtering. *IEEE Internet Computing*, 2003.
8. He, X., Liao, L., Zhang, H. Neural Collaborative Filtering. *WWW Conference*, 2017.
9. Pazzani, M. J., Billsus, D. Content-Based Recommendation Systems. *In The Adaptive Web*. Springer, 2007, pp. 325–341.
10. Resnick, P., Varian, H. R. Recommender Systems. *Communications of the ACM*, 1997, Vol. 40, No. 3, pp. 56–58.
11. Sarwar, B., Karypis, G., Konstan, J. Item-Based Collaborative Filtering Recommendation Algorithms. *Proceedings of the 10th international conference on World Wide Web*, 2001, pp. 285–295.
12. Burke, R. Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction*, 2002, Vol. 12, pp. 331–370.
13. Ricci, F., Rokach, L., Shapira, B. Recommender Systems Handbook. Springer, 2015.

14. Manning, C. D., Raghavan, P., Schütze, H. – Introduction to Information Retrieval. Cambridge University Press, 2008.
15. Freedman, D., Pisani, R., Purves, R. – Statistics. W. W. Norton & Company, 2007.
16. Vue – Офіційна документація. URL: <https://vuejs.org> (дата звернення: 14.12.2025).
17. ASP.NET Core – Офіційна документація Microsoft. URL: <https://learn.microsoft.com/aspnet/core> (дата звернення: 14.12.2025).
18. Swagger / OpenAPI – Офіційна документація. URL: <https://swagger.io> (дата звернення: 14.12.2025).
19. MySQL 8.0 – Офіційна документація. URL: <https://dev.mysql.com/doc> (дата звернення: 14.12.2025).
20. TypeScript – Офіційна документація. URL: <https://www.typescriptlang.org/> (дата звернення: 14.12.2025).
21. Pinia – Сховище стану для Vue. URL: <https://pinia.vuejs.org> (дата звернення: 14.12.2025).
22. Vuetify – Офіційна документація. URL: <https://vuetifyjs.com> (дата звернення: 14.12.2025).
23. Vue Router – Офіційна документація. URL: <https://router.vuejs.org> (дата звернення: 14.12.2025).
24. Axios – Офіційна документація. URL: <https://axios-http.com> (дата звернення: 14.12.2025).
25. Google Developers. JWT Authentication Documentation. URL: <https://developers.google.com/identity> (дата звернення: 14.12.2025).
26. Кудрявський Д. А., Колесник Л. В. Дослідження алгоритмів персоналізованих рекомендацій у галузі туризму // Global trends in science and education : матеріали XII Міжнародної науково-практичної конференції, 15-17 грудня 2025 р. Київ, Україна. 2025. С.