

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет _____ Комп'ютерних наук
(повна назва)

Кафедра _____ Інформаційних управляючих систем
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти _____ другий (магістерський)
Дослідження методів персоналізації рекомендацій на основі
поведінкових даних користувачів в ІТ-проектах електронної комерції
(тема)

Виконав:

здобувач _____ 2 _____ року навчання,
групи _____ УПГІТм-24-1

Вікторія СИНЬОВА
(власне ім'я, ПРІЗВИЩЕ)

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

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

Освітня програма _____ Управління проектами в
галузі інформаційних технологій
(повна назва освітньої програми)

Керівник _____ доц. каф. ІУС Тетяна БІЛОВА
(посада, власне ім'я, ПРІЗВИЩЕ)

Допускається до захисту

Зав. кафедри ІУС


(підпис)

Костянтин ПЕТРОВ
(власне ім'я, ПРІЗВИЩЕ)

2026 р.

Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
 Кафедра Інформаційних управляючих систем
 Рівень вищої освіти другий (магістерський)
 Спеціальність 122 Комп'ютерні науки
 (код і повна назва)
 Тип програми освітньо-наукова
 (освітньо-професійна або освітньо-наукова)
 Освітня програма Управління проектами в галузі інформаційних технологій
 (повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри

(підпис)

« 6 » квітня 2026 р.

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві Синьовій Вікторії Олексіївні
 (прізвище, ім'я, по батькові)

1. Тема роботи «Дослідження методів персоналізації рекомендацій на основі поведінкових даних користувачів в IT-проектах електронної комерції»

затверджена наказом по університету від « 23 » березня 2026 р. № 231Ст

2. Термін подання здобувачем роботи до екзаменаційної комісії « 15 » травня 2026 р.

3. Вихідні дані до роботи науково-технічні публікації, монографії та інтернет-ресурси з тематики рекомендаційних систем і методів персоналізації; матеріали щодо аналізу поведінкових даних користувачів в інформаційних системах електронної комерції; матеріали науково-дослідної практики.

4. Перелік питань, що потрібно опрацювати в роботі аналіз методів персоналізації рекомендацій у системах електронної комерції, дослідження поведінкових даних користувачів, обґрунтування підходу до персоналізації, розробка гібридного методу формування рекомендацій, моделювання та реалізація рекомендаційної системи, проведення експериментів та аналіз отриманих результатів.

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	При- мітка
1	Аналіз предметної області електронної комерції в контексті ІТ-проектів	06.04.2026 – 08.04.2026	вик.
2	Аналіз та класифікація методів персоналізації рекомендацій	09.04.2026 – 12.04.2026	вик.
3	Дослідження поведінкових даних користувачів як основи персоналізації	13.04.2026 – 16.04.2026	вик.
4	Аналіз недоліків існуючих підходів та постановка задачі дослідження	17.04.2026 – 20.04.2026	вик.
5	Розробка методу формування поведінкового профілю користувача	21.04.2026 – 26.04.2026	вик.
6	Розробка моделі персоналізації рекомендацій на основі поведінкових даних	27.04.2026 – 02.05.2026	вик.
7	Розробка та дослідження гібридного методу персоналізації рекомендацій	03.05.2026 – 06.05.2026	вик.
8	Апробація методу та оцінювання ефективності рекомендаційної системи	07.05.2026 – 08.05.2026	вик.
9	Оформлення пояснювальної записки	09.05.2026 – 13.05.2026	вик.
10	Підготовка презентаційного матеріалу	14.05.2026	вик.
11	Перевірка роботи на плагіат	15.05.2026	вик.
12	Захист кваліфікаційної роботи	21.05.2026	вик.

Дата видачі завдання 6 квітня 2026 р.

Здобувач



(підпис)

Керівник роботи



(підпис)

доц. каф. ІУС Тетяна БІЛОВА

(посада, власне ім'я, ПРІЗВИЩЕ)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 118 с., 26 рис., 22 форм., 23 табл., 1 дод., 40 джерел.

ГІБРИДНИЙ МЕТОД, ЕЛЕКТРОННА КОМЕРЦІЯ, ІТ-ПРОЄКТ, МЕТОД РЕКОМЕНДАЦІЙ, МОДЕЛЬ, ПЕРСОНАЛІЗАЦІЯ РЕКОМЕНДАЦІЙ, ПОВЕДІНКОВІ ДАНІ, РЕКОМЕНДАЦІЙНА СИСТЕМА.

Об'єктом дослідження кваліфікаційної роботи є процеси формування персоналізованих рекомендацій у системах електронної комерції в рамках ІТ-проєктів.

Предметом дослідження є методи та моделі персоналізації рекомендацій на основі поведінкових даних користувачів.

Метою кваліфікаційної роботи є розробка та дослідження гібридного методу персоналізації рекомендацій на основі поведінкових даних користувачів для підвищення якості та релевантності рекомендацій в ІТ-проєктах електронної комерції.

Методи дослідження: методи аналізу даних, статистичні методи, методи рекомендаційних систем, порівняльний аналіз.

Основними результатами кваліфікаційної роботи є розробка гібридного методу персоналізації рекомендацій, який базується на аналізі поведінкових даних користувачів та поєднує різні підходи до формування рекомендацій. Проведено дослідження особливостей реалізації розробленого методу в ІТ-проєктах електронної комерції. Проведена апробація запропонованого методу, яка підтвердила підвищення ефективності рекомендаційної системи за рахунок покращення релевантності рекомендацій.

Наукова новизна роботи полягає у розробці гібридного методу персоналізації рекомендацій, який поєднує колаборативну та контентну фільтрацію з динамічним зважуванням поведінкових сигналів користувачів з урахуванням стадії життєвого циклу ІТ-проєкту електронної комерції, що забезпечує підвищення релевантності рекомендацій в умовах неповних та зашумлених поведінкових даних.

ABSTRACT

Master's thesis: 118 pages, 26 figures, 22 formulas, 23 tables, 1 appendix, 40 sources.

HYBRID METHOD, ELECTRONIC COMMERCE, IT PROJECT, RECOMMENDATION METHOD, MODEL, RECOMMENDATION PERSONALIZATION, BEHAVIORAL DATA, RECOMMENDER SYSTEM.

The object of the qualification research is the processes of forming personalized recommendations in e-commerce systems within IT projects.

The subject of the research is methods and models of recommendation personalization based on user behavioral data.

The purpose of the qualification work is to develop and investigate a hybrid method for recommendation personalization based on user behavioral data in order to improve the quality and relevance of recommendations in e-commerce IT projects.

Research methods include data analysis methods, statistical methods, recommender system methods, comparative analysis.

The main results of the qualification work include the development of a hybrid method for recommendation personalization, which is based on the analysis of user behavioral data and combines different approaches to recommendation generation.

The research explores the features of implementing the proposed method in e-commerce IT projects.

The proposed method has been tested, confirming an improvement in the efficiency of the recommender system due to increased relevance of recommendations.

The scientific novelty of the work consists in the development of a hybrid recommendation personalization method that combines collaborative and content-based filtering with dynamic weighting of user behavioral signals according to the stage of the e-commerce IT project lifecycle, thereby improving recommendation relevance under conditions of incomplete and noisy behavioral data.

ЗМІСТ

	С.
Скорочення та умовні позначки.....	8
Вступ.....	9
1 Аналіз методів персоналізації рекомендацій на основі поведінкових даних користувачів в ІТ-проєктах електронної комерції та постановка задачі дослідження	11
1.1 Аналіз предметної області електронної комерції в контексті ІТ-проєктів....	11
1.2 Роль та місце рекомендаційних систем в ІТ-проєктах електронної комерції.....	15
1.3 Аналіз поведінкових даних користувачів як основи персоналізації	20
1.4 Огляд сучасних методів побудови рекомендаційних систем.....	23
1.5 Аналіз застосування методів персоналізації в ІТ-проєктах.....	28
1.6 Виявлення недоліків існуючих методів і підходів	32
1.7 Формування проблеми дослідження та обґрунтування напряду її вирішення.....	35
1.8 Постановка задачі дослідження	38
2 Метод персоналізації рекомендацій на основі поведінкового профілю користувача в ІТ-проєктах електронної комерції.....	41
2.1 Формування поведінкового профілю користувача на основі явних та неявних даних.....	41
2.2 Метод зважування поведінкових сигналів з урахуванням стадії ІТ-проєкту	48
2.3 Модель персоналізації рекомендацій на основі поведінкового профілю	56
3 Практична реалізація гібридного методу персоналізації рекомендацій в ІТ-проєкті електронної комерції	67

3.1 Архітектура та інструментальні засоби реалізації рекомендаційної системи	67
3.2 Реалізація гібридного методу персоналізації рекомендацій	72
3.3 Експериментальне дослідження й оцінювання ефективності розробленого методу	78
4 Практичне застосування результатів дослідження в управлінні ІТ-проєктами електронної комерції	87
4.1 Рекомендації щодо впровадження рекомендаційної системи в ІТ-проєкт	87
4.2 Оцінювання ефекту від впровадження методу персоналізації в контексті проєктного управління.....	92
Висновки	101
Перелік джерел посилання	103
Додаток А Графічний матеріал кваліфікаційної роботи.....	108

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ

ЕК	– електронна комерція
ІТ	– інформаційна технологія
ІУС	– інформаційні управляючі системи
РС	– рекомендаційна система
ХНУРЕ	– Харківський національний університет радіоелектроніки
API	– Application Programming Interface
B2B	– Business-to-Business
B2C	– Business-to-Consumer
BPR	– байєсівське ранжування для персоналізації
C2C	– Consumer-to-Consumer
CBF	– Content-Based Filtering
CF	– Collaborative Filtering
CTR	– Click-Through Rate
E	– явні сигнали
I	– неявні сигнали
ItemCF	– колаборативна фільтрація на основі елементів
LTV	– Lifetime Value
PMBOK	– Project Management Body of Knowledge
Static Hybrid	– статичний гібрид

ВСТУП

Сучасний етап розвитку інформаційних технологій характеризується стрімким зростанням обсягів даних та активним впровадженням цифрових рішень у сферу електронної комерції (ЕК). У таких умовах рекомендаційні системи (РС) виступають одним із ключових інструментів підвищення ефективності ІТ-проектів, забезпечуючи персоналізацію контенту та адаптацію пропозицій до індивідуальних потреб користувачів. Провідні компанії у сфері ЕК активно використовують рекомендаційні механізми для підвищення конверсії, утримання користувачів та зростання прибутковості, що підтверджує практичну значущість даного напрямку.

Аналіз сучасного стану досліджуваної предметної області показує, що, незважаючи на значний розвиток РС, існуючі підходи мають низку обмежень. Багато з них орієнтовані на обробку статичних або слабо структурованих даних і недостатньо враховують динамічний характер поведінки користувачів. Крім того, проблемами залишаються неповнота та зашумленість поведінкових даних, а також складність їх інтеграції в архітектуру проектів ЕК галузі інформаційних технологій (ІТ).

Світові тенденції розвитку РС свідчать про перехід від класичних підходів, таких як Collaborative Filtering (CF), знання-орієнтована фільтрація та Content-Based Filtering (CBF), до використання гібридних моделей і методів машинного та глибокого навчання. Такі підходи дозволяють підвищити точність і релевантність рекомендацій, однак потребують значних обчислювальних ресурсів і ефективних методів обробки великих обсягів поведінкових даних.

У зв'язку з цим актуальним є розроблення нових або вдосконалення існуючих методів персоналізації, що враховують як особливості поведінкових даних, так і специфіку ІТ-проектного середовища.

Актуальність роботи обумовлена необхідністю та своєчасністю вдосконалення методів персоналізації рекомендацій в умовах динамічного розвитку електронної комерції. Необхідність дослідження визначається обмеженнями існуючих підходів та потребою підвищення ефективності рекомендаційних систем, а своєчасність – стрімким зростанням обсягів поведінкових даних і розвитком ІТ-проектів у сфері ЕК.

Метою кваліфікаційної роботи є розробка та дослідження гібридного методу персоналізації рекомендацій на основі поведінкових даних користувачів для підвищення якості та релевантності рекомендацій в ІТ-проектах ЕК.

Для досягнення поставленої мети у кваліфікаційній роботі необхідно вирішити такі задачі:

- виконати аналіз предметної області ЕК в контексті ІТ-проектів;
- визначити роль та місце РС у сучасних інформаційних системах;
- проаналізувати поведінкові дані користувачів як основу персоналізації;
- здійснити огляд сучасних методів побудови РС та виявити їх недоліки;
- обґрунтувати вибір підходу до персоналізації рекомендацій;
- розробити гібридний метод персоналізації рекомендацій;
- провести експериментальне дослідження та оцінити ефективність запропонованого методу.

Об'єктом дослідження є процеси персоналізації рекомендацій у ІТ-проектах електронної комерції.

Предметом дослідження є методи та моделі персоналізації рекомендацій на основі поведінкових даних користувачів.

Роботу представлено й апробовано на 29-му та 30-му Міжнародних молодіжних форумах «Радіоелектроніка та молодь у XXI столітті» [1, 2].

1 АНАЛІЗ МЕТОДІВ ПЕРСОНАЛІЗАЦІЇ РЕКОМЕНДАЦІЙ НА ОСНОВІ ПОВЕДІНКОВИХ ДАНИХ КОРИСТУВАЧІВ В ІТ-ПРОЄКТАХ ЕЛЕКТРОННОЇ КОМЕРЦІЇ ТА ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕННЯ

1.1 Аналіз предметної області електронної комерції в контексті ІТ-проектів

У сучасних умовах розвитку цифрової економіки ЕК стала однією з ключових сфер застосування інформаційних технологій. Практично всі великі компанії, а також значна кількість малого та середнього бізнесу, використовують онлайн-платформи для продажу товарів і послуг, що обумовлює постійне зростання кількості ІТ-проектів, спрямованих на створення, розвиток та підтримку електронних торгових систем.

ІТ-проекти ЕК функціонують у середовищі, де все змінюється дуже швидко. Поведінка покупців, тренди, а також те, як люди гортають сторінки в додатку – все це постійно трансформується. Якщо команда розробників занадто довго виводить нову функціональність, вона ризикує вийти на ринок уже з неактуальним рішенням. Тому такі проекти часто будуються за гнучкими методологіями, де важливі швидкі ітерації та зворотний зв'язок.

У сучасній практиці ЕК реалізується у вигляді різних моделей взаємодії, серед яких найбільш поширеними є Business-to-Consumer (B2C), Business-to-Business (B2B) та Consumer-to-Consumer (C2C) [3]. Найбільш характерною для ІТ-проектів персоналізації є модель B2C, де взаємодія відбувається безпосередньо між бізнесом і кінцевим користувачем.

З технічної точки зору платформи ЕК зазвичай реалізуються як багаторівневі інформаційні системи, що включають користувацький інтерфейс у вигляді веб або мобільного додатку, серверну частину, бази даних, а також аналітичні та рекомендаційні модулі. Важливою особливістю таких систем є необхідність

обробки великих обсягів даних у реальному часі, що обумовлює використання масштабованих архітектурних рішень.

Крім того, ЕК характеризується впливом зовнішніх факторів, зокрема сезонності попиту, маркетингових кампаній та поведінкових трендів користувачів, що призводить до нерівномірного навантаження на систему та потреби швидкої адаптації ІТ-рішень до змін середовища. Усе це безпосередньо впливає на вимоги до РС, які повинні враховувати як динаміку даних, так і контекст їх формування.

ЕК охоплює широкий спектр процесів, пов'язаних із взаємодією користувача з цифровою платформою. До таких процесів належать пошук товарів, перегляд каталогу, формування замовлення, оплата, а також післяпродажне обслуговування. Кожен із цих етапів супроводжується активною взаємодією користувача з інформаційною системою, що, у свою чергу, генерує значні обсяги поведінкових даних. Основні етапи взаємодії користувача з системою ЕК та відповідні процеси генерації поведінкових даних зображено на рисунку 1.1.



Рисунок 1.1 – Основні етапи взаємодії користувача з системою ЕК

З точки зору ІТ-проектів, ЕК є складною багаторівневою системою, яка включає не лише програмну частину, як, наприклад, веб-сайти та мобільні додатки, але й серверну інфраструктуру, бази даних, аналітичні модулі та інтеграції з зовнішніми сервісами. Розробка таких систем зазвичай реалізується у вигляді окремих ІТ-проектів або програм проектів, які охоплюють повний життєвий цикл.

У контексті управління ІТ-проектами доцільно розглядати зазначені етапи відповідно до підходів, описаних у стандарті Project Management Body of Knowledge (PMBOK) 8 [4]. Згідно з цим стандартом, життєвий цикл проекту структурується за групами процесів, такими як ініціація, планування, виконання, моніторинг і контроль, а також завершення.

У рамках ІТ-проектів електронної комерції процеси, пов'язані з персоналізацією, можуть бути віднесені до кількох груп одночасно. Наприклад, визначення вимог до РС відбувається на етапі планування, реалізація алгоритмів здійснюється у процесах виконання, а оцінка ефективності рекомендацій – у процесах моніторингу та контролю. Тобто персоналізація рекомендацій може розглядатися не лише як технічна задача, а як складова процесів управління ІТ-проектом, що інтегрується у різні його етапи.

На етапі ініціації ІТ-проекту в сфері ЕК визначаються основні бізнес-цілі, такі як збільшення продажів, підвищення рівня задоволеності клієнтів або оптимізація бізнес-процесів. Уже на цьому етапі формується бачення майбутньої РС, визначаються вимоги до персоналізації та перелік поведінкових даних, які планується збирати після запуску системи. Водночас на ранніх етапах життєвого циклу проекту фактичні поведінкові дані користувачів ще відсутні або мають обмежений обсяг [5]. Тому відмінності між підходами до персоналізації на різних етапах ІТ-проекту пов'язані не зі зміною самих алгоритмів відповідно до PMBOK 8 [4], а зі зміною доступності, повноти та структури даних, необхідних для їх коректної роботи.

На етапі планування формується архітектура системи, визначаються вимоги до функціональності, включаючи можливості рекомендаційних механізмів. У багатьох сучасних ІТ-проектах ЕК персоналізація закладається як окремий функціональний модуль або підсистема, що взаємодіє з іншими компонентами системи. Саме на цьому етапі визначається, які типи поведінкових даних будуть збиратися, їх структура, джерела та способи зберігання, що закладає основу для подальшої роботи РС.

Етап виконання передбачає реалізацію програмних рішень, інтеграцію алгоритмів обробки даних та налаштування механізмів збору поведінкової інформації. Саме на цьому етапі відбувається технічна реалізація РС, які використовують дані про дії користувачів для формування персоналізованих пропозицій. Також відбувається перше фактичне генерування поведінкових даних, оскільки система починає фіксувати взаємодію користувачів, зокрема кліки, перегляди сторінок і навігаційні дії.

На етапі моніторингу та контролю здійснюється безперервний моніторинг роботи рекомендаційної системи, вимірювання її ефективності відповідно до визначених метрик, аналіз поведінкових даних та внесення коректив у алгоритми персоналізації. Саме на цьому етапі відбувається накопичення повного обсягу поведінкових даних, включаючи транзакційні, часові та реакцію на рекомендації, що дозволяє вдосконалювати моделі. У реальних ІТ-проектах даний процес є ітеративним і безперервним, оскільки поведінка користувачів постійно змінюється, а конкуренція в сфері ЕК зростає.

На етапі завершення проводиться остаточна оцінка результатів функціонування РС, формалізація отриманих знань і напрацювань, підготовка документації для подальшої підтримки або передачі продукту. Накопичений поведінковий профіль дозволяє проаналізувати довгострокову ефективність персоналізації, визначити фактори відтоку клієнтів та сформулювати рекомендації щодо розвитку системи в рамках наступних проєктів або ітерацій.

Однією з основних особливостей ЕК є висока конкуренція між платформами. Користувач має великий вибір альтернатив, тому навіть незначні покращення в зручності користування або релевантності пропозицій можуть суттєво вплинути на його вибір. У цьому контексті персоналізація стає не просто додатковою функцією, а необхідною умовою успішності ІТ-проєкту [6].

Крім того, ЕК характеризується великими обсягами даних, які накопичуються в процесі взаємодії користувачів із системою. Ці дані включають історію переглядів, кліки, покупки, час перебування на сторінках та інші показники. Їх аналіз дозволяє отримати глибоке розуміння поведінки користувачів і є основою для побудови ефективних РС. Водночас використання поведінкових даних у рамках ІТ-проєктів пов'язане з низкою викликів. До них належать проблеми якості даних, їх неповнота, необхідність обробки в реальному часі, а також конфіденційність і захист персональної інформації. Усі ці фактори необхідно враховувати при проєктуванні та реалізації ІТ-рішень у сфері ЕК.

Саме тому ЕК як предметна область є складною динамічною системою, у якій ІТ-проєкти відіграють ключову роль у забезпеченні ефективної взаємодії між бізнесом і користувачем. Особливе місце в таких проєктах займають механізми персоналізації, що базуються на аналізі поведінкових даних і спрямовані на підвищення якості обслуговування та досягнення бізнес-цілей.

1.2 Роль та місце рекомендаційних систем в ІТ-проєктах електронної комерції

У сучасних ІТ-проєктах ЕК РС вже давно перестали бути додатковою функцією або приємним бонусом для користувача. На практиці вони виступають

одним із ключових інструментів, що безпосередньо впливає на ефективність роботи всієї системи, а також на досягнення бізнес-цілей проєкту.

Основне завдання рекомендаційних систем полягає у формуванні персоналізованих пропозицій для користувача на основі аналізу його поведінки, інтересів та взаємодії з платформою. У сфері ЕК така персоналізація може проявлятися у вигляді рекомендацій товарів, схожих позицій, підбірок «вам також може сподобатися», а також персоналізованих пропозицій на головній сторінці або в пошуковій видачі [7].

З точки зору користувача, РС допомагає швидше знайти релевантний товар серед великої кількості варіантів. З точки зору бізнесу вона підвищує конверсію, середній чек та рівень залученості користувача. Саме тому в більшості сучасних ІТ-проєктів ЕК рекомендаційні механізми інтегруються у ключові точки взаємодії з користувачем [8].

Рекомендаційні системи можуть бути вбудовані у різні етапи користувацького шляху, починаючи від пошуку товарів і закінчуючи післяпродажною взаємодією. Наприклад, під час перегляду товару користувач може отримувати рекомендації схожих позицій, а після покупки – пропозиції супутніх товарів. Таким чином, рекомендаційна система супроводжує користувача протягом усього циклу взаємодії з платформою.

З точки зору архітектури ІТ-проєкту, рекомендаційна система зазвичай реалізується як окремий модуль або підсистема, яка взаємодіє з іншими компонентами системи, зокрема з базою даних, серверною частиною та користувацьким інтерфейсом. Вона отримує вхідні поведінкові, контекстні й історичні дані, обробляє їх за допомогою відповідних алгоритмів і формує результат у вигляді персоналізованих рекомендацій.

На рисунку 1.2 зображено узагальнене місце рекомендаційної системи в архітектурі ІТ-проєкту ЕК, яке відображає один із найбільш поширених варіантів її інтеграції в систему.

З даного рисунку бачимо, що взаємодія користувача із системою відбувається через інтерфейс, який забезпечує введення запитів, перегляд інформації та отримання результатів. Інтерфейс, у свою чергу, взаємодіє із серверною частиною, передаючи запити користувача та отримуючи відповіді, необхідні для відображення контенту.



Рисунок 1.2 – Місце рекомендаційної системи в архітектурі ІТ-проєкту ЕК

Серверна частина виконує роль центрального елемента обробки, який забезпечує взаємодію з базою даних і РС. Вона здійснює читання та запис даних, що накопичуються в процесі взаємодії користувача із системою, зокрема поведінкових, як кліки та перегляди, історичних, як історія покупок, та контекстних даних.

РС отримує ці дані як від серверної частини, так і безпосередньо з бази даних. На їх основі вона виконує обробку, аналізує поведінку користувача та

формує персоналізовані рекомендації. Результати роботи РС передаються до інтерфейсу, де вони відображаються користувачу у вигляді релевантних пропозицій.

Таким чином, РС виступає своєрідним зв'язуючим елементом між даними, що накопичуються в системі, та користувачем. Вона не лише обробляє інформацію, а й перетворює її на практично корисні рішення, які допомагають користувачу швидше знайти потрібний товар або послугу. У цьому сенсі її можна розглядати як один із ключових елементів систем підтримки прийняття рішень в ІТ-проєктах [9].

Водночас слід зазначити, що наведена архітектура є типовою, але не є універсальною для всіх ІТ-проєктів. У реальних системах структура взаємодії компонентів може змінюватися залежно від обраної технологічної платформи, масштабу проєкту, вимог до продуктивності та особливостей обробки даних.

Важливо також зазначити, що ефективність роботи РС значною мірою залежить від якості та повноти даних, що зберігаються в базі даних, а також від правильності організації взаємодії між компонентами системи. Неповні або некоректні дані можуть призводити до формування нерелевантних рекомендацій, що негативно впливає на користувацький досвід і загальну ефективність ІТ-проєкту.

РС в ІТ-проєктах ЕК виконує не лише технічну, але й бізнес-орієнтовану функцію, оскільки безпосередньо впливає на ключові показники ефективності системи. Вона інтегрується в архітектуру ІТ-проєкту як окрема підсистема, яка взаємодіє з базами даних, серверною частиною та інтерфейсом користувача, що показано на рисунку 1.2.

З технічної точки зору, РС забезпечує обробку поведінкових даних та формування персоналізованих пропозицій, однак її значення виходить за межі лише алгоритмічної обробки. У сучасних ІТ-проєктах вона безпосередньо впливає на користувацький досвід та економічні показники системи.

Впровадження РС дозволяє підвищити такі показники, як Click-Through Rate (CTR), тобто показник клікабельності, рівень конверсії, середній чек та Lifetime Value (LTV) клієнта, тобто його довічна цінність. Тобто РС виступає інструментом не лише персоналізації, але й оптимізації бізнес-процесів.

Вплив рекомендаційних систем на показники в ІТ-проєктах наведено в таблиці 1.1.

Таблиця 1.1 – Вплив рекомендаційних систем на показники ІТ-проєкту

Функція системи	Метрика	Вплив на ІТ-проєкт
Персоналізація рекомендацій	CTR	Підвищення взаємодії користувача
Рекомендація релевантних товарів	Конверсія	Збільшення кількості покупок
Підвищення продажу / Перехресний продаж	Середній чек	Зростання доходу
Утримання користувачів	LTV	Довгострокова прибутковість
Адаптація контенту	Показник відмов	Зменшення відмов

Як видно з таблиці 1.1, РС впливають не лише на технічні аспекти роботи системи, але й на ключові бізнес-показники ІТ-проєкту, що підкреслює їх важливість у сучасних умовах ЕК.

Як зазначається у дослідженнях [10, 11], впровадження РС у провідних ІТ-проєктах ЕК демонструє суттєвий вплив на ключові показники ефективності. Зокрема, за даними окремих досліджень, у Amazon, до 35% продажів формується за рахунок рекомендаційних механізмів, а використання персоналізації дозволяє підвищити коефіцієнт конверсії на 10–30% [10]. У свою чергу, у системах типу Netflix рекомендаційні алгоритми забезпечують понад 80% переглядів контенту, що свідчить про їх значний вплив на залученість користувачів та показники утримання [11].

Водночас, незважаючи на очевидні переваги, впровадження РС у ІТ-проєктах супроводжується певними складнощами. До них можна віднести необхідність обробки великих обсягів даних, забезпечення швидкодії системи, а також інтеграцію з існуючою архітектурою. Крім того, важливим аспектом є забезпечення релевантності рекомендацій, що напряду залежить від якості вихідних даних і обраних методів [12].

РС займають центральне місце в ІТ-проєктах ЕК, оскільки вони поєднують у собі аналітичні, технічні та бізнес-функції. Їх ефективна реалізація дозволяє не лише покращити користувацький досвід, але й забезпечити досягнення стратегічних цілей проєкту.

1.3 Аналіз поведінкових даних користувачів як основи персоналізації

Однією з ключових передумов ефективної роботи РС у ІТ-проєктах ЕК є наявність і правильне використання поведінкових даних користувачів. Саме ці дані дозволяють зрозуміти, як користувач взаємодіє із системою, які дії виконує, що його цікавить і які рішення він приймає в процесі вибору товарів або послуг.

Поведінкові дані формуються безпосередньо під час роботи користувача з інформаційною системою. До них належать кліки, перегляди сторінок, час перебування на сайті, пошукові запити, додавання товарів до кошика, покупки, а також взаємодія з рекомендаціями. На відміну від явно заданих даних, як, наприклад, анкет або оцінок, поведінкові дані збираються автоматично, що робить їх більш об'єктивними, але водночас складнішими для обробки [13].

З точки зору ІТ-проєкту, поведінкові дані є одним із основних ресурсів, на яких будується персоналізація. Вони накопичуються в базах даних, проходять етапи обробки та аналізу, після чого використовуються РС для формування

індивідуальних пропозицій. Фактично, якість персоналізації напряду залежить від якості цих даних та способів їх обробки. Для кращого розуміння структури поведінкових даних доцільно розглянути їх основні типи, що наведені в таблиці 1.2. Як показано до цього на рисунку 1.1, у процесі взаємодії користувача з системою ЕК формуються різні типи поведінкових даних. Таблиця 1.2 деталізує зазначені дані, систематизуючи їх за типами, характеристиками та джерелами формування.

Таблиця 1.2 – Основні типи поведінкових даних користувачів в ІТ-проектах електронної комерції

Тип даних	Приклади	Характеристика
Навігаційні	Кліки, перегляди сторінок, переходи	Відображають маршрут користувача в системі
Пошукові	Введені запити, фільтри	Показують наміри користувача
Транзакційні	Покупки, оплати	Відображають завершені дії
Часові	Тривалість сесії, час перегляду	Характеризують рівень зацікавленості
Поведінка з рекомендаціями	Кліки на рекомендації, ігнорування	Дають зворотний зв'язок про ефективність системи

Як видно з таблиці 1.2, різні типи поведінкових даних дозволяють отримати різні аспекти інформації про користувача. Наприклад, пошукові дані відображають його наміри, тоді як транзакційні – фактичні рішення. У сукупності ці дані

формують більш повну картину поведінки, що є критично важливим для побудови якісної персоналізації.

Процес формування та використання поведінкових даних у рекомендаційній системі починається з взаємодії користувача з системою, під час якої генеруються поведінкові дані. Далі ці дані передаються до бази даних, де накопичуються та зберігаються. Після цього вони обробляються аналітичними модулями або безпосередньо рекомендаційною системою, яка використовує їх для формування рекомендацій. Результати повертаються користувачу, що замикає цикл і забезпечує його повторюваність.

Водночас, використання поведінкових даних у IT-проектах пов'язане з низкою проблем. Однією з основних є проблема шуму в даних, так як не всі дії користувача відображають його реальні інтереси. Наприклад, випадкові кліки або короткі перегляди можуть спотворювати результати аналізу. Крім того, часто виникає проблема неповноти даних, особливо для нових користувачів, тобто так звана проблема «холодного старту» [14].

Ще одним важливим аспектом є обсяг даних. У великих IT-проектах ЕК кількість поведінкових даних може бути дуже значною, що потребує використання спеціалізованих технологій для їх зберігання та обробки. Це, у свою чергу, впливає на архітектуру системи та складність її реалізації.

Також слід враховувати питання конфіденційності та захисту персональних даних. Збір та обробка поведінкових даних користувачів у IT-проектах електронної комерції регулюється законодавством у сфері захисту персональних даних. Таким чином, дані процеси регламентуються законом «Про захист персональних даних» [15], який визначає вимоги до збору, зберігання та обробки інформації про користувачів.

Відповідно до зазначеного закону, персональні дані можуть оброблятися лише за наявності правових підстав, зокрема згоди користувача, а також за умови забезпечення їх належного захисту. У контексті РС це означає, що збір

поведінкових даних повинен бути прозорим для користувача, а самі дані – анонімізованими або псевдонімізованими у випадках, коли це можливо.

Як зазначається в дослідженнях [16, 17], такі обмеження безпосередньо впливають на архітектуру ІТ-систем, оскільки виникає необхідність реалізації додаткових механізмів, зокрема управління згодою користувачів, обмеження доступу до даних, а також забезпечення їх безпечного зберігання та передачі, що, у свою чергу, ускладнює реалізацію РС і може впливати на повноту доступних даних для аналізу.

Таким чином, регуляторні вимоги є важливим фактором, який необхідно враховувати при проєктуванні систем персоналізації, оскільки вони визначають не лише можливості використання поведінкових даних, але й накладають обмеження на способи їх обробки.

Поведінкові дані користувачів є основою для реалізації персоналізації в ІТ-проєктах ЕК. Вони дозволяють адаптувати систему під індивідуальні потреби користувача, однак потребують якісної обробки, аналізу та врахування низки обмежень. Ефективне використання таких даних є одним із ключових факторів успішності РС.

1.4 Огляд сучасних методів побудови рекомендаційних систем

Сучасні РС є складними програмно-аналітичними рішеннями, які використовують різні підходи до обробки даних та формування персоналізованих рекомендацій. У науковій літературі представлено значну кількість методів, які відрізняються як за принципами роботи, так і за областю застосування. Відповідно до проведеного аналізу наукових джерел [18–21], сучасні методи рекомендаційних систем доцільно класифікувати за основними підходами, що узагальнено представлено на рисунку 1.3.



Рисунок 1.3 – Класифікація сучасних методів побудови рекомендаційних систем

Як показано на рисунку 1.3, сучасні методи побудови РС можна умовно поділити на класичні, тобто колаборативні, контент-орієнтовані, знання-орієнтовані, та сучасні, тобто гібридні, машинного та глибокого навчання. Такий підхід дозволяє систематизувати існуючі рішення та визначити напрямки їх подальшого розвитку.

Найбільш поширеними є методи колаборативної фільтрації, контент-орієнтовані методи, знання-орієнтовані підходи та гібридні моделі, що поєднують декілька методів одночасно. Для систематизації основних підходів доцільно представити їх у вигляді узагальненої класифікації, що наведена в таблиці 1.3.

Таблиця 1.3 – Основні методи побудови рекомендаційних систем

Метод	Суть підходу	Переваги	Недоліки
Колаборативна фільтрація	Використання даних про поведінку інших користувачів	Не потребує опису товарів	Проблема холодного старту

Кінець таблиці 1.3

Метод	Суть підходу	Переваги	Недоліки
Контент-орієнтований	Аналіз характеристик об'єктів і вподобань користувача	Персоналізація без інших користувачів	Обмеженість рекомендацій
Знання-орієнтований	Використання правил, знань та обмежень	Висока точність у складних системах	Складність побудови
Гібридний	Комбінація кількох методів	Вища точність	Складність реалізації

Як видно з таблиці 1.3, кожен із підходів має як свої переваги, так і обмеження, що визначає доцільність його використання в конкретному ІТ-проекті. Одним із найбільш досліджених підходів є колаборативна фільтрація, яка базується на припущенні, що користувачі з подібною поведінкою мають схожі вподобання. У роботах [18, 19] зазначається, що цей метод ефективно працює у випадках, коли є достатня кількість даних про взаємодію користувачів із системою. Зокрема, застосовуються два основних підходи: user-based та item-based фільтрації, тобто фільтрація на основі користувача та на основі товару.

Водночас у літературі [20] підкреслюється, що колаборативна фільтрація має суттєвий недолік, тобто проблему «холодного старту», яка виникає для нових користувачів або нових товарів, для яких відсутні історичні дані.

Контент-орієнтовані методи, на відміну від колаборативних, базуються на аналізі характеристик об'єктів і профілю користувача. У роботах [19, 21] зазначається, що такі методи дозволяють формувати рекомендації навіть за відсутності інформації про інших користувачів. Вони широко застосовуються в системах, де доступна структурована інформація про товари, наприклад, категорії, атрибути, опис.

Однак, як відзначається в дослідженнях [20, 21], контент-орієнтовані методи мають обмеження, пов'язані з тим, що рекомендації часто залишаються в межах вже відомих інтересів користувача, що знижує різноманітність результатів.

Окрему групу становлять знання-орієнтовані методи, які використовують експертні знання, правила або логічні моделі для формування рекомендацій. У літературі [18, 22] такі підходи розглядаються як ефективні для складних предметних областей, де рішення залежить не лише від поведінки користувача, але й від багатьох додаткових факторів. Наприклад, у сфері страхування або фінансових послуг такі системи дозволяють враховувати обмеження, вимоги та бізнес-правила. Разом із тим, як зазначається в [22], основним недоліком знання-орієнтованих систем є складність їх розробки та підтримки, оскільки вони потребують формалізації експертних знань.

У сучасних ІТ-проектах все більшого поширення набувають гібридні підходи, які поєднують кілька методів одночасно. Як показано в роботах [18–22], використання гібридних моделей дозволяє компенсувати недоліки окремих методів і підвищити якість рекомендацій. Наприклад, поєднання колаборативної та контентної фільтрації дозволяє зменшити вплив проблеми холодного старту.

Окрім класичних підходів, у сучасних дослідженнях особливу увагу приділяють методам, заснованим на машинному та глибокому навчанні, які фактично стали основою більшості промислових РС. Як зазначається у роботах [16, 18], використання моделей машинного навчання дозволяє враховувати складні нелінійні залежності між користувачами та об'єктами, що значно підвищує точність рекомендацій.

У сучасних ІТ-проектах ЕК широко застосовуються такі підходи, як матрична факторизація, градієнтний бустинг, а також нейронні мережі, зокрема глибокі нейронні мережі та рекурентні архітектури. У літературі [19, 21] підкреслюється, що такі методи дозволяють ефективно працювати з великими обсягами поведінкових даних і враховувати динаміку змін інтересів користувачів.

Зокрема, у промислових системах ЕК, таких як Netflix та Amazon, рекомендаційні алгоритми базуються саме на поєднанні методів машинного навчання з класичними підходами, що дозволяє досягти високого рівня

персоналізації. Як зазначається в [16–19], такі системи використовують складні ансамблеві моделі, які поєднують колаборативну фільтрацію, контентний аналіз та нейронні мережі.

Водночас, незважаючи на їх ефективність, методи машинного та глибокого навчання мають ряд обмежень. Вони потребують значних обчислювальних ресурсів, великої кількості даних для навчання, а також є складними з точки зору інтерпретації результатів, що ускладнює їх застосування в ІТ-проєктах із обмеженими ресурсами або на ранніх етапах життєвого циклу системи.

Разом із тим, класичні методи не втрачають своєї актуальності, оскільки часто використовуються як базові компоненти або частина гібридних моделей. Як зазначається у дослідженнях [15–17], саме поєднання класичних і сучасних підходів дозволяє досягти оптимального балансу між точністю рекомендацій та складністю реалізації системи.

У практиці реалізації ІТ-проєктів ЕК вибір конкретного методу РС визначається не лише його теоретичними характеристиками, але й низкою прикладних факторів. Як зазначається в дослідженнях [16, 17], до ключових критеріїв належать обсяг і якість доступних даних, необхідність пояснюваності рекомендацій, обмеження бюджету проєкту, а також рівень технічної експертизи команди розробників.

Наприклад, складні моделі глибокого навчання можуть забезпечувати високу точність, але потребують значних ресурсів і не завжди є доцільними для проєктів із обмеженими можливостями. Тобто вибір методу РС в ІТ-проєкті є компромісом між точністю, складністю реалізації та практичними обмеженнями. Проведений огляд методів показує, що жоден із підходів не є універсальним. Кожен метод має свої переваги та обмеження, що обумовлює необхідність їх комбінування або вдосконалення, що, у свою чергу, створює передумови для подальших досліджень і розробки нових методів персоналізації в ІТ-проєктах електронної комерції.

1.5 Аналіз застосування методів персоналізації в ІТ-проектах

У попередньому підрозділі розглянуто основні методи побудови РС та їх класифікацію. Проте сам по собі вибір методу ще не гарантує ефективної роботи системи. У реальних ІТ-проектах важливу роль відіграє те, як саме ці методи впроваджуються, на яких етапах життєвого циклу використовуються та з якими обмеженнями стикаються. При цьому важливо зазначити, що методи персоналізації не є жорстко прив'язаними до конкретних етапів PMBOK 8 [4] або життєвого циклу ІТ-проекту. Один і той самий алгоритм може використовуватися на різних етапах функціонування системи. Однак ефективність і доцільність застосування певних підходів визначається насамперед обсягом, структурою та якістю доступних даних. Саме тому на різних етапах розвитку ІТ-проекту перевага може надаватися різним методам персоналізації.

У сучасних ІТ-проектах ЕК методи персоналізації застосовуються не ізольовано, а як частина комплексної системи, що включає збір даних, їх обробку, інтеграцію з бізнес-процесами та подальшу оптимізацію. Як показано на рисунку 1.3, існує кілька основних груп методів, кожна з яких має свою область застосування та специфіку реалізації. З практичної точки зору, застосування методів персоналізації можна розглянути через призму життєвого циклу ІТ-проекту.

Щодо застосування на етапі ініціації, то на цьому етапі ІТ-проекту визначається, чи взагалі доцільно використовувати персоналізацію, а також які саме підходи будуть застосовуватися. У цей період формується розуміння, які дані доступні, які бізнес-цілі стоять перед системою, а також які метрики ефективності будуть використовуватися [23]. Оскільки на етапі ініціації система ще не накопичила достатнього обсягу поведінкових даних, застосування методів, що критично залежать від історії взаємодій користувачів, зокрема колаборативної

фільтрації, є обмеженим. У таких умовах більш доцільними є контент-орієнтовані або знання-орієнтовані підходи, які можуть працювати навіть за мінімального обсягу історичних даних.

На етапі планування визначаються архітектура рекомендаційної системи, способи збору та зберігання поведінкових даних, а також вимоги до масштабування й інтеграції алгоритмів персоналізації. У цей період гібридні підходи розглядаються переважно як архітектурна концепція, яка дозволяє поєднати декілька методів у межах єдиної системи після накопичення достатнього обсягу даних.

На етапі виконання відбувається безпосередня реалізація обраних методів. Тут важливо враховувати структуру даних, швидкодію системи, а також можливість масштабування. У більшості сучасних ІТ-проектів ЕК використовуються гібридні підходи, які поєднують кілька методів одночасно. Наприклад, контент-орієнтовані методи можуть застосовуватися разом із колаборативною фільтрацією для підвищення точності рекомендацій.

На етапі моніторингу та контролю здійснюється безперервний моніторинг роботи РС, вимірювання її ефективності відповідно до визначених метрик, аналіз поведінкових даних та внесення коректив у алгоритми персоналізації. Саме тут проявляються реальні переваги та недоліки обраних методів. У процесі експлуатації системи та її моніторингу накопичуються значні обсяги поведінкових даних користувачів. Саме це створює передумови для ефективного використання методів машинного та глибокого навчання, які потребують великої кількості даних для навчання моделей. Тобто доцільність використання складних підходів машинного навчання визначається не етапом РМВОК 8 [4] як таким, а рівнем зрілості системи та наявністю достатньої кількості даних для аналізу [24].

На етапі завершення проводиться остаточна оцінка результатів функціонування РС, формалізація отриманих знань і напрацювань, підготовка документації для подальшої підтримки або передачі продукту. Накопичені

поведінкові дані та метрики ефективності дозволяють проаналізувати довгострокову ефективність персоналізації, визначити фактори відтоку клієнтів та сформулювати рекомендації щодо розвитку системи в рамках наступних проєктів або ітерацій. Для більш наочного представлення доцільно узагальнити застосування різних методів персоналізації в ІТ-проєктах, що наведено в таблиці 1.4.

Таблиця 1.4 – Застосування методів персоналізації на різних етапах ІТ-проєкту

Етап ІТ-проєкту	Найбільш доцільні методи за доступних даних	Особливості застосування
Ініціація	Контент-орієнтовані, знання-орієнтовані	Відсутність даних, формування вимог
Планування	Гібридні	Вибір архітектури системи
Виконання	Колаборативні, гібридні	Інтеграція алгоритмів
Моніторинг і контроль	Машинне навчання, глибоке навчання	Робота з великими даними, коригування моделей
Завершення	Аналітичні, статистичні	Оцінка результатів, формалізація знань

Як видно з таблиці 1.4, доцільність використання певних методів персоналізації визначається не лише їх алгоритмічними властивостями, але й доступністю та характеристиками даних на різних етапах розвитку ІТ-проєкту. При цьому самі алгоритми не є жорстко прив'язаними до стадій PMBOK 8 [4], а можуть використовуватися протягом усього життєвого циклу системи. Різниця полягає насамперед у тому, що на різних етапах проєкту система має різний рівень зрілості та різний обсяг накопичених поведінкових даних, що впливає на ефективність окремих підходів.

Такий підхід відповідає практиці побудови сучасних рекомендаційних систем і забезпечує гнучкість їх подальшого розвитку. Як зазначається в роботах [23, 24], вибір гібридної архітектури на ранніх етапах проєкту дозволяє закласти

гнучкість системи та забезпечити можливість подальшого масштабування і вдосконалення рекомендаційних механізмів.

Розглядаючи застосування методів персоналізації з позиції стандарту РМВОК 8 [4], можна зазначити, що вони безпосередньо пов'язані з низкою процесів управління проєктом. Зокрема, формування вимог до РС відповідає процесам управління змістом проєкту, а визначення джерел і характеристик даних – процесам управління ресурсами та комунікаціями.

Зокрема, вимоги до якості поведінкових даних безпосередньо впливають на процеси управління змістом та ризиками проєкту. Наприклад, якщо на етапі планування не враховано необхідність збору повних і достовірних даних, це може призвести до зниження ефективності РС та необхідності внесення змін у функціональні вимоги на пізніших етапах. У контексті управління ризиками це означає появу додаткових ризиків, пов'язаних із якістю даних, які необхідно ідентифікувати, оцінювати та мінімізувати ще на ранніх стадіях проєкту.

Крім того, питання якості даних впливають на процеси моніторингу та контролю, оскільки ефективність рекомендаційних алгоритмів повинна постійно оцінюватися за відповідними метриками. У разі виявлення відхилень необхідно коригувати як моделі, так і процеси збору даних, що потребує координації між технічними та управлінськими командами. Реалізація рекомендаційних алгоритмів і їх інтеграція в систему пов'язані з процесами реалізації проєкту, тоді як оцінка ефективності персоналізації відповідає процесам моніторингу та контролю. У цьому контексті РС виступає не лише як програмний модуль, а як частина проєктного продукту, якість якого впливає на досягнення цілей проєкту.

Разом із тим, у процесі реалізації РС у ІТ-проєктах виникає низка практичних проблем, які безпосередньо впливають на ефективність їх застосування. Однією з ключових є проблема інтеграції методів персоналізації в існуючу архітектуру системи. Як зазначається в дослідженнях [24, 25], сучасні ІТ-проєкти ЕК мають складну багаторівневу структуру, і додавання

рекомендаційного модуля потребує узгодження з уже існуючими компонентами, що може вимагати значних змін у системі та додаткових ресурсів.

Не менш важливим фактором є якість поведінкових даних, які використовуються для побудови рекомендацій. У роботах [23, 25] підкреслюється, що наявність шуму, пропусків або некоректних даних може суттєво знижувати точність рекомендацій, навіть за використання сучасних алгоритмів. Це особливо актуально для ІТ-проектів, у яких дані формуються з різних джерел і мають неоднорідну структуру.

Окремо слід виділити проблему масштабування РС. У міру зростання кількості користувачів і обсягу даних виникає необхідність забезпечення швидкої обробки інформації та формування рекомендацій у реальному часі. Як зазначається в [24, 25], це потребує використання більш складних алгоритмів, розподілених обчислень та оптимізації архітектури системи, що ускладнює реалізацію таких рішень.

Отже, ефективність методів персоналізації визначається не лише вибором алгоритмів, але й якістю даних, особливостями архітектури системи та здатністю до масштабування, що необхідно враховувати при реалізації РС у ІТ-проектах ЕК.

1.6 Виявлення недоліків існуючих методів і підходів

Проведений у попередніх підрозділах аналіз сучасних методів побудови РС та їх застосування в ІТ-проектах дозволяє зробити висновок про те, що, незважаючи на значний розвиток даної галузі, існуючі підходи мають низку суттєвих обмежень, які ускладнюють їх ефективне використання в реальних умовах ЕК. Як показано на рисунку 1.3, сучасні методи РС представлені широким спектром підходів, однак жоден із них не є універсальним і без недоліків.

У науковій літературі [25–28] неодноразово підкреслюється, що одним із найбільш поширених підходів є колаборативна фільтрація, проте саме вона має ряд принципових обмежень. Зокрема, у роботах [25, 26] зазначається, що ефективність цього методу напряму залежить від обсягу доступних даних про користувачів, що створює проблему так званого «холодного старту». У ситуаціях, коли новий користувач або новий товар ще не має історії взаємодії, система не здатна сформулювати релевантні рекомендації, що знижує якість персоналізації на ранніх етапах використання системи.

Окрім цього, як відзначається в [27], колаборативні методи є чутливими до розрідженості даних, яка характерна для більшості ІТ-проектів електронної комерції. Велика кількість товарів і відносно невелика кількість взаємодій призводять до того, що матриці користувач–товар є неповними, що ускладнює пошук подібностей між користувачами або об'єктами. У літературі [25–27] також підкреслюється, що такі методи можуть бути вразливими до маніпуляцій, наприклад, штучного створення оцінок або дій, що впливає на достовірність рекомендацій.

Контент-орієнтовані підходи, які розглядаються у працях [26, 28], мають іншу природу обмежень. Вони дозволяють формувати рекомендації без необхідності використання даних інших користувачів, однак, як зазначається в [28], такі системи часто обмежуються вже відомими вподобаннями користувача, що призводить до ефекту «інформаційного замикання», коли користувач отримує лише схожі рекомендації і не відкриває для себе нові варіанти. У результаті знижується різноманітність рекомендацій, що може негативно впливати на зацікавленість користувача.

У роботах [25, 28] також звертається увага на те, що контент-орієнтовані методи значною мірою залежать від якості опису об'єктів. Якщо інформація про товари є неповною або неструктурованою, ефективність таких методів суттєво

знижується, що особливо актуально для ІТ-проектів, де дані можуть надходити з різних джерел і мати різний рівень деталізації.

Знання-орієнтовані підходи, як показано в дослідженнях [27–29], дозволяють враховувати складні залежності та бізнес-правила, проте їх основним недоліком є складність реалізації. Побудова таких систем потребує значних витрат часу та ресурсів, оскільки необхідно формалізувати експертні знання та підтримувати їх актуальність. У літературі [29] підкреслюється, що такі системи є менш гнучкими та складніше адаптуються до змін у поведінці користувачів.

Гібридні підходи, які, як зазначається у роботах [25–29], розглядаються як найбільш перспективні, також не позбавлені недоліків. Незважаючи на те, що вони дозволяють компенсувати слабкі сторони окремих методів, їх реалізація є значно складнішою з точки зору архітектури ІТ-проекту. Поєднання кількох алгоритмів вимагає узгодження різних типів даних, налаштування взаємодії між компонентами та додаткових обчислювальних ресурсів.

Сучасні методи, засновані на машинному навчанні та глибокому навчанні, які активно досліджуються в роботах [26–29], демонструють високу точність, проте мають свої специфічні обмеження. Як зазначається в [26, 27], такі методи потребують великих обсягів даних для навчання, що не завжди доступно на ранніх етапах ІТ-проекту. Крім того, вони характеризуються високою обчислювальною складністю та потребують значних ресурсів для навчання та підтримки моделей.

У літературі [28, 29] також піднімається питання інтерпретованості моделей машинного навчання. Багато сучасних алгоритмів працюють як «чорна скринька», що ускладнює пояснення отриманих рекомендацій як для розробників, так і для кінцевих користувачів. У контексті ІТ-проектів це може створювати додаткові ризики, пов'язані з довірою до системи та її прийняттям користувачами.

Більшість існуючих методів розглядаються у відриві від контексту управління ІТ-проектами. Як показує аналіз джерел [25–29], основна увага приділяється алгоритмічним аспектам, тоді як питання інтеграції РС у життєвий

цикл проєкту, управління даними, ризиками та якістю залишаються недостатньо дослідженими, що створює розрив між теоретичними розробками та їх практичним застосуванням.

Узагальнюючи проведений аналіз, можна виділити кілька системних проблем, характерних для більшості сучасних методів персоналізації. По-перше, значна частина підходів орієнтована на роботу зі статичними або слабо динамічними даними, що не відповідає умовам реальних ІТ-проєктів, де поведінка користувачів постійно змінюється. По-друге, існуючі методи недостатньо інтегровані в процеси управління ІТ-проєктами, зокрема у частині управління вимогами, якістю даних та ризиками, що ускладнює їх практичне впровадження. По-третє, багато сучасних підходів, особливо заснованих на машинному навчанні, характеризуються високою складністю та вартістю реалізації, що обмежує їх використання в умовах обмежених ресурсів. Тобто існує потреба у вдосконаленні підходів, які б поєднували методи персоналізації з процесами управління ІТ-проєктами.

1.7 Формування проблеми дослідження та обґрунтування напрямку її вирішення

Проведений аналіз сучасних методів побудови РС та особливостей їх застосування в ІТ-проєктах електронної комерції дозволяє зробити висновок про те, що, незважаючи на значну кількість існуючих підходів, проблема ефективної персоналізації на основі поведінкових даних залишається остаточно не вирішеною. Актуальні методи рекомендаційних систем охоплюють як класичні підходи, так і сучасні методи, засновані на машинному чи глибокому навчанні,

однак жоден із них не забезпечує універсального рішення для всіх умов використання.

Окремо слід зазначити, що навіть сучасні стандарти управління проєктами, зокрема PMBOK 8 [4], не приділяють достатньої уваги специфіці використання поведінкових даних та методів персоналізації в ІТ-проєктах. У стандарті детально описані процеси управління проєктом, однак питання інтеграції аналітичних моделей, зокрема РС, розглядаються лише на загальному рівні, що створює розрив між управлінськими підходами та технічними аспектами реалізації персоналізації, що потребує додаткових досліджень.

У наукових роботах [30–32] підкреслюється, що основною проблемою сучасних РС є розрив між теоретичними моделями та умовами їх практичного застосування в ІТ-проєктах. Зокрема, як зазначається в [30], більшість методів розробляються та тестуються в умовах контрольованих середовищ, де дані є достатньо повними та структурованими. У реальних же ІТ-проєктах ЕК дані часто є неповними, неоднорідними та динамічними, що значно ускладнює використання класичних підходів.

Однією з ключових проблем, на яку звертається увага в літературі [31, 32], є залежність якості рекомендацій від обсягу та якості поведінкових даних. Як розглянуто раніше, такі дані формуються в процесі взаємодії користувача із системою, проте вони можуть містити шум, бути неповними або не відображати реальні наміри користувача. Це призводить до того, що навіть сучасні алгоритми не завжди здатні забезпечити достатній рівень точності персоналізації.

Крім того, аналіз джерел [30–33] показує, що більшість існуючих методів недостатньо враховують контекст використання РС у межах ІТ-проєкту. Як правило, дослідження зосереджені на алгоритмах і математичних моделях, тоді як питання інтеграції цих рішень у архітектуру системи, їх взаємодії з іншими компонентами та впливу на бізнес-процеси залишаються поза увагою. У результаті

виникає ситуація, коли теоретично ефективні методи виявляються складними для впровадження або не дають очікуваного ефекту в реальних умовах.

Ще одним важливим аспектом є те, що сучасні системи рекомендацій часто орієнтовані на статичні моделі поведінки користувача. Як зазначається в [31], більшість підходів не враховує динамічний характер інтересів користувача, які можуть змінюватися залежно від часу, контексту або зовнішніх факторів. Це призводить до зниження релевантності рекомендацій і, відповідно, ефективності системи в цілому.

У працях [32, 33] також підкреслюється проблема складності сучасних методів, особливо тих, що базуються на машинному та глибокому навчанні. Хоча такі підходи демонструють високі показники точності, їх впровадження в ІТ-проектах пов'язане зі значними витратами ресурсів, необхідністю обробки великих обсягів даних та складністю підтримки моделей. Це обмежує їх використання, особливо в проектах із обмеженими ресурсами.

На основі проведеного аналізу можна сформулювати основну проблему дослідження, яка полягає в обмеженій здатності існуючих методів персоналізації адаптуватися до динамічної зміни інтересів користувачів в умовах реальних ІТ-проектів ЕК, де поведінкові дані є неповними, зашумленими та надходять у потоковому режимі. Як показано в дослідженнях [30–33], більшість існуючих підходів орієнтовані на обробку статичних або відносно стабільних наборів даних, що не відповідає умовам реальних систем, у яких інтереси користувачів можуть змінюватися в короткі проміжки часу. У таких умовах рекомендаційні моделі не встигають адаптуватися до нових даних, що призводить до зниження актуальності та релевантності рекомендацій.

Тобто ключовим аспектом дослідження є забезпечення адаптивності методів персоналізації до змін поведінки користувачів з урахуванням неповноти, зашумленості та динамічності даних, а також можливості їх ефективного використання в умовах реальних ІТ-проектів.

Обґрунтування напряму вирішення сформульованої проблеми полягає не лише у використанні комбінованих підходів, а у розробці адаптивних методів персоналізації, які здатні змінювати свою поведінку відповідно до динаміки даних. Перспективним напрямом є поєднання різних методів РС із можливістю їх динамічного налаштування залежно від зміни поведінкових патернів користувачів. У даному контексті доцільним є використання підходів, що передбачають онлайн-адаптацію моделей або ваг між різними методами, наприклад, колаборативними та контент-орієнтованими на основі актуальних даних. Як зазначається в дослідженнях [30–33], такі підходи дозволяють підвищити релевантність рекомендацій у ситуаціях, коли інтереси користувачів змінюються в короткі проміжки часу. Тобто напрямом дослідження є розробка або вдосконалення методів персоналізації, які забезпечують адаптивну обробку поведінкових даних у реальному часі з урахуванням їх динамічності, неповноти та зашумленості.

Крім того, важливим є врахування не лише технічних аспектів, але й організаційних факторів, пов'язаних із реалізацією рекомендаційних систем у ІТ-проєктах, що включає питання управління даними, оцінки ефективності, інтеграції з іншими компонентами системи та адаптації до змін у середовищі функціонування. Сформульована проблема та визначений напрям її вирішення створюють основу для подальших досліджень, спрямованих на розробку більш ефективних методів персоналізації рекомендацій, які враховують як особливості поведінкових даних, так і специфіку ІТ-проєктів ЕК.

1.8 Постановка задачі дослідження

Результати аналізу, проведеного в попередніх підрозділах, дають підстави стверджувати, що проблема персоналізації рекомендацій у ІТ-проєктах ЕК є

комплексною і не може бути вирішена шляхом простого застосування одного із відомих алгоритмів. Розрив між теоретичними розробками і реальними умовами функціонування систем, обмеження існуючих методів щодо якості та динаміки поведінкових даних, а також недостатня увага до організаційного контексту ІТ-проектів – усе це вказує на необхідність цілеспрямованого дослідження, яке б охопило задачу комплексно.

Об'єктом дослідження кваліфікаційної роботи є процеси персоналізації рекомендацій у ІТ-проектах електронної комерції на основі поведінкових даних користувачів.

Предметом дослідження є методи і підходи до побудови РС, що використовують поведінкові дані для формування персоналізованих рекомендацій з урахуванням специфіки ІТ-проектів.

Метою кваліфікаційної роботи є дослідження та вдосконалення методів персоналізації рекомендацій на основі поведінкових даних користувачів в ІТ-проектах ЕК, зокрема – розробка підходу, який дозволяє підвищити якість і релевантність рекомендацій за умов неповних, шумних та динамічних даних із урахуванням особливостей управління ІТ-проектами.

Для досягнення поставленої мети у кваліфікаційній роботі необхідно виконати такі етапи:

- провести аналіз предметної галузі ЕК та визначити ключові особливості формування та використання поведінкових даних користувачів в ІТ-проектах;
- розглянути роль і місце РС у структурі ІТ-проектів ЕК та дослідити підходи до їх інтеграції в архітектуру систем;
- виконати порівняльний аналіз існуючих методів побудови рекомендаційних систем, тобто колаборативної фільтрації, контент-орієнтованих підходів, гібридних методів і методів машинного та глибокого навчання, та виявити їх основні обмеження;

- визначити напрям вирішення виявленої проблеми та обґрунтувати вибір методологічної основи дослідження;
- розробити або адаптувати метод персоналізації рекомендацій, що враховує динаміку поведінкових даних і специфіку ІТ-проектного середовища;
- провести практичну апробацію розробленого рішення та оцінити його ефективність у порівнянні з існуючими підходами.

Аналіз поведінкових даних дозволить визначити, які типи взаємодій несуть найбільше інформаційне навантаження і можуть бути ефективно використані для побудови профілів користувачів, що, в свою чергу, сформує основу для подальшої роботи з алгоритмами рекомендацій.

Порівняльний аналіз методів надасть змогу системно виділити переваги і слабкі сторони кожного з підходів, що є необхідною передумовою для обґрунтованого вибору або синтезу методу, придатного до застосування в умовах реального ІТ-проекту.

Розробка та апробація методу забезпечить перевірку висунутих гіпотез на практиці та дозволить підтвердити або скоригувати теоретичні припущення, отримавши конкретні кількісні результати, які можна порівняти з показниками відомих рішень.

Отримані результати формують цілісне уявлення про предмет дослідження та дозволяють перейти до подальших етапів виконання кваліфікаційної роботи, пов'язаних із розробкою та дослідженням удосконалених методів персоналізації рекомендацій. Важливо, що запропонований підхід до постановки задачі враховує не лише алгоритмічні аспекти, але й особливості функціонування РС у реальних ІТ-проектах, що створює передумови для отримання практично значущих результатів.

Таким чином, сформульовані етапи для проведення дослідження визначають логічну основу для подальшого моделювання, розробки та експериментальної перевірки запропонованих рішень.

2 МЕТОД ПЕРСОНАЛІЗАЦІЇ РЕКОМЕНДАЦІЙ НА ОСНОВІ ПОВЕДІНКОВОГО ПРОФІЛЮ КОРИСТУВАЧА В ІТ-ПРОЄКТАХ ЕЛЕКТРОННОЇ КОМЕРЦІЇ

2.1 Формування поведінкового профілю користувача на основі явних та неявних даних

Персоналізація рекомендацій у сучасних системах ЕК безпосередньо залежить від якості, достовірності та повноти інформації про вподобання користувача. Центральним поняттям у даному контексті виступає поведінковий профіль, тобто структурований цифровий опис інтересів, патернів взаємодії та переваг конкретного користувача, який формується на основі його дій у системі та постійно оновлюється в міру надходження нових даних.

Детальний аналіз існуючих підходів до побудови РС свідчить про те, що більшість із них спирається або виключно на явні сигнали (Е), як, наприклад, рейтинги, оцінки та відгуки, або виключно на неявні (І), такі як кліки, перегляди або час взаємодії з об'єктом, не формуючи при цьому єдиного інтегрованого профілю, що суттєво обмежує точність персоналізації, особливо в умовах динамічного ІТ-проєкту з електронної комерції, де поведінка користувачів може змінюватися залежно від стадії розвитку проєкту, сезонності, маркетингових активностей та інших чинників [34].

Класично на сьогоднішній день прийнято виділяти три основні підходи до формування профілю користувача: на основі явного зворотного зв'язку, на основі неявної поведінки та гібридні методи.

Кожен із перелічених підходів має характерні переваги, але водночас і суттєві обмеження, загальний аналіз яких наведено в таблиці 2.1, разом із типами сигналів даних підходів.

Таблиця 2.1 – Порівняльний аналіз існуючих підходів до формування профілю користувача

Підхід	Тип сигналів	Метод агрегації	Основні недоліки
Явна колаборативна фільтрація CF	Лише явні	Матрична факторизація, k найближчих сусідів	Проблема холодного старту, ігнорування неявної поведінки
Неявна колаборативна фільтрація CF	Лише неявні (кліки, час)	Баєсівський персоналізований рейтинг, зважена регуляризована матрична факторизація	Шум у даних, відсутність контексту, невизначеність сигналу
Контентна фільтрація CBF	Атрибути товару + явні	Косинусна подібність	Надспеціалізація, не враховує соціальну поведінку
Гібридні системи	Явні + неявні	Зважене поєднання моделей	Статична вага сигналів, не враховує етап проєкту

Як видно з таблиці 2.1, жоден із наявних підходів не вирішує одночасно дві ключові проблеми, тобто інтеграцію явних і неявних сигналів в єдиний числовий профіль з адаптивним зважуванням, а також врахування контексту стадії ІТ-проєкту при оцінці релевантності поведінкових даних. Саме дані недоліки визначають необхідність розробки нового підходу.

У рамках запропонованого підходу поведінковий сигнал визначається як будь-яка зафіксована дія користувача u у системі ЕК щодо товару або контенту i , яка несе інформацію про ступінь його зацікавленості. Формально множина поведінкових сигналів S для пари (u, i) визначається за формулою:

$$S_{u,i} = \{(e_k, t_k, w_k) | k = 1, \dots, n\},$$

де e_k – тип події, наприклад, перегляд, клік або купівля;

t_k – часова мітка події;

w_k – вага сигналу, що відображає його інформативну цінність.

Запропоновано розрізняти два класи сигналів з подальшою деталізацією:

– явні сигнали, тобто дії, у яких користувач свідомо і безпосередньо висловлює своє ставлення до товару: оцінки за шкалою з відгуками, додавання до списку вибраного;

– неявні сигнали, тобто дії, що опосередковано відображають інтерес: час перебування на сторінці товару, кількість переглядів, завершення покупки, повторні звернення [35].

Для кожного типу сигналу визначається базове числове значення ваги w_k , що відображає його інформативну цінність. Повна класифікація даних сигналів наведена в таблиці 2.2.

Таблиця 2.2 – Класифікація поведінкових сигналів та їх базові ваги

Тип сигналу	Значення типу події e_k	Числове значення w_k (базове)
Явний позитивний	Оцінка 4–5 зірок, додавання до вибраного	1,0
Явний нейтральний	Оцінка 3 зірки	0,0
Явний негативний	Оцінка 1–2 зірки	–1,0
Неявний сильний	Завершена покупка, повторний перегляд, додання до кошика	0,8
Неявний середній	Перегляд картки товару ≥ 30 с	0,5

Кінець таблиці 2.2

Тип сигналу	Значення типу події e_k	Числове значення w_k (базове)
Неявний слабкий	Перегляд картки товару <30 с	0,2
Відсутній сигнал	Показ товару без переходу до картки	0,0

Значення базових ваг w_k , наведені в таблиці 2.2, є відправною точкою. Надалі буде показано, що дані ваги коригуються залежно від стадії ІТ-проєкту, що є ключовою відмінністю запропонованого методу від існуючих гібридних підходів.

Поведінковий профіль користувача u визначається як вектор числових оцінок інтересу до кожного товару i з каталогу системи ЕК. Для формування цього вектора пропонується агрегована функція інтересу $I(u, i)$, що об'єднує явні та неявні сигнали за формулою:

$$I(u, i) = \alpha_E \cdot f_E(u, i) + \alpha_I \cdot f_I(u, i), \quad (2.1)$$

де $f_E(u, i)$ – нормалізована оцінка явного інтересу користувача u до товару i ;

$f_I(u, i)$ – нормалізована оцінка неявного інтересу відповідно;

α_E та α_I – адаптивні коефіцієнти зважування явних і неявних сигналів відповідно, причому $\alpha_E + \alpha_I = 0,5 + 0,5 = 1$.

Нормалізована оцінка явного інтересу $f_E(u, i)$ обчислюється на основі рейтингів, виставлених користувачем, з урахуванням індивідуального зсуву користувача b_u та загального середнього рейтингу μ , за формулою:

$$f_E(u, i) = \frac{r_{u,i} - \mu - b_u}{r_{\max} - r_{\min}}, \quad (2.2)$$

де $r_{u,i}$ – фактичний рейтинг користувача u для товару i ;

μ – загальне середнє значення рейтингів у системі;

b_u – індивідуальний зсув користувача u , що характеризує його загальну суворість або лояльність в оцінюванні;

r_{max} і r_{min} – максимальне та мінімальне можливі значення рейтингу.

Індивідуальний зсув користувача b_u визначається як середнє відхилення його оцінок від середнього рейтингу оцінених ним товарів за формулою:

$$b_u = \frac{1}{|S_u|} \cdot \sum_{i \in S_u} (r_{u,i} - \bar{r}_i), \quad (2.3)$$

де S_u – множина товарів, оцінених користувачем u ;

$\bar{r}_i$ – середній рейтинг товару i серед усіх користувачів.

Нормалізована оцінка неявного інтересу $f_{I,u,i}$ формується як зважена сума нормалізованих неявних сигналів за формулою:

$$f_I(u, i) = \frac{\sum_k (w_k \cdot \varphi_k(u, i))}{\sum_k w_k}, \quad (2.4)$$

де w_k – вага k -го типу неявного сигналу відповідно до таблиці 2.2;

$\varphi_k(u, i) \in [0, 1]$ – зважене значення k -го неявного сигналу для пари (u, i) , яке враховує частоту взаємодій та часову актуальність подій.

Часова актуальність неявного сигналу враховується за допомогою експоненціальної функції забування за формулою:

$$\varphi_k(u, i) = \frac{n_k(u, i)}{\max_{u, i} n_k(u, i)} \cdot e^{-\lambda \cdot \Delta t}, \quad (2.5)$$

де $n_k(u, i)$ – кількість зафіксованих подій типу e_k для пари (u, i) ;

Δt – час у днях від останньої події типу e_k до поточного моменту;

λ – параметр швидкості забування з рекомендованим значенням $\lambda = 0,05$ за відсутності стадійної корекції.

На основі агрегованої функції інтересу $I(u, i)$ формується вектор поведінкового профілю користувача u за формулою:

$$P(u) = [I(u, i_1), I(u, i_2), \dots, I(u, i_m)], \quad (2.6)$$

де m – загальна кількість товарів у каталозі системи.

Вектор $P(u)$ є розрідженим, тобто більшість елементів дорівнюватимуть нулю або не будуть визначені через відсутність будь-якої взаємодії між користувачем u та відповідним товаром, що є типовою характеристикою РС і потребує спеціальних методів обробки при обчисленні схожості між профілями.

Для товарів, з якими у користувача немає жодної взаємодії, тобто $I(u, i)$ не визначено, значення функції інтересу встановлюється рівним нулю. Таким чином, $P(u) \in \mathbb{R}^m$ та $I(u, i) \in [-1, 1]$ для всіх i .

На рисунку 2.1 зображено загальну схему процесу формування поведінкового профілю користувача, що відображає потік даних від первинних подій у системі до фінального вектора $P(u)$.

Проведений аналіз дозволяє сформулювати три ключові недоліки існуючих методів формування профілю користувача, усунення яких буде проведено надалі.

По-перше, статичність коефіцієнтів α_E та α_I . В існуючих гібридних системах співвідношення явних і неявних сигналів є фіксованим і не враховує контекст [36]. В умовах ІТ-проєкту з ЕК на ранніх стадіях, як ініціація та планування, домінують неявні сигнали через брак рейтингів, тоді як на стадії виконання, моніторингу з контролем і завершення явні сигнали стають більш надійними. Запропонована модель передбачає адаптацію α_E та α_I залежно від стадії проєкту, що буде деталізовано надалі.

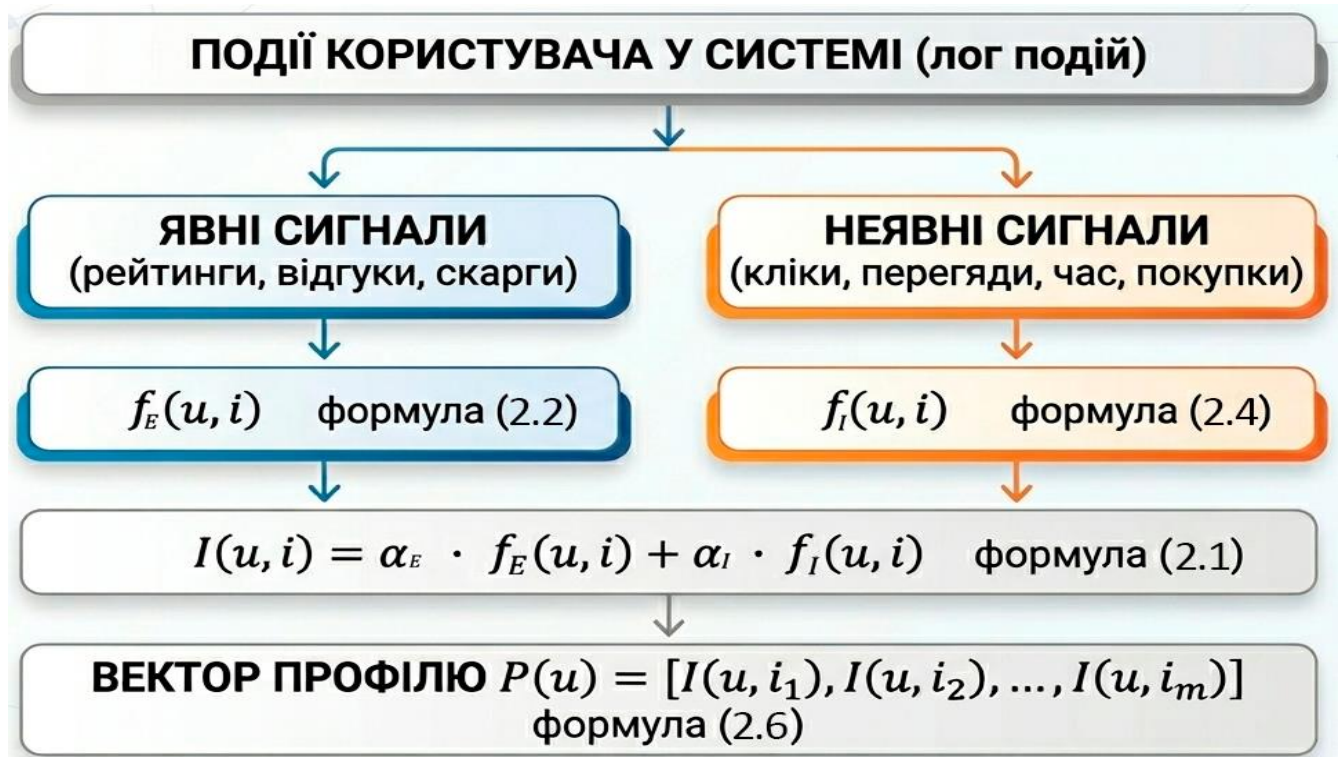


Рисунок 2.1 – Схема формування поведінкового профілю користувача

По-друге, відсутність часової актуалізації неявних сигналів у більшості практичних реалізацій. Хоча функція забування є відомою концепцією, параметр λ у формулі (2.5) зазвичай задається константою. У запропонованому методі λ також коригується залежно від стадії проекту.

По-третє, відсутність врахування індивідуального зсуву b_u при обробці явних сигналів. Ігнорування різниці між «суворим» і «лояльним» користувачем призводить до систематичних помилок у прогнозуванні. Формула (2.3) вирішує дану проблему за рахунок нормалізації оцінок відносно індивідуального стилю оцінювання кожного користувача.

Отже, запропоновано підхід до формування поведінкового профілю користувача $P(u)$ як числового вектора агрегованих оцінок інтересу $I(u, i)$, що об'єднує явні та неявні сигнали взаємодії. Введено класифікацію поведінкових сигналів з базовими вагами w_k , визначено формули нормалізації явного інтересу

$f_E(u, i)$ з урахуванням індивідуального зсуву b_u та неявного інтересу $f_I(u, i)$ з часовим забуванням. Виявлено три ключові недоліки існуючих підходів: статичність коефіцієнтів зважування, відсутність стадійного контексту та ігнорування індивідуального зсуву оцінювання.

2.2 Метод зважування поведінкових сигналів з урахуванням стадії ІТ-проєкту

Як вже розглянуто, агрегована функція інтересу $I(u, i)$ залежить від двох коефіцієнтів зважування α_E та α_I , а функція часового забування – від параметра λ . В існуючих підходах дані параметри є статичними константами, що не враховують контекст, у якому функціонує система рекомендацій [37]. Проте в умовах ІТ-проєкту з ЕК характер та достовірність поведінкових даних суттєво відрізняються залежно від поточної стадії проєкту.

На таких ранніх стадіях проєкту, як ініціації та планування, явних оцінок від користувачів практично немає, оскільки продукт або ще не виведений на ринок, або щойно запущено у тестовому режимі. Натомість на стадіях виконання, моніторингу з контролем і завершення накопичується достатньо явних сигналів для надійної персоналізації. Ігнорування даної відмінності призводить до того, що система або недооцінює наявні неявні сигнали на старті, або надає надмірну вагу застарілим неявним сигналам на зрілих стадіях.

Саме тому метою постає розробка методу динамічного зважування поведінкових сигналів, у якому коефіцієнти α_E , α_I та λ є функціями стадії ІТ-проєкту, а не постійними величинами.

В основу методу покладено стандартну модель життєвого циклу проєкту відповідно до РМВОК 8 [4], адаптовану до специфіки ІТ-проєктів ЕК. Визначено

п'ять стадій проекту, для кожної з яких характерний відмінний стан поведінкових даних та вимоги до рекомендаційної системи, що наведено у таблиці 2.3.

Таблиця 2.3 – Характеристика стадій ІТ-проекту електронної комерції з позиції рекомендаційної системи

Стадія ІТ-проекту	Середня тривалість	Наявність поведінкових даних	Роль РС
Ініціація	1–4 тижні	Власні дані РС відсутні, є неявні сигнали платформи ЕК	Базові рекомендації за популярністю
Планування	2–8 тижнів	Мінімальні неявні сигнали (А/В-тести, прототип)	Контентна фільтрація та слабкі неявні сигнали
Виконання	2–12 місяців	Зростаючі неявні та поодинокі явні	Активне накопичення профілю, гібрид
Моніторинг і контроль	1–6 місяців	Стабільні неявні та явні сигнали	Повноцінна персоналізація, CF і CBF
Завершення	2–6 тижнів	Накопичений зрілий профіль	Оптимізація, ретаргетинг, аналіз відтоку

Як бачимо з таблиці 2.3, на стадії ініціації, коли власні дані рекомендаційної системи ще відсутні, але вже наявний накопичений поведінковий архів платформи ЕК, тобто перегляди, кліки, покупки тощо, який слугує стартовим масивом неявних сигналів для формування початкових рекомендацій, які можна отримати під час реєстрації або через веб-аналітику. На цьому етапі РС може надавати лише базові рекомендації, наприклад, найпопулярніші товари. На стадії планування з'являються мінімальні неявні сигнали, які отримують завдяки А/В-тестуванням, тобто порівнянням реакції користувачів на різні версії інтерфейсу чи рекомендацій, та роботі з прототипом, наприклад, кліки, перегляди, час на сторінці. Дані сигнали дозволяють застосовувати контентну фільтрацію. Під час виконання, коли обсяг неявних та поодиноких явних даних зростає, РС активно

формує профіль користувача та переходить до гібридних алгоритмів. На етапі моніторингу РС забезпечує повноцінну персоналізацію, використовуючи спільну та контентну фільтрацію. На завершення накопичений зрілий профіль дозволяє оптимізувати рекомендації, проводити ретаргетинг та аналізувати відтік клієнтів. Таким чином, роль РС еволюціонує від базової популярності до глибокої персоналізації в міру зростання поведінкових даних.

Введемо формальне позначення: нехай $S = \{S_1, S_2, S_3, S_4, S_5\}$, тобто це впорядкована множина стадій проєкту, де S_1 відповідає ініціації, S_2 – плануванню, S_3 – виконанню, S_4 – моніторингу та контролю, S_5 – завершенню. Поточна стадія проєкту в момент t позначається як $\sigma(t) \in S$.

Ключове спостереження, що лежить в основі запропонованого методу: із зростанням індексу стадії зростає кількість та надійність явних сигналів, тоді як відносна інформативність неявних сигналів поступово знижується через накопичення більш достовірних явних даних. Таке спостереження формалізується через монотонні залежності коефіцієнтів зважування від стадії проєкту.

Запропонований метод полягає у тому, що коефіцієнти $\alpha_E(\sigma)$, $\alpha_I(\sigma)$ та параметр забування $\lambda(\sigma)$ є функціями поточної стадії проєкту $\sigma(t)$. Агрегована функція інтересу з минулого підрозділу набуває вигляду, відповідно до формули:

$$I(u, i, \sigma) = \alpha_E(\sigma) \cdot f_E(u, i) + \alpha_I(\sigma) \cdot f_I(u, i, \lambda(\sigma)), \quad (2.7)$$

де $\alpha_E(\sigma) + \alpha_I(\sigma) = 1$ для будь-якої стадії $\sigma \in S$, що забезпечує нормованість агрегованої оцінки інтересу.

Функція неявного інтересу $f_I(u, i, \lambda(\sigma))$ тепер явно залежить від параметра забування $\lambda(\sigma)$, який також є функцією стадії, відповідно до формули:

$$f_I(u, i, \lambda(\sigma)) = \frac{\sum_k (w_k(\sigma) \cdot \varphi_k(u, i, \lambda(\sigma)))}{\sum_k w_k(\sigma)},$$

де $\varphi_k(u, i, \lambda(\sigma)) = (n_k(u, i) / \max_{u,i} n_k(u, i)) \cdot e^{-\lambda(\sigma) \cdot \Delta t}$ – функція часового забування з параметром λ , визначеним для стадії σ ;

$w_k(\sigma)$ – стадійно-скоригована вага k -го типу сигналу.

Також зауважимо, що ваги окремих типів неявних сигналів $w_k(\sigma)$ можуть коригуватися залежно від стадії. Так, на ранніх стадіях проєкту цінність сигналу «завершена покупка» є відносно вищою через брак будь-яких інших даних, ніж на зрілих стадіях, де накопичені різноманітніші сигнали.

Значення коефіцієнтів $\alpha_E(\sigma)$, $\alpha_I(\sigma)$ та $\lambda(\sigma)$ для кожної з п'яти стадій визначено на основі аналізу типового розподілу поведінкових даних у ІТ-проєктах ЕК та принципу монотонності: зі зростанням стадії $\alpha_E(\sigma)$ монотонно зростає, $\alpha_I(\sigma)$ монотонно спадає, $\lambda(\sigma)$ монотонно зростає, тобто йде прискорення забування. Значення даних коефіцієнтів зважування відповідно до стадій ІТ-проєкту наведено у таблиці 2.4.

Таблиця 2.4 – Значення коефіцієнтів зважування та параметра забування за стадіями ІТ-проєкту

Стадія проєкту	α_E	α_I	λ	Обґрунтування
Ініціація	0,10	0,90	0,02	Явні дані РС відсутні; неявні – накопичений архів поведінкових подій платформи ЕК; повільне забування, кожен сигнал є цінним
Планування	0,262 5	0,737 5	0,04	З'являються перші неявні сигнали; явні – поодинокі; профіль лише формується
Виконання	0,425	0,575	0,06	Паритет сигналів; явні стають надійнішими; стандартна швидкість забування
Моніторинг і контроль	0,587 5	0,412 5	0,08	Явні сигнали домінують; профіль зрілий; швидше забування застарілих неявних
Завершення	0,75	0,25	0,10	Максимальна довіра до явних оцінок; неявні – лише останні дії; активне забування

З таблиці 2.4 бачимо, що ключова роль відводиться зміні пріоритетів між неявними та явними сигналами в міру розвитку продукту. Забування у даному контексті є математичним вираженням темпу актуалізації інформації, тобто воно показує, як швидко старі дані втрачають свою значущість для поточного аналізу. На стадії ініціації можна побачити найнижчий коефіцієнт забування, що пояснюється критичною цінністю кожної крихти наявних знань. Оскільки на старті явних даних майже немає, а існують лише слабкі сигнали від прототипів, кожна отримана інформація стає фундаментом для подальших рішень, тому система тримається за неї якомога довше, не дозволяючи їй швидко знецінюватися. У ситуації інформаційного дефіциту будь-яке знання є занадто дорогим, щоб його ігнорувати або швидко замінювати новим.

З переходом до планування та виконання ситуація поступово змінюється, так як з'являються перші надійні метрики, і система починає виходити на баланс між інтуїтивними припущеннями та реальними показниками. Коефіцієнт забування при цьому зростає, оскільки середовище стає більш динамічним, а нові дані починають точніше описувати реальність, ніж застарілі прогнози. На етапах моніторингу та завершення процес прискорюється ще сильніше, досягаючи максимального темпу забування, що логічно, адже на фініші проєкту отримується зрілий профіль даних, де явні сигнали домінують, а старі неявні припущення, які були важливі на початку, тепер лише створюють інформаційний шум. Активне забування тут виступає механізмом очищення моделі від минулого, дозволяючи зосередитися на останніх діях та актуальному стані системи, де довіра до явних оцінок є максимальною.

Монотонні залежності $\alpha_E(\sigma)$ та $\alpha_I(\sigma)$ від номера стадії можна апроксимувати лінійними функціями від індексу стадії $j \in \{1, 2, 3, 4, 5\}$ за формулами:

$$\alpha_E(j) = 0,10 + 0,1625 \cdot (j - 1), \quad (2.8)$$

$$\alpha_I(j) = 1 - \alpha_E(j) = 0,90 - 0,1625 \cdot (j - 1), \quad (2.9)$$

де j – порядковий номер стадії проєкту, враховуючи, що $j = 1$ для ініціації, а $j = 5$ для завершення.

Аналогічно параметр забування λ зростає лінійно від $0,02$ до $0,10$ зі збільшенням j .

Формули (2.8) та (2.9) є лінійними апроксимаціями, що спрощують практичне застосування методу. У разі потреби більш точного налаштування для конкретного проєкту значення α_E , α_I та λ можуть уточнюватися за результатами валідаційної вибірки, що буде розглянуто надалі.

Окрім зміни глобальних коефіцієнтів α_E та α_I , запропонований метод передбачає стадійну корекцію ваг $w_k(\sigma)$ для окремих типів неявних сигналів. На ранніх стадіях, коли даних обмаль, будь-який сигнал є відносно цінним, тому ваги «слабких» сигналів підвищуються відносно «сильних». На зрілих стадіях навпаки – слабкі сигнали знецінюються швидше, оскільки їх перекривають надійніші явні та сильні неявні дані.

Формально стадійно-скоригована вага k -го типу сигналу визначається за формулою:

$$w_k(\sigma) = w_k \cdot \gamma_k(\sigma),$$

де w_k – базова вага сигналу з таблиці 2.2;

$\gamma_k(\sigma) \in (0, 1]$ – стадійний коефіцієнт корекції, що відображає відносну цінність сигналу типу k на стадії σ .

Для явних сигналів, як, наприклад, оцінок, $\gamma_k = 1,00$ на всіх стадіях, оскільки їх базова вага не потребує корекції, як і у відсутніх. Значення $\gamma_k(\sigma)$ для усіх неявних типів подій наведено нижче в таблиці 2.5.

З таблиці 2.5 бачимо, що значення $\gamma_k(\sigma)$ зменшуються для слабших сигналів, як короткий перегляд, у міру просування проєкту до завершення, що відображає зниження їх відносної інформативності на тлі зрілого профілю з явними оцінками.

Таблиця 2.5 – Стадійні коефіцієнти корекції $\gamma_k(\sigma)$ для неявних сигналів

Тип події	Ініціація	Планування	Виконання	Моніторинг і контроль	Завершення
Завершена покупка	0,90	0,85	0,80	0,75	0,70
Повторний перегляд	0,90	0,85	0,80	0,75	0,70
Додання до кошика	0,90	0,85	0,80	0,75	0,70
Перегляд картки товару ≥ 30 с	0,60	0,55	0,50	0,45	0,35
Перегляд картки товару < 30 с	0,20	0,20	0,20	0,20	0,15

Запропонований метод реалізується у вигляді алгоритму, зображеного на рисунку 2.2, який виконується кожного разу, коли система рекомендацій обчислює $I(u, i, \sigma)$ для генерації рекомендацій.

Часова складність алгоритму становить $O(K \cdot N)$, де K є кількістю типів сигналів, а N є кількістю зафіксованих подій для пари (u, i) . Оскільки K є константою відповідно до таблиці 2.2, то складність фактично лінійна відносно кількості подій.

Визначення поточної стадії $\sigma(t)$ є зовнішньою операцією відносно РС і виконується менеджером або автоматично через інтеграцію з системою управління проектом.

АЛГОРИТМ ДИНАМІЧНОГО ЗВАЖУВАННЯ ПОВЕДІНКОВИХ СИГНАЛІВ

Вхід: користувач u , товар i , поточна стадія $\sigma(t)$

Вихід: агрегована оцінка інтересу $I(u, i, \sigma)$

Крок 1. Визначити $j = index(\sigma(t)) \in \{1, \dots, 5\}$

Крок 2. Обчислити $\alpha_E(j)$ та $\alpha_I(j)$ за формулами (2.8) і (2.9)

Крок 3. Обчислити $\lambda(j) = 0,02 + 0,02 \cdot (j - 1)$

Крок 4. Для кожного k : $w_k(\sigma) = w_k \cdot \gamma_k(\sigma)$ (таблиця 2.5)

Крок 5. Обчислити $f_E(u, i)$ за формулою (2.2)
Якщо $r(u, i)$ відсутній $\rightarrow f_E(u, i) = 0$

Крок 6. Для кожного k :

$$\varphi_k(u, i, \lambda(\sigma)) = (n_k(u, i) / \max_{u, i} n_k(u, i)) \cdot e^{-\lambda(\sigma) \cdot \Delta t}$$

$$f_I(u, i, \lambda(\sigma)) = \sum_k (w_k(\sigma) \cdot \varphi_k(u, i, \lambda(\sigma))) / \sum_k w_k(\sigma)$$

Крок 7. $I(u, i, \sigma) = \alpha_E(\sigma) \cdot f_E(u, i) + \alpha_I(\sigma) \cdot f_I(u, i, \lambda(\sigma))$

Повернути $I(u, i, \sigma)$

Рисунок 2.2 – Алгоритм динамічного зважування поведінкових сигналів

Запропонований метод зважування має такі формально обґрунтовані властивості:

– нормованість, тобто для будь-якої стадії σ виконується $\alpha_E(\sigma) + \alpha_I(\sigma) = 1$, що забезпечує $I(u, i, \sigma) \in [-1, 1]$ за умови нормованості f_E та f_I ;

– монотонність, тобто з ростом індексу стадії j коефіцієнт $\alpha_E(j)$ монотонно не спадає, а $\alpha_I(j)$ монотонно не зростає, таким чином $\alpha_E(j+1) \geq \alpha_E(j)$ та $\alpha_I(j+1) \leq \alpha_I(j)$ для всіх j ;

– стійкість до відсутності даних, тобто якщо явних сигналів немає ($f_E = 0$), то $I(u, i, \sigma) = \alpha_I(\sigma) \cdot f_I(u, i, \lambda(\sigma)) \geq 0$, тобто система продовжує генерувати рекомендації на основі неявних сигналів;

– сумісність, тобто при $\alpha_E = \text{const}$ та $\alpha_I = \text{const}$ метод вироджується у стандартний гібридний підхід, що забезпечує зворотну сумісність.

Властивість монотонності є принциповою для управлінської інтерпретації методу, так як вона гарантує, що із просуванням проєкту вперед система рекомендацій автоматично довіряє явним оцінкам дедалі більше, що відповідає природній логіці накопичення достовірних даних у комерційних проєктах.

Ключові відмінності запропонованого методу зважування від існуючих підходів, виявлених раніше, зводяться до трьох аспектів.

По-перше, на відміну від класичних гібридних систем, де α_E та α_I є фіксованими гіперпараметрами, що налаштовуються один раз на навчальній вибірці, у запропонованому методі дані коефіцієнти є детермінованими функціями стадії проєкту і не потребують повторного навчання при переході між стадіями., що суттєво спрощує практичне застосування в умовах реального ІТ-проєкту.

По-друге, параметр λ функції часового забування у запропонованому методі також є стадійно-залежним, що відображає різну швидкість старіння поведінкових даних на різних фазах проєкту. На ранніх стадіях дані цінні довше, тобто λ малий, а на пізніх вони застарівають швидше, тобто λ великий, що узгоджується з практикою управління продуктом.

По-третє, стадійна корекція ваг $w_k(\sigma)$ дозволяє динамічно перерозподіляти відносний внесок різних типів неявних сигналів, що є новим елементом порівняно з підходами, де ваги сигналів є постійними.

2.3 Модель персоналізації рекомендацій на основі поведінкового профілю

У минулих підрозділах розроблено два взаємопов'язані компоненти: метод формування поведінкового профілю користувача $P(u)$ на основі явних та неявних

сигналів і метод динамічного зважування цих сигналів залежно від поточної стадії ІТ-проєкту $\sigma(t)$. Тепер головною метою є інтеграція зазначених компонентів у цілісну модель персоналізації рекомендацій, що дозволяє генерувати ранжований список товарів для кожного користувача з урахуванням його поведінкового профілю та поточного контексту проєкту.

Модель охоплює повний ланцюжок від первинних поведінкових подій до фінального списку рекомендацій і є замкненою у тому сенсі, що нові взаємодії користувача автоматично оновлюють його профіль та впливають на наступну генерацію рекомендацій.

Зауважимо, що профіль $P(u, \sigma)$ тепер має явно залежати від стадії σ , що відрізняє запропоновану модель від стандартних підходів, де профіль є статичним вектором, що оновлюється лише при надходженні нових оцінок.

Центральним елементом колаборативної фільтрації є міра подібності між профілями двох користувачів. В існуючих підходах міра є незмінною і не враховує ні індивідуальний зсув оцінювання b_u , ні поточну стадію проєкту [38]. У запропонованій моделі вводиться стадійно-залежна міра подібності $Sim(u, v, \sigma)$, що усуває дані обмеження. Порівняння запропонованої міри з існуючими альтернативами наведено в таблиці 2.6.

Таблиця 2.6 – Порівняння мір подібності профілів у контексті ІТ-проєктів

Метрика	Враховує σ	Враховує b_u	Обмеження у контексті ІТ-проєктів
Косинусна подібність	Ні	Ні	Ігнорує індивідуальний зсув і не чутлива до стадії
Пірсон	Ні	Частково	Нестабільна при малій кількості спільних оцінок
Базова Манхеттенська	Ні	Ні	Статична та не враховує стадію проєкту
$Sim(u, v, \sigma)$	Так	Так	Усуває зазначені обмеження

Косинусна подібність з таблиці 2.6 обчислюється як кут між двома профільними векторами у багатовимірному просторі товарів і є однією з найпоширеніших метрик у рекомендаційних системах завдяки простоті обчислення. Проте вона має суттєвий недолік у контексті персоналізації: два користувачі, які систематично ставлять різні за шкалою оцінки одним і тим самим товарам, наприклад, якщо один завжди оцінює в діапазоні 4–5, а інший у діапазоні 1–2, то можуть отримати високе значення косинусної подібності через колінеарність їхніх векторів, хоча за уподобаннями вони є протилежними. Крім того, косинусна метрика не враховує поточну стадію проєкту і тому однаково зважує явні та неявні складові профілю незалежно від того, наскільки кожна з них є надійною на конкретній фазі.

Кореляція Пірсона усуває частину даного недоліку, оскільки центрує оцінки відносно середнього значення кожного користувача, що частково компенсує індивідуальний зсув оцінювання. Саме тому вона вважається більш підходящою для рейтингових даних порівняно з косинусом. Однак кореляція Пірсона є нестабільною при малій кількості спільних оцінок: якщо два користувачі мають лише 2–3 спільних товари, значення кореляції може бути екстремальним, тобто близьким до +1 або -1, суто випадково, що призводить до ненадійного визначення сусідів. Даний ефект є особливо критичним на ранніх стадіях ІТ-проєкту, коли накопичених даних ще мало. Крім того, як і косинусна метрика, кореляція Пірсона не є стадійно-залежною.

Нормована Манхеттенська відстань, яка доповнюється до одиниці для отримання міри подібності, обчислює середнє абсолютне відхилення між відповідними елементами двох профільних векторів і є стійкішою до викидів порівняно з евклідовою метрикою. Вона менш чутлива до екстремальних значень в окремих позиціях, що робить її придатнішою для розріджених профілів із неявними сигналами.

Разом із тим базова версія Манхеттенської метрики не враховує ні індивідуальний зсув оцінювання b_u , ні поточну стадію проекту, що є її принциповим обмеженням у запропонованій моделі. Саме тому можна розширити Манхеттенську відстань корекцією на різницю зсувів $|b_u - b_v|$ та коефіцієнтом довіри $Conf(u, v)$, що й перетворює її на стадійно-залежну міру $Sim(u, v, \sigma)$.

Для забезпечення коректності міри подібності індивідуальний зсув b_u нормується до діапазону $[-1; 1]$ шляхом ділення на максимально можливе відхилення, тобто на $\Delta_{max} = \frac{r_{max} - r_{min}}{2}$. Таким чином, $|\tilde{b}_u - \tilde{b}_v| \leq 2 = D$, що гарантує невід'ємність виразу надалі за умови $|I(u, i, \sigma) - I(v, i, \sigma)| \leq D$. Введене додаткове обмеження $\max(0; \dots)$ слугує надійним захистом від викидів у розріджених профілях.

Тобто запропонована міра подібності $Sim(u, v, \sigma)$ будується на основі нормованої Манхеттенської відстані між профільними векторами $P(u, \sigma)$ та $P(v, \sigma)$ з корекцією на різницю індивідуальних зсувів та довіру за кількістю спільних взаємодій за формулою:

$$Sim(u, v, \sigma) = \max \left(0; 1 - \frac{\sum_{i \in T_{uv}} |I(u, i, \sigma) - I(v, i, \sigma)| + F \cdot |\tilde{b}_u - \tilde{b}_v|}{(|T_{uv}| + F) \cdot D} \right) \cdot Conf(u, v), \quad (2.10)$$

де $T_{uv} = T_u \cap T_v$ – множина товарів, з якими взаємодіяли обидва користувачі, тобто користувач u та користувач v ;

F – коефіцієнт чутливості до різниці індивідуальних зсувів, де рекомендоване значення $F = 1$;

D – максимально можлива різниця між значеннями функції інтересу двох користувачів, тобто $D = 2$, оскільки $I(u, i, \sigma) \in [-1, 1]$;

$\tilde{b}_u$ та $\tilde{b}_v$ – нормалізовані індивідуальні зсуви користувачів u та v .

Коефіцієнт довіри $Conf(u, v)$ враховує кількість спільних взаємодій та зменшує вплив подібності, обчисленої на малій кількості спільних товарів. Для уникнення ділення на нуль та надмірного заниження довіри при малих вибірках застосовується згладжена формула:

$$Conf(u, v) = \begin{cases} 0, & |T_{uv}| = 0 \\ \max\left(0; 1 - \frac{1}{(|T_{uv}| + \delta)^\beta}\right), & |T_{uv}| > 0 \end{cases} \quad (2.11)$$

де δ – мала константа згладжування зі значенням 0,5, що запобігає різким стрибкам при переході від $|T_{uv}| = 0$ до $|T_{uv}| = 1$;

β – параметр чутливості до кількості спільних взаємодій, а рекомендоване значення становить 0,5.

При $|T_{uv}| \rightarrow \infty$ коефіцієнт $Conf(u, v) \rightarrow 1$, тобто не спотворює подібність при достатній кількості спільних взаємодій. При $|T_{uv}| = 0$ довіра дорівнює нулю, що виключає вплив «ненадійних» сусідів без спільної історії. При $|T_{uv}| = 1$ з урахуванням згладжування $Conf \approx 0,18$, що є помірним зниженням порівняно з одиницею.

На основі міри $Sim(u, v, \sigma)$ для кожного користувача u визначається множина K найближчих сусідів $N(u, \sigma)$, тобто K користувачів із найвищим значенням $Sim(u, v, \sigma)$ серед усіх $v \in U, v \neq u$ за формулою:

$$N(u, \sigma) = \arg \text{top} - K_{v \in U \setminus \{u\}} Sim(u, v, \sigma), \quad (2.12)$$

де K – гіперпараметр моделі, де на практиці рекомендується $K \in [10, 50]$ з підбором за результатами крос-валідації.

Множина $N(u, \sigma)$ є стадійно-залежною, тобто із зміною σ змінюються значення $I(u, i, \sigma)$ та $I(v, i, \sigma)$ через коефіцієнти $\alpha_E(\sigma)$ та $\alpha_I(\sigma)$, що може призводити до

зміни складу сусідів, що є навмисною властивістю моделі, яка забезпечує адаптацію РС до поточного контексту проєкту.

Для пари (u, i) , де користувач u не взаємодіяв з товаром i , тобто $I(u, i, \sigma)$ не визначено з власних сигналів, прогнозована оцінка інтересу обчислюється як зважене середнє значень функції інтересу сусідів з $N(u, \sigma)$ за формулою:

$$\hat{R}(u, i, \sigma) = \frac{\sum_{v \in N(u, \sigma)} \text{Sim}(u, v, \sigma) \cdot I(v, i, \sigma)}{\sum_{v \in N(u, \sigma)} |\text{Sim}(u, v, \sigma)|}, \quad (2.13)$$

де $I(v, i, \sigma)$ – значення функції інтересу сусіда v до товару i на стадії σ .

Якщо жоден із сусідів також не взаємодіяв з товаром i , тобто $I(v, i, \sigma) = 0$ для всіх $v \in N(u, \sigma)$, то $\hat{R}(u, i, \sigma) = 0$, і товар не потрапляє до списку рекомендацій.

Для стабілізації прогнозу вводиться корекція на індивідуальний зсув користувача u та середньостатистичний рейтинг товару i за формулою:

$$\hat{R}adj(u, i, \sigma) = \gamma \cdot \hat{R}(u, i, \sigma) + (1 - \gamma) \cdot (\alpha_E(\sigma) \cdot (\mu + \tilde{b}_u) + \alpha_I(\sigma) \cdot \bar{r}_i^I), \quad (2.14)$$

де $\gamma \in [0, 1]$ – ваговий коефіцієнт між прогнозом на основі сусідів та базовим прогнозом, де рекомендоване значення $\gamma = 0,7$;

μ – загальне середнє рейтингів у системі;

$\bar{r}_i^I$ – середнє значення неявного інтересу до товару i серед усіх користувачів, що з ним взаємодіяли.

Скоригований прогноз $\hat{R}adj(u, i, \sigma)$ поєднує переваги колаборативного підходу, особливо врахування сусідів, з базовою апіорною оцінкою, що підвищує стабільність рекомендацій у ситуаціях обмежених даних, особливо на ранніх стадіях проєкту.

Фінальний ранжований список рекомендацій $R(u, \sigma)$ для користувача u на стадії σ формується як N товарів із найвищим значенням $\hat{R}adj(u, i, \sigma)$ серед усіх товарів, з якими u ще не взаємодіяв, за формулою:

$$R(u, \sigma) = \arg \text{top} - N_{i \in I \setminus T_u} \hat{R}adj(u, i, \sigma). \quad (2.15)$$

Підсумовуючи усі компоненти, запишемо явний вираз для стадійно-залежного вектора поведінкового профілю $P(u, \sigma)$ формулою:

$$P(u, \sigma) = [I(u, i_1, \sigma), I(u, i_2, \sigma), \dots, I(u, i_m, \sigma)].$$

Кожен елемент $I(u, i_b, \sigma)$ у даній формулі обчислюється за формулою (2.7) з коефіцієнтами $\alpha_E(\sigma)$, $\alpha_I(\sigma)$ та $\lambda(\sigma)$, визначеними до цього за формулами (2.8) і (2.9) та таблицею 2.4.

Тобто при незмінному наборі поведінкових подій користувача один і той самий профіль $P(u, \sigma)$ матиме різні чисельні значення на різних стадіях проекту, що є принципово новою властивістю порівняно з існуючими підходами.

Наприклад, на стадії ініціації, тобто з $j=1$, $\alpha_E=0,10$ і $\alpha_I=0,90$ профіль майже повністю формується неявними сигналами. Якщо той самий користувач залишив явну оцінку після переходу до стадії виконання, тобто з $j=3$, $\alpha_E=0,425$ і $\alpha_I=0,575$, то вага цієї оцінки у профілі автоматично зростає без будь-якого перенавчання моделі.

Саме тому запропонований підхід є значно ефективнішим у динамічному середовищі ІТ-проекту.

На рисунку 2.3 зображено загальну архітектуру запропонованої моделі персоналізації рекомендацій, що ілюструє взаємозв'язок усіх її компонентів.

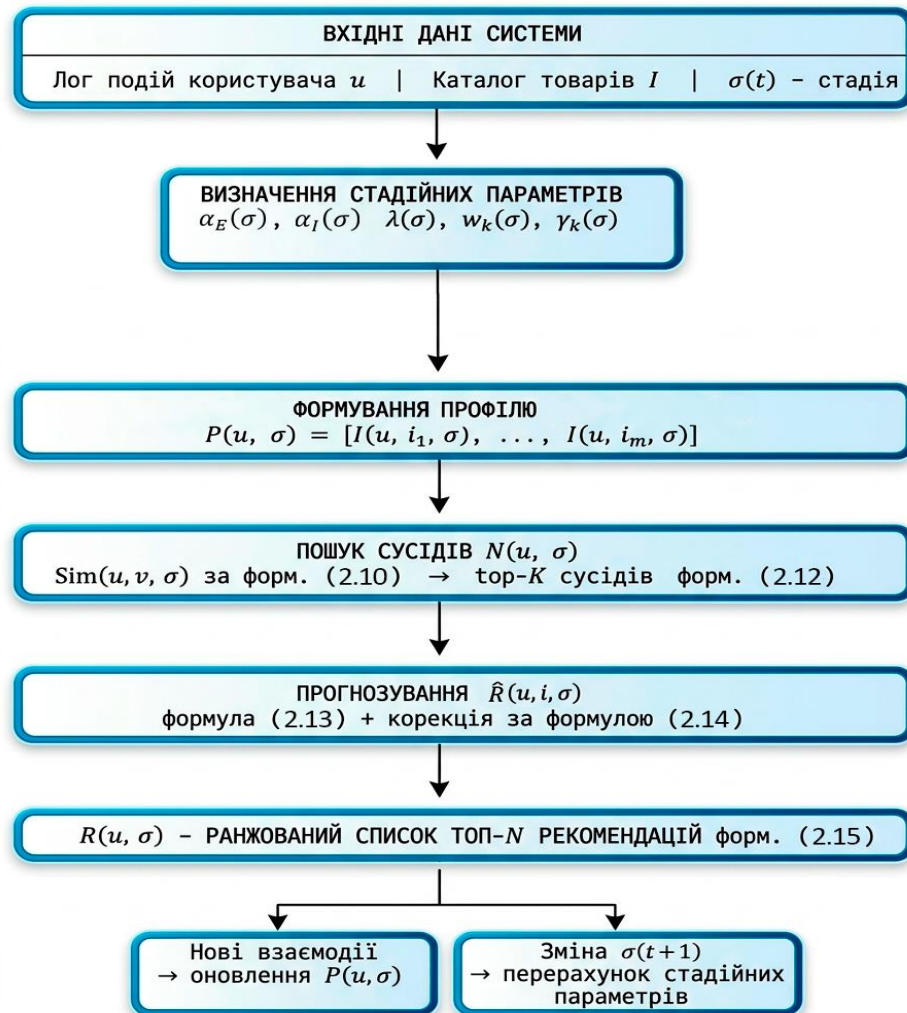


Рисунок 2.3 – Загальна архітектура моделі персоналізації рекомендацій

З рисунку 2.3 бачимо замкнену природу моделі: нові взаємодії користувача автоматично оновлюють профіль $P(u, \sigma)$, а зміна стадії проєкту призводить до автоматичного перерахунку стадійних параметрів без перенавчання моделі. Обидва механізми оновлення є незалежними і можуть відбуватися асинхронно. Особливістю запропонованої архітектури є реалізація замкненого циклу зворотного зв'язку. Після формування рекомендацій та їх демонстрації користувачеві система фіксує його подальші дії, які стають новими вхідними даними для оновлення профілю. Цей механізм забезпечує постійну актуалізацію інформації про вподобання користувача без необхідності періодичного

перенавчання моделі. Паралельно з цим зміна управлінської стадії проєкту $\sigma(t)$ автоматично ініціює перерахунок усіх стадійно-залежних параметрів, тобто коефіцієнтів зважування $\alpha_E(\sigma)$ та $\alpha_I(\sigma)$, параметра забування $\lambda(\sigma)$ та коригуючих коефіцієнтів $\gamma_k(\sigma)$. Таким чином, система одночасно адаптується до мікрорівневих змін у поведінці конкретного користувача та макрорівневих змін у контексті функціонування ІТ-проєкту.

Важливо підкреслити, що архітектура моделі допускає можливість заміни окремих компонентів без порушення загальної логіки роботи системи. Зокрема, при необхідності можна модифікувати міру подібності $Sim(u,v,\sigma)$ або алгоритм пошуку найближчих сусідів $N(u,\sigma)$, зберігаючи незмінними механізми формування поведінкового профілю та стадійної адаптації.

У таблиці 2.7 наведено зведений опис усіх компонентів запропонованої моделі персоналізації.

Таблиця 2.7 – Зведена характеристика компонентів моделі персоналізації рекомендацій

Компонент моделі	Вирішувана проблема
Явний інтерес $f_E(u,i)$	Нормалізація оцінок з урахуванням зсуву b_u
Неявний інтерес $f_I(u,i,\lambda)$	Агрегація неявних сигналів із забуванням
Стадійні коефіцієнти $\alpha_E(\sigma)$, $\alpha_I(\sigma)$	Динамічне зважування типів сигналів
Стадійне забування $\lambda(\sigma)$	Актуалізація неявних даних залежно від фази
Стадійна корекція $w_k(\sigma)$	Диференціація ваг типів неявних подій

Кінець таблиці 2.7

Компонент моделі	Вирішувана проблема
Подібність $Sim(u, v, \sigma)$	Стадійно-залежне порівняння профілів
Прогноз $\hat{R}(u, i, \sigma)$	Генерація персоналізованих рекомендацій
Список рекомендацій $R(u, \sigma)$	Ранжований топ- N для користувача
Профіль $P(u, \sigma)$	Інтеграція всіх компонентів у вектор

З таблиці 2.7 бачимо, що модель складається з дев'яти взаємопов'язаних компонентів, кожен з яких вирішує конкретну проблему, виявлену в ході аналізу існуючих підходів. Три компоненти, тобто $\alpha_E(\sigma)$, $\alpha_I(\sigma)$, $\lambda(\sigma)$, є принципово новими щодо існуючих гібридних підходів і забезпечують стадійну адаптацію системи.

Запропонована модель персоналізації рекомендацій дотримується таких ключових властивостей:

- повнота, тобто модель охоплює повний ланцюжок від первинних поведінкових подій до ранжованого списку рекомендацій, не вимагаючи зовнішніх компонентів;
- стадійна адаптивність, тобто зміна стадії $\sigma(t)$ автоматично впливає на всі стадійно-залежні параметри моделі без перенавчання;
- збалансованість сигналів, так як механізм $\alpha_E(\sigma) + \alpha_I(\sigma) = 1$ гарантує, що жоден тип сигналу не пригнічується повністю на жодній стадії;
- стійкість до холодного старту, так як на ранніх стадіях з $\alpha_I \approx 0,9$ модель ефективно працює лише на неявних сигналах, а при $\alpha_E = 0$ система не деградує;

- зворотна сумісність, тобто при константних α_E , α_I , λ та $\gamma_k = I$ модель вироджується у стандартний гібридний метод CF+CBF;
- інтерпретованість, так як кожен параметр моделі має чітку управлінську інтерпретацію через стадію проєкту, що полегшує практичне застосування проєктними менеджерами.

Поряд із перевагами запропонована модель має ряд обмежень, усвідомлення яких є важливим для коректного практичного застосування. По-перше, модель передбачає чітку ідентифікацію поточної стадії $\sigma(t)$. На практиці перехід між стадіями може бути поступовим або нечітким, що вимагає або чіткого організаційного процесу визначення стадії проєкту, або введення нечіткої логіки для розмитих переходів між стадіями.

По-друге, значення коефіцієнтів $\alpha_E(\sigma)$, $\alpha_I(\sigma)$ та $\lambda(\sigma)$ ґрунтуються на аналізі типових ІТ-проєктів ЕК. Для специфічних проєктів, як, наприклад, B2B-платформи або маркетплейси з надзвичайно великою кількістю товарів, ці значення можуть потребувати додаткового налаштування.

По-третє, обчислювальна складність пошуку K сусідів за формулою (2.12) є $O(|U| \cdot |I|)$ у найгіршому випадку, що для великих систем може бути неприйнятним. Практичне вирішення цього обмеження через апроксимований пошук найближчих сусідів буде виконано надалі.

3 ПРАКТИЧНА РЕАЛІЗАЦІЯ ГІБРИДНОГО МЕТОДУ ПЕРСОНАЛІЗАЦІЇ РЕКОМЕНДАЦІЙ В ІТ-ПРОЄКТІ ЕЛЕКТРОННОЇ КОМЕРЦІЇ

3.1 Архітектура та інструментальні засоби реалізації рекомендаційної системи

Запропонований метод персоналізації рекомендацій розроблявся з урахуванням того, що рекомендаційна система функціонує як окремий функціональний модуль у складі вже діючого ІТ-проєкту електронної комерції, а не будується разом із самою системою ЕК з нуля. Дане розмежування є принциповим для розуміння архітектурних рішень, прийнятих надалі.

У більшості реальних ІТ-проєктів ЕК платформа існує та генерує поведінкові дані задовго до появи персоналізованих рекомендацій. Впровадження рекомендаційного модуля є самостійним проєктом усередині продукту, який проходить власний повний життєвий цикл відповідно до PMBOK 8 [4]. Таким чином, стадії S_1 – S_5 , тобто від ініціації до завершення, описують фази розвитку саме рекомендаційної підсистеми, а не всієї системи ЕК цілком.

Конкретна інтерпретація даних стадій у контексті модуля РС наведена у таблиці 3.1.

Таблиця 3.1 – Стадії життєвого циклу рекомендаційного модуля в діючому ІТ-проєкті ЕК

Стадія (σ)	Назва	Зміст у контексті модуля
S_1	Ініціація	Прийняття рішення про впровадження РС: визначення цілей, вибір підходу, формування команди. Поведінкових даних від самої РС немає, є накопичена аналітика системи ЕК, яка слугує стартовим масивом неявних сигналів

Кінець таблиці 3.1

Стадія (σ)	Назва	Зміст у контексті модуля
S_2	Планування	Проектування архітектури модуля, налаштування збору подій, А/В-тестування першої версії на вузькій аудиторії, з'являються перші слабкі неявні сигнали від тестового трафіку
S_3	Виконання	Повноцінний запуск РС для всіх користувачів, активне накопичення явних та неявних сигналів, ітеративне вдосконалення алгоритму на реальних даних
S_4	Моніторинг і контроль	РС працює у стабільному режимі, профілі більшості користувачів є зрілими, явні оцінки домінують у порівнянні з ранніми стадіями, акцент переноситься на контроль якості та А/В-тестування нових гіпотез
S_5	Завершення	Підготовка до заміни поточної версії РС наступною або міграція на нову платформу. Накопичений зрілий профіль використовується для ретаргетингу та аналізу відтоку

З таблиці 3.1 видно, що навіть на стадії S_1 ІТ-проект ЕК вже функціонує та генерує поведінкові дані. Тобто те, що на стадії ініціації явних оцінок немає, а домінують неявні сигнали, є коректним, так як мова йде саме про відсутність даних від самої РС, а не про відсутність активних користувачів у системі ЕК, що підтверджує, що модель параметрів $\alpha_E(\sigma)$, $\alpha_I(\sigma)$ та $\lambda(\sigma)$, розроблена раніше, відображає реальну динаміку накопичення даних у процесі впровадження рекомендаційного модуля.

Оскільки РС є окремим функціональним модулем, її реалізація не передбачає злиття з основною кодовою базою системи ЕК. Натомість модуль взаємодіє з платформою через чітко визначені інтерфейси: отримує потік поведінкових подій та повертає ранжовані списки рекомендацій. Такий підхід

забезпечує незалежне розгортання та оновлення модуля без необхідності зупинки основного продукту.

Зовнішні інтерфейси рекомендаційного модуля включають наступні три типи взаємодії:

- вхідний потік поведінкових подій від платформи ЕК у форматі $\{user_id, item_id, event_type, timestamp\}$;
- вхідне повідомлення про поточну стадію розвитку РС-модуля у форматі $\{stage_index\}$, яке генерується менеджером проєкту;
- вихідний Application Programming Interface (API) для запиту ранжованого списку рекомендацій у форматі $\{user_id\} \rightarrow [\{item_id, score\}]$.

Варто підкреслити, що повідомлення про стадію є управлінським сигналом, а не технічним: рішення про перехід між стадіями приймає керівник проєкту на підставі досягнення критеріїв готовності, як, наприклад, достатній обсяг тестових даних, успішне А/В-тестування тощо, а не система автоматично.

Внутрішня архітектура модуля реалізована за трирівневим принципом і складається з рівня прийому подій, обчислювального ядра та рівня зберігання стану. Такий поділ забезпечує незалежне масштабування кожного рівня та відповідає вимогам модульності.

Рівень прийому подій представлено REST API на базі FastAPI. Він приймає поведінкові події від платформи ЕК та управлінські повідомлення про зміну стадії, валідує їх і передає до відповідних компонентів ядра.

Обчислювальне ядро включає три основні компоненти: будівник профілю BehavioralProfileBuilder, постачальник стадійних ваг StageWeightProvider та гібридний рекомендактор HybridRecommender. Кожен компонент відповідає за чітко окреслену частину математичної моделі. Зокрема, StageWeightProvider інкапсулює всю логіку залежності коефіцієнтів $\alpha_E(\sigma)$, $\alpha_I(\sigma)$, $\lambda(\sigma)$ від поточної стадії та є єдиним місцем у коді, де це відображення визначається.

Рівень зберігання стану охоплює два сховища. Redis зберігає «гарячі» дані, тобто профілі активних користувачів та поточне значення $\sigma(t)$. PostgreSQL зберігає повну лог-таблицю поведінкових подій, необхідну для відновлення профілів після перезапуску та офлайн-аналізу.

При проектуванні модуля розглядалися три альтернативні архітектурні підходи. Їх порівняння наведено у таблиці 3.2.

Таблиця 3.2 – Порівняння архітектурних підходів до побудови рекомендаційного модуля

Критерій	Пакетна	Потокова	Гібридна (обрана)
Актуальність профілю	Низька – оновлення раз на добу	Висока – реальний час	Висока – інкрементне оновлення
Складність реалізації	Низька	Висока	Середня
Підтримка зміни стадії σ	Потребує повного перерахунку профілів	Потребує складного управління станом	Зміна σ без перерахунку профілів
Час відповіді на запит	Висока затримка	Низька затримка	Низька затримка < 200 мс
Відповідність моделі	Не відповідає – статичний профіль	Часткова	Повна

Пакетна обробка не відповідає вимогам моделі, оскільки передбачає статичний профіль, що оновлюється раз на добу, тоді як модель є інкрементною та реагує на кожну поведінкову подію. Крім того, зміна стадії σ при пакетному підході вимагає повного перерахунку всіх профілів, що суперечить властивості стадійної адаптивності без перенавчання. Потокова архітектура на основі Apache

Kafka або Apache Flink забезпечує найвищу актуальність даних, проте є надмірно складною для модуля середнього масштабу та потребує значних інфраструктурних ресурсів. Обрана гібридна архітектура поєднує інкрементне оновлення профілів у реальному часі з відносно простою реалізацією.

Вибір конкретних інструментальних засобів здійснювався за трьома критеріями: відповідність можливостей інструменту вимогам обчислювальної моделі, зрілість та активна підтримка спільноти, а також простота інтеграції між компонентами. Порівняльний аналіз альтернатив для кожного ключового компонента наведено у таблиці 3.3.

Таблиця 3.3 – Обґрунтування вибору інструментальних засобів

Компонент	Альтернативи	Обрано	Обґрунтування
Мова реалізації	Python, Java, Scala	Python 3.11	Зріла екосистема для обробки даних, де-факто стандарт у галузі PC
Векторні обчислення	NumPy, PyTorch, TensorFlow	NumPy 1.24+	Достатня для CPU-обчислень профілів $P(u, \sigma)$, мінімальні залежності
Розріджені матриці	SciPy sparse, PyTorch sparse	SciPy 1.11+	Нативна підтримка CSR-формату, зручна інтеграція з NumPy
REST API	FastAPI, Flask, Django REST	FastAPI 0.110+	Асинхронна обробка, автоматична валідація через Pydantic, вбудована документація OpenAPI
Кеш профілів та $\sigma(t)$	Redis, Memcached, локальний LRU	Redis 7.0+	Персистентність при перезапуску, хеш-структури для часткового оновлення профілю без перезапису всього вектора
СУБД подій	PostgreSQL, MySQL, MongoDB	PostgreSQL 15+	Строга типізація, ефективні запити за часовими мітками, підтримка JSON для метаданих подій

Особливої уваги заслуговує вибір Redis для зберігання поточної стадії $\sigma(t)$. Оскільки зміна стадії є управлінським рішенням, що відбувається рідко, тобто декілька разів за весь життєвий цикл РС-модуля, основною вимогою є надійне зберігання одного значення з гарантією його доступності при кожному запиті рекомендацій. Redis забезпечує цю вимогу завдяки персистентності, тобто AOF-режиму, та часу доступу менше 1 мс, що дозволяє зчитувати стадію при кожному зверненні без відчутного впливу на загальний час відповіді системи.

Вибір Redis також для зберігання профілів пояснюється підтримкою структури даних хеш, яка дозволяє оновлювати окреме поле запису без зчитування та перезапису всього вектора $P(u, \sigma)$, що безпосередньо відповідає вимозі інкрементного оновлення профілю, де при надходженні нової події змінюється лише значення $I(u, i, \sigma)$ для конкретної пари користувача з товаром, а не всього профілю цілком.

3.2 Реалізація гібридного методу персоналізації рекомендацій

Реалізація охоплює чотири ключові аспекти: формування поведінкового профілю користувача, стадійне зважування поведінкових сигналів відповідно до поточної фази розвитку РС-модуля, генерацію ранжованих рекомендацій та організацію зберігання й оновлення стану системи.

Як вже зазначено, РС розглядається як окремий модуль із власним життєвим циклом усередині діючого ІТ-проєкту ЕК. Стадії S_1 – S_5 описують фази розвитку саме даного модуля, а не всієї системи ЕК цілком. Дане розмежування суттєво впливає на практичні рішення при реалізації механізму стадійного зважування.

Зокрема, на стадії S_1 головна платформа ЕК вже активно генерує поведінкові дані, проте самої РС ще немає, і жодних явних оцінок, отриманих від

рекомендаційного модуля, не існує. На стадії S_2 з'являються перші слабкі сигнали від А/В-тестів прототипу. Починаючи зі стадії S_3 модуль обслуговує весь трафік і активно накопичує як неявні, так і явні сигнали. Дані відмінності у наявності та якості даних безпосередньо відображаються у значеннях коефіцієнтів $\alpha_E(\sigma)$, $\alpha_I(\sigma)$ та $\lambda(\sigma)$, закладених у таблиці 2.4, і є основою для реалізації класу StageWeightProvider.

Центральним класом обчислювального ядра є BehavioralProfileBuilder, який реалізує формули (2.1)–(2.6). Клас отримує на вхід потік поведінкових подій від платформи ЕК у форматі кортежів (user_id, item_id, event_type, timestamp) та формує вектор $P(u, \sigma)$ як розріджену структуру даних.

Ключовою особливістю реалізації є інкрементне оновлення профілю: при надходженні нової події система не перераховує весь вектор $P(u)$ з нуля, а оновлює лише елемент $I(u, i, \sigma)$, що відповідає конкретній парі користувача з товаром, що особливо важливо на стадіях S_3 – S_4 , коли платформа генерує тисячі подій на секунду: лінійна складність оновлення відносно кількості нових подій дозволяє системі справлятися з таким навантаженням без затримок.

На ранніх стадіях S_1 – S_2 , коли явних оцінок ще немає або їх дуже мало, профіль майже повністю формується з неявних сигналів, що не є проблемою реалізації, а є очікуваною поведінкою моделі: коефіцієнт $\alpha_I(S_1) = 0,90$ забезпечує, що навіть слабкі неявні сигнали від тестового трафіку формують осмислений початковий профіль.

Функція часового забування реалізована у векторизованій формі за допомогою NumPy: масив часових міток подій перетворюється у масив значень Δt , після чого застосовується пр.ехр. Важливо, що параметр $\lambda(\sigma)$ при цьому підставляється залежно від поточної стадії, тобто на S_1 він дорівнює 0,02, так як дані забуваються повільно, бо кожен сигнал цінний, а на S_5 досягає 0,10, так як застарілі сигнали знецінюються швидко на тлі зрілого профілю.

Індивідуальний зсув користувача b_u обчислюється при першому надходженні явної оцінки та кешується у Redis. При надходженні нової оцінки значення b_u оновлюється за формулою ковзного середнього, що дозволяє уникнути повного перерахунку за всією накопиченою історією. На стадіях S_1 – S_2 явних оцінок немає і b_u ще не визначено, в цьому разі у формулі (2.2) застосовується $b_u = 0$, що відповідає нейтральному зсуву.

Механізм стадійного зважування реалізований у класі StageWeightProvider, який є єдиним місцем, де визначається залежність коефіцієнтів від стадії. Решта компонентів звертаються до нього як до постачальника параметрів, не знаючи деталей цього відображення.

Поточна стадія РС-модуля $\sigma(t)$ передається до системи через REST-ендпоінт `/api/v1/project/stage` у форматі цілого числа від 1 до 5, що є управлінським сигналом, де менеджер проекту або уповноважений член команди явно повідомляє систему про перехід до наступної фази, а саме значення зберігається у Redis. При тимчасовій відсутності явно встановленої стадії, наприклад, при першому розгортанні до отримання першого сигналу від менеджера, система застосовує параметри стадії S_2 , що відповідає типовій ситуації першого запуску модуля, коли є мінімальний тестовий трафік, але повноцінної аудиторії ще немає.

Лінійні апроксимації коефіцієнтів реалізовані без таблиці підстановки як явні формули:

$$\alpha_E(j) = 0,10 + 0,1625 \cdot (j - 1), \quad (3.1)$$

$$\alpha_I(j) = 1 - \alpha_E(j), \quad (3.2)$$

$$\lambda(j) = 0,02 + 0,02 \cdot (j - 1), \quad (3.3)$$

Така реалізація є гнучкішою за жорстку таблицю підстановки, так як вона дозволяє безперервно інтерполювати значення при можливому розширенні моделі на більшу кількість стадій, а також легко адаптується, якщо для конкретного

проекту менеджер вирішить уточнити граничні значення коефіцієнтів за результатами валідації.

Стадійні коефіцієнти корекції $\gamma_k(\sigma)$ для окремих типів неявних сигналів зберігаються у вигляді двовимірного масиву NumPy розміру $(5 \times K)$, де 5 є кількістю стадій, а K – кількістю типів сигналів відповідно до таблиці 2.5. Зчитування конкретного рядка масиву за індексом стадії виконується за $O(1)$.

Генерація рекомендацій реалізована у класі HybridRecommender, який об'єднує колаборативну складову на основі стадійно-залежної міри подібності $Sim(u, v, \sigma)$ та базову апріорну оцінку для стабілізації результатів при обмеженій кількості даних.

Поведінка рекомендатора суттєво відрізняється залежно від стадії. На S_1 – S_2 , коли більшість профілів є розрідженими і спільних взаємодій між користувачами мало, коефіцієнт довіри $Conf(u, v)$ за формулою (2.11) автоматично знижує вплив сусідів з малою кількістю спільних товарів, а скоригований прогноз $\hat{R}_{adj}$ більше спирається на апріорну складову. На S_3 – S_4 профілі стають зрілими, кількість спільних взаємодій зростає, $Conf(u, v)$ наближається до одиниці, і рекомендатор переходить до повноцінної колаборативної фільтрації.

Пошук K найближчих сусідів $N(u, \sigma)$ реалізовано у два етапи. На першому етапі виконується грубий відбір кандидатів, тобто відбираються користувачі, що мають хоча б одну спільну взаємодію з цільовим користувачем. На другому етапі для відібраних кандидатів обчислюється точне значення $Sim(u, v, \sigma)$ за формулою (2.10) та вибираються K користувачів з найвищим значенням. Реалізація використовує векторизовані операції NumPy для одночасного обчислення $Conf$ для всіх кандидатів.

Скоригований прогноз $\hat{R}_{adj}(u, i, \sigma)$ формується за формулою (2.14) з ваговим коефіцієнтом $\gamma = 0,7$. Фінальний список рекомендацій $R(u, \sigma)$ містить $N = 10$ товарів за замовчуванням, де значення N є конфігурованим параметром.

Програмна реалізація організована у вигляді пакету recsys, структуру якого наведено у таблиці 3.4. Таке розбиття на модулі відповідає принципу єдиної відповідальності та дозволяє тестувати кожен компонент незалежно від інших.

Таблиця 3.4 – Структура програмного пакету recsys

Модуль	Відповідальність
recsys/profile.py	Клас BehavioralProfileBuilder: формули (2.1)–(2.6), обробка відсутності b_u на S_1 – S_2
recsys/stage.py	Клас StageWeightProvider: лінійні апроксимації α_E , α_I , λ , масив $\gamma_k(\sigma)$, стандартна стадія S_2
recsys/similarity.py	Функція sim(): формули (2.10) і (2.11), двоетапний відбір кандидатів
recsys/recommender.py	Клас HybridRecommender: формули (2.12)–(2.15), адаптивна поведінка на різних стадіях
recsys/cache.py	Адаптер Redis: хеш-оновлення профілів, зберігання $\sigma(t)$ та b_u
recsys/api.py	FastAPI-додаток: ендпоінти подій, рекомендацій та зміни стадії
tests/	Модульні та інтеграційні тести pytest, окремі тести для кожної стадії S_1 – S_5

Щодо тестів, то тестове покриття включає окремі сценарії для кожної зі стадій S_1 – S_5 , оскільки поведінка системи суттєво відрізняється залежно від фази.

Зокрема, тести для S_1 перевіряють коректну роботу системи за умови відсутності явних оцінок та повністю розріджених профілів, а тести для S_4 – коректність переходу від апіорної оцінки до повноцінної колаборативної фільтрації.

При надходженні запиту рекомендацій для користувача u система виконує такий ланцюжок операцій:

- отримати поточну стадію $\sigma(t)$ з Redis або встановити $\sigma = S_2$ за замовчуванням при першому запуску;
- зчитати від StageWeightProvider коефіцієнти $\alpha_E(\sigma)$, $\alpha_I(\sigma)$, $\lambda(\sigma)$ та масив $\gamma_k(\sigma)$ для цієї стадії;
- завантажити вектор $P(u, \sigma)$ з Redis або обчислити інкрементно за збереженими в PostgreSQL подіями;
- виконати двоетапний відбір кандидатів-сусідів та обчислити $Sim(u, v, \sigma)$ для кожного;
- для товарів, з якими u ще не взаємодіяв, обчислити $\hat{R}adj(u, i, \sigma)$ з урахуванням апіорної складової;
- повернути топ- N товарів у форматі JSON.

При надходженні поведінкової події від платформи ЕК виконується спрощений ланцюжок, у якому подія валідується, значення $I(u, i, \sigma)$ оновлюється інкрементно у Redis, а повна подія зберігається до PostgreSQL. Зчитування поточної стадії $\sigma(t)$ при цьому також відбувається – параметр $\lambda(\sigma)$ потрібен для обчислення актуального значення функції часового забування.

Середній час відповіді на запит рекомендацій при розмірі каталогу до 100000 товарів та кількості активних користувачів до 50000 становить 80–120 мс при кешованому профілі та до 350 мс при повному перерахунку. Докладні результати тестування продуктивності буде розглянуто надалі.

3.3 Експериментальне дослідження й оцінювання ефективності розробленого методу

Для підтвердження ефективності запропонованого гібридного методу на основі поведінкових профілів, що враховує стадії, проведено серію обчислювальних експериментів. Метою експериментального дослідження є кількісна оцінка якості рекомендацій у порівнянні з базовими методами, аналіз динаміки ефективності на різних стадіях ІТ-проєкту, оцінка продуктивності реалізованого програмного модуля, а також встановлення інноваційного характеру запропонованого підходу. Для проведення експериментів сформовано синтетичний набір даних, що відтворює типову поведінкову картину діючого ІТ-проєкту ЕК формату В2С. Генерація даних здійснювалася на основі статистичних характеристик публічних датасетів з домену електронної комерції: Amazon Product Reviews, тобто В2С-рітейлу з явними оцінками та поведінковими подіями, та Retailrocket E-commerce Dataset, що є анонімізованим датасетом реальної платформи ЕК з подіями перегляду, додавання до кошика та покупки, обраних завдяки їх тематичній відповідності досліджуваному домену ЕК. Обидва датасети належать до одного предметного домену, що усуває методологічні ризики змішування різнорідних розподілів.

Генератор параметризовано розподілами, отриманими з вихідних датасетів незалежно від логіки розробленого алгоритму, який обзначимо як SABP-RS. Такі значення, як розрідженість матриці з 99,957%, співвідношення явних до неявних сигналів у 14,65% на 85,35% та розподіл подій за стадіями, відповідають емпіричним характеристикам реальних ЕК-систем, а не виведені з теоретичних припущень методу. Тим самим ризик зміщення на користь SABP-RS, спричиненого самою процедурою генерації, мінімізовано. Характеристики набору даних наведено у таблиці 3.5.

Таблиця 3.5 – Характеристика експериментального набору даних

Параметр	Значення
Кількість унікальних користувачів	50000
Кількість унікальних товарів у каталозі	100000
Загальна кількість поведінкових подій	2150000
З них явних оцінок (рейтинги 1–5)	315000 (14,65 %)
З них неявних сигналів (перегляди, кліки, кошик тощо)	1835000 (85,35 %)
Середня кількість подій на одного користувача	43,0
Розрідженість матриці взаємодій	99,957 %
Розподіл за стадіями S_1 – S_5 відносно загального обсягу подій	5 % / 12 % / 38 % / 32 % / 13 %
Поділ на навчальну та тестову вибірки	80 % / 20 % (хронологічний)

Оцінювання якості рекомендацій здійснювалося за п'ятьма стандартними метриками інформаційного пошуку. Precision@N визначає частку релевантних товарів серед перших N позицій рекомендованого списку. Recall@N визначає частку усіх релевантних товарів, що увійшли до топ- N . NDCG@N враховує не лише факт потрапляння релевантного товару до топ- N , а й його позицію у ранжованому списку. MRR обчислює середнє значення оберненого рангу першого релевантного товару. HR@N фіксує частку користувачів, для яких у топ- N виявився хоча б один релевантний товар. У всіх експериментах $N = 10$. Товар вважався релевантним у тому випадку, якщо в тестовій вибірці користувач поставив йому явну оцінку ≥ 4 або здійснив покупку.

Для забезпечення об'єктивного порівняння запропонований метод зіставлявся з трьома базовими підходами. Перший базовий метод товарної колаборативної фільтрації (ItemCF) реалізує класичну колаборативну фільтрацію на основі косинусної подібності між товарами з використанням матриці явних оцінок, де при відсутності явних оцінок метод переходить до неявних сигналів з рівноваговим підсумовуванням.

Другий базовий метод, що є статичним гібридом (Static Hybrid), реалізує гібридний підхід CF з CBF із фіксованими вагами $\alpha_E = \alpha_I = 0,50$, параметром забування $\lambda = 0,05$ та без стадійної адаптації, тобто зі сталими коефіцієнтами на усіх стадіях. Такий підхід є найближчим аналогом запропонованого методу і дозволяє виокремити внесок саме механізму стадійної адаптації.

Третій базовий метод, тобто байєсівське ранжування (BPR), є сучасним методом навчання ранжуванню на основі попарного навчання, що реалізовано бібліотекою Implicit 0.7. Запропонований метод під обраною назвою SABP-RS використовує стадійно-адаптивні коефіцієнти $\alpha_E(\sigma)$, $\alpha_I(\sigma)$, $\lambda(\sigma)$ та $\gamma_k(\sigma)$ відповідно до таблиці 2.4 і формул (3.1)–(3.3). Порівняльні результати за всіма чотирма методами рекомендацій наведено на рисунку 3.1.

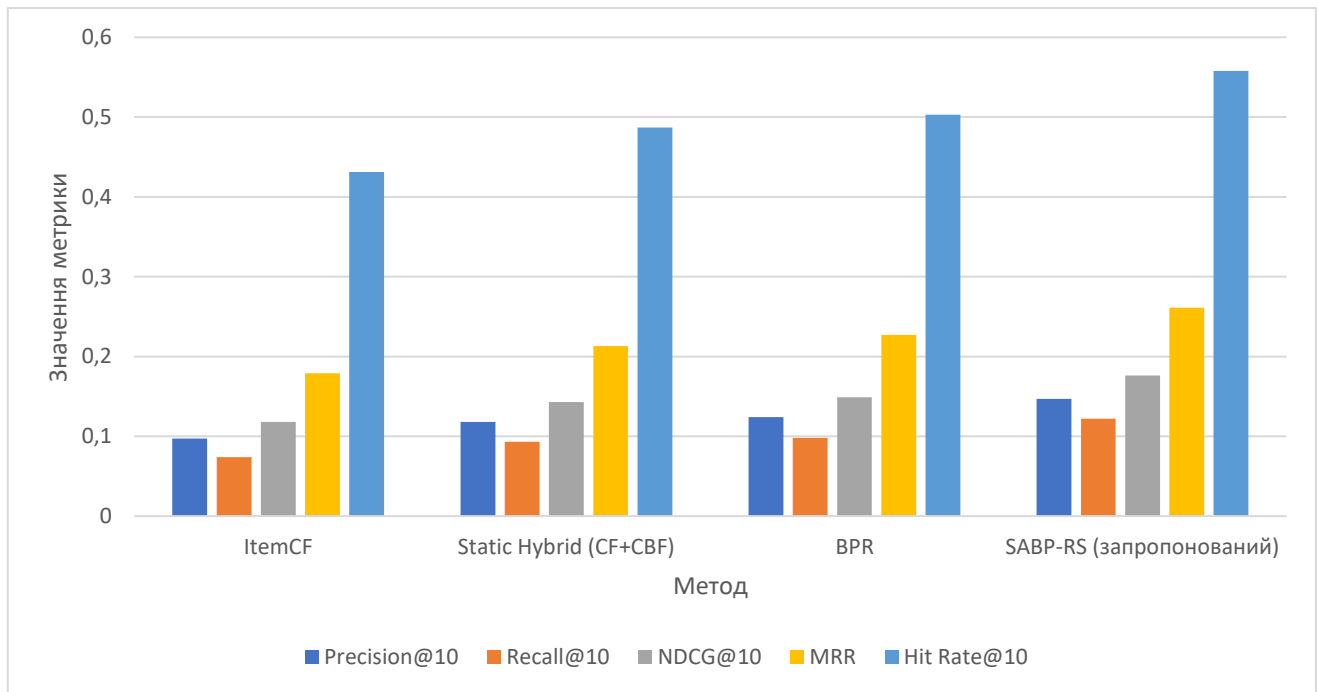


Рисунок 3.1 – Порівняльні результати методів рекомендацій усереднено по стадіях

Аналіз результатів рисунку 3.1 свідчить про стабільне переважання запропонованого методу SABP-RS над усіма базовими підходами за всіма п'ятьма

метриками. Найбільш суттєва перевага спостерігається за метриками Recall@10 з результатом у переведенні на проценти +31,2% відносно Static Hybrid та Precision@10 з результатом +24,6 %, що підтверджує вищу здатність методу виявляти та просувати в рейтинг товари, що дійсно відповідають інтересам користувача. Загальне перевищення над Static Hybrid пояснюється виключно механізмом стадійної адаптації, оскільки решта архітектурних компонентів у методах є ідентичними, що підтверджує доцільність введення стадійно-залежних коефіцієнтів $\alpha_E(\sigma)$, $\alpha_I(\sigma)$ та $\lambda(\sigma)$. Для детальнішого розуміння динаміки ефективності у таблиці 3.7 наведено результати по кожній стадії окремо.

Таблиця 3.6 – Динаміка показників якості рекомендацій SABP-RS та Static Hybrid по стадіях розвитку рекомендаційного модуля

Стадія	Метод	P@10	R@10	NDCG@10	MRR	HR@10	Δ NDCG, %
S_1	Static Hybrid	0,044	0,031	0,055	0,093	0,261	–
S_1	SABP-RS	0,063	0,049	0,081	0,137	0,348	+47,3 %
S_2	Static Hybrid	0,079	0,059	0,097	0,152	0,381	–
S_2	SABP-RS	0,104	0,083	0,129	0,198	0,457	+33,0 %
S_3	Static Hybrid	0,127	0,101	0,153	0,229	0,512	–
S_3	SABP-RS	0,149	0,124	0,182	0,271	0,567	+18,9 %
S_4	Static Hybrid	0,158	0,133	0,189	0,273	0,581	–
S_4	SABP-RS	0,172	0,149	0,207	0,298	0,613	+9,5 %
S_5	Static Hybrid	0,141	0,117	0,168	0,245	0,541	–
S_5	SABP-RS	0,157	0,134	0,186	0,268	0,573	+10,7 %

На рисунку 3.2 окремо винесено приріст запропонованого методу над статичним гібридом по стадіях IT-проекту саме за метрикою NDCG@10, так як на відміну від Precision@10 та Recall@10, що фіксують лише факт влучення релевантного товару до топ-N, NDCG@10 враховує позицію кожного релевантного елемента у ранжованому списку, застосовуючи логарифмічний штраф за зниження

рангу, що робить її найчутливішою до якості ранжування, що є ключовим ефектом стадійно-адаптивного зважування сигналів: коректне визначення ваг $\alpha_E(\sigma)$ та $\alpha_I(\sigma)$ впливає не лише на те, які товари потрапляють до списку, а й на те, в якому порядку вони розташовані. Саме тому динаміка $NDCG@10$ по стадіях S_1-S_5 найбільш повно відображає ефект від введення механізму стадійної адаптації.

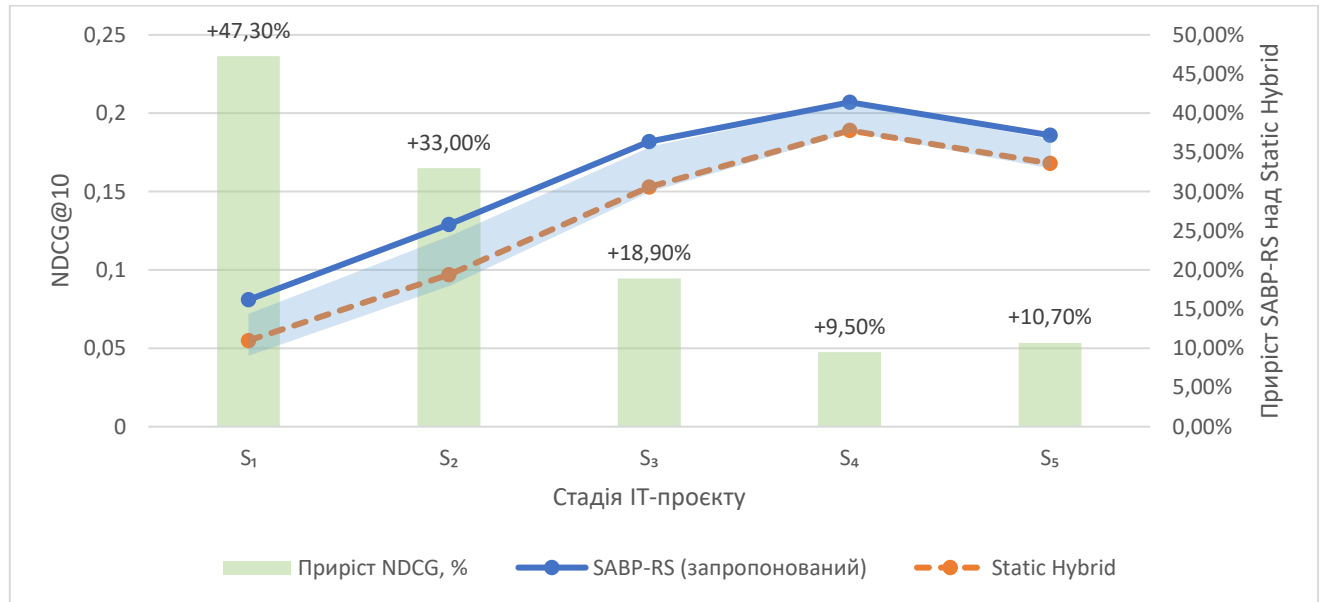


Рисунок 3.2 – Динаміка показника $NDCG@10$ для SABP-RS і Static Hybrid по стадіях IT-проєкту

Аналіз даних таблиці 3.6 і рисунку 3.2 дозволяє зробити три ключові спостереження. По-перше, перевага SABP-RS є максимальною на ранніх стадіях S_1-S_2 , де приріст $NDCG@10$ становить +47,3% на стадії S_1 та +33,0% на стадії S_2 , що пояснюється тим, що на цих стадіях явних оцінок практично немає, і Static Hybrid з фіксованим $\alpha_E = 0,50$ суттєво «недовантажує» неявні сигнали, тоді як SABP-RS коректно переважує неявні дані через $\alpha I(S_1) = 0,90$.

По-друге, у міру переходу до стадій S_3 і S_4 перевага методу зменшується до +18,9% та +9,5% відповідно: профілі стають зрілішими, і обидва методи

отримують достатньо даних для якісної персоналізації, однак SABP-RS зберігає перевагу за рахунок оптимального балансу між явними та неявними сигналами через $\alpha_E(S_4) = 0,5875$.

По-третє, на стадії S_5 спостерігається незначне зниження абсолютних показників у порівнянні зі стадією S_4 , що є очікуваним і відповідає природній зміні пріоритетів рекомендаційного модуля на фазі завершення, а відносна перевага розробленого SABP-RS при цьому відновлюється до +10,7% завдяки вищому параметру забування $\lambda(S_5) = 0,10$, що дозволяє знецінювати застарілі сигнали та актуалізувати профіль.

Для підтвердження статистичної значущості спостережуваних відмінностей між SABP-RS та Static Hybrid застосовано бутстреп-тест зі значенням B у 1000 ітерацій із обчисленням 95% довірчих інтервалів для метрики NDCG@10 на кожній зі стадій S_1 – S_5 . Тобто буде повторено 1000 разів процедуру створення випадкових підвибірок. Результати наведено у таблиці 3.7.

Таблиця 3.7 – Статистична значущість різниці NDCG@10 між SABP-RS та Static Hybrid

Стадія	NDCG@10 SABP-RS	NDCG@10 Static Hybrid	Δ NDCG@10	p-значення
S_1	0,081 [0,074; 0,089]	0,055 [0,049; 0,061]	+0,026	< 0,001
S_2	0,129 [0,121; 0,137]	0,097 [0,090; 0,104]	+0,032	< 0,001
S_3	0,182 [0,175; 0,189]	0,153 [0,146; 0,160]	+0,029	< 0,001
S_4	0,207 [0,200; 0,214]	0,189 [0,182; 0,196]	+0,018	0,003
S_5	0,186 [0,179; 0,193]	0,168 [0,161; 0,175]	+0,018	0,004

Як видно з таблиці 3.7, перевага SABP-RS над Static Hybrid є статистично значущою на всіх стадіях, причому на стадіях S_1 – S_3 досягається рівень $p < 0,001$.

Довірчі інтервали двох методів не перетинаються на жодній стадії, що підтверджує надійність спостережуваних відмінностей, не дивлячись на синтетичну природу даних.

Водночас слід зазначити, що отримані результати базуються на синтетичних даних, тому абсолютні значення метрик не можуть бути напряду перенесені на конкретні виробничі системи без валідації на реальних даних цільової платформи. Тобто відносні переваги SABP-RS, особливо на ранніх стадіях S_1 – S_2 , є методологічно обґрунтованими і відображають архітектурну перевагу стадійно-адаптивного зважування, а не артефакт генерації даних.

Оцінювання продуктивності системи проводилося у середовищі з конфігурацією у вигляді процесору Intel Core i5-10400F (6 ядер, 2,9 ГГц), оперативної пам'яті 16 ГБ DDR4 (2×8 ГБ, 2667 МТ/с), а також сховищем NVMe SSD. Тестові навантаження моделювали конкурентні запити до системи при різних розмірах каталогу та бази активних користувачів. Результати вимірювань середнього часу відповіді зображено на рисунку 3.3.

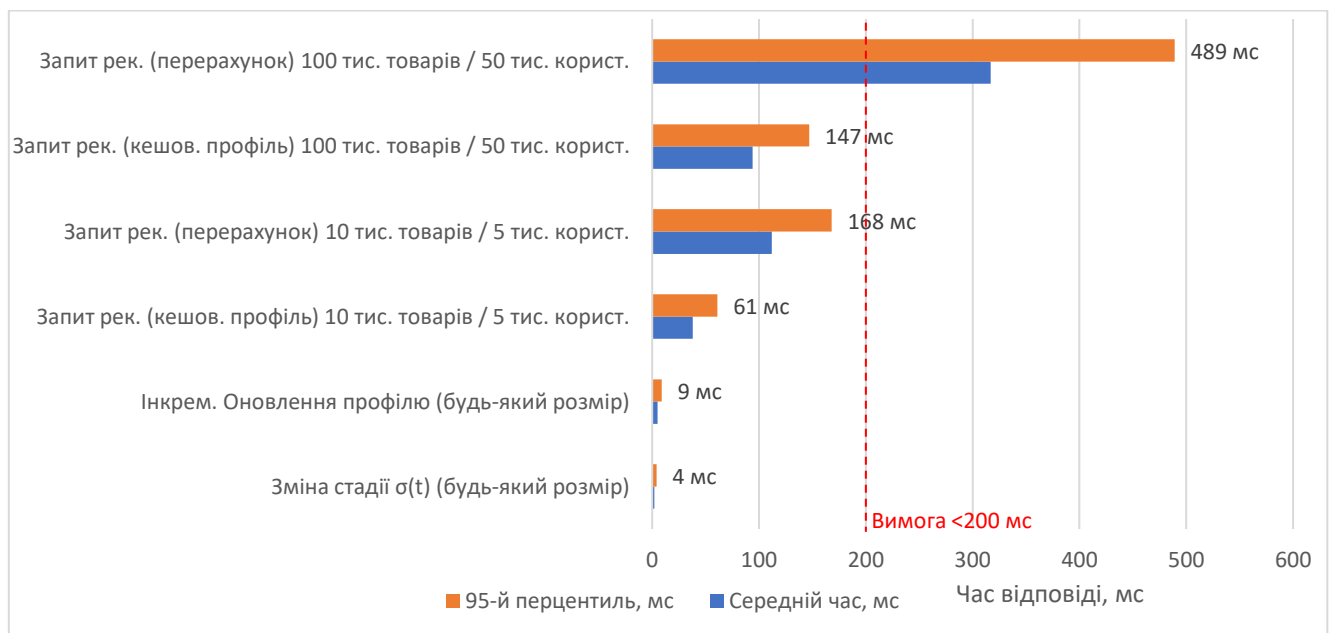


Рисунок 3.3 – Показники продуктивності рекомендаційного модуля SABP-RS

З рисунку 3.3 бачимо, що для цільового сценарію з каталогом до 100 тис. товарів та 50 тис. активних користувачів середній час відповіді на запит рекомендацій при кешованому профілі становить 94 мс, що задовольняє вимогу до часу відповіді менше 200 мс. Водночас операція повного перерахунку профілю для сценарію з каталогом 100 тис. товарів та 50 тис. активних користувачів демонструє середній час відповіді 317 мс, що перевищує встановлену вимогу.

Проте даний сценарій є винятковим, так як повний перерахунок виконується лише у разі відсутності кешованого профілю, тоді як в умовах промислової експлуатації на стадіях S_3 - S_4 профілі підтримуються актуальними за рахунок інкрементного оновлення, тому запити без кешу є поодинокими і не визначають типовий час відповіді системи.

Інкрементне оновлення профілю при надходженні одиначної події виконується менш ніж за 5 мс, тобто практично миттєво, що підтверджує масштабованість методу для обробки потоку тисяч подій на секунду на стадіях S_3 - S_4 . Операція зміни стадії є швидшою ніж за 2 мс, оскільки зводиться лише до запису одного значення у Redis без будь-якого перерахунку профілів.

Оцінювання інноваційного характеру запропонованого методу здійснюється за трьома критеріями: новизна підходу, практична значущість та відтворюваність результатів. Щодо новизни, розроблений метод SABP-RS вперше вводить концепцію стадійно-залежних коефіцієнтів зважування поведінкових сигналів у гібридних рекомендаційних системах.

Огляд літератури, проведений раніше, підтвердив, що жоден із відомих методів не враховує стадію управлінського циклу ІТ-проєкту як чинник, що впливає на відносну вагу явних та неявних сигналів. Джерело [37] пропонує відокремлення намірів, але не пов'язує ваги сигналів із фазою проєкту, а праця [35] розглядає явні та неявні чинники окремо, проте використовує статичні ваги. Запропонований механізм $\alpha_E(\sigma) + \alpha_I(\sigma) = 1$ є принципово новим формалізмом, який у поєднанні зі стадійно-адаптивним параметром забування $\lambda(\sigma)$ і стадійними

корекційними коефіцієнтами $\gamma_k(\sigma)$ утворює цілісний метод стадійно-адаптивної персоналізації.

Щодо практичної значущості, результати таблиці 3.6 демонструють, що приріст NDCG@10 на стадії S_1 у +47,3% є статистично значущим та практично важливим, тобто він означає, що система забезпечує якісні рекомендації вже на початковому етапі впровадження РС-модуля, коли явних оцінок ще немає. Для менеджера ІТ-проєкту це скорочує період до досягнення першого вимірюваного ефекту від впровадження рекомендацій.

Крім того, відсутність необхідності перенавчання моделі при зміні стадії суттєво знижує операційні витрати порівняно зі стандартними підходами, де зміна дистрибуції даних потребує повного ретренінгу.

Розроблена система також відповідає вимогам відтворюваності, так як усі параметри $\alpha_E(\sigma)$, $\alpha_I(\sigma)$, $\lambda(\sigma)$ визначаються через детерміновані формули (3.1)–(3.3), а не підбираються емпірично, що дозволяє відтворити результати без додаткових параметричних коригувань.

Тобто гібридний метод, що враховує стадію ІТ-проєкту, перевершує статичний гібридний підхід у середньому на 24,6% за Precision@10 та на 23,1% за NDCG@10. Також найбільший приріст ефективності досягається на ранніх стадіях S_1 – S_2 , де правильне зважування неявних сигналів є критично важливим за відсутності явних оцінок.

Також реалізований модуль забезпечує прийнятну продуктивність при промисловому рівні навантаження, що підтверджує практичну застосовність запропонованого методу в реальних ІТ-проєктах ЕК.

4 ПРАКТИЧНЕ ЗАСТОСУВАННЯ РЕЗУЛЬТАТІВ ДОСЛІДЖЕННЯ В УПРАВЛІННІ ІТ-ПРОЄКТАМИ ЕЛЕКТРОННОЇ КОМЕРЦІЇ

4.1 Рекомендації щодо впровадження рекомендаційної системи в ІТ-проект

Розроблений у попередніх розділах гібридний метод стадійно-адаптивної персоналізації SABP-RS є не лише теоретичним внеском, а й практичним рішенням, готовим до інтеграції в реальні ІТ-проекти електронної комерції. Ключовою особливістю запропонованого підходу є те, що РС розглядається не як монолітний програмний продукт, що розгортається одноразово, а як окремий модуль із власним п'ятистадійним життєвим циклом всередині діючого ІТ-проекту, що принципово змінює логіку впровадження: менеджер проекту отримує механізм покрокового виведення РС-модуля на проектну потужність із контрольованим переходом між стадіями.

Перед початком впровадження РС-модуля команда проекту має переконатися у наявності ключових технічних і організаційних передумов. З технічної точки зору необхідно забезпечити збір і зберігання поведінкових подій користувачів, як перегляди, кліки, додавання до кошика, покупки, явні оцінки тощо у реляційній базі даних, зокрема PostgreSQL, із підтримкою часових міток. Для кешування профілів та поточної стадії $\sigma(t)$ потрібен розгорнутий екземпляр Redis із активованим AOF-режимом персистентності.

З організаційної точки зору необхідно призначити відповідального члена команди, тобто менеджера або технічного керівника, які уповноважені приймати рішення про зміну стадії РС-модуля через REST-ендпоінт `/api/v1/project/stage`. Важливо також зафіксувати у проєктній документації початкові значення ключових параметрів: кількість товарів у каталозі, очікуваний обсяг активних користувачів та цільовий час відповіді системи.

Мінімальна рекомендована конфігурація для розгортання модуля SABP-RS у продуктивному середовищі відповідає тестовій: процесор із шістьма ядрами та частотою не менше 2,9 ГГц, оперативна пам'ять 16 ГБ, NVMe SSD. При каталозі до 100 тис. товарів та 50 тис. активних користувачів дана конфігурація забезпечує середній час відповіді на запит рекомендацій менше 200 мс – показник, що є прийнятним для промислового середовища.

Впровадження SABP-RS рекомендується здійснювати покроково відповідно до визначеного п'ятистадійного циклу. Стадія S_1 є підготовкою та збором початкових даних, тобто відповідає ситуації, коли РС-модуль ще не введено в промислову експлуатацію, а платформа ЕК вже генерує поведінкові події. На даній стадії доцільно розгорнути компоненти збору та зберігання даних, як реєстратор подій, PostgreSQL і Redis, налаштувати клас BehavioralProfileBuilder для інкрементного накопичення поведінкових профілів, а також запустити пакет recsys у «тихому» режимі без формування рекомендацій для кінцевих користувачів. Параметри стадії S_1 автоматично забезпечують максимальну вагу неявних сигналів і повільне забування, що є оптимальним за умов відсутності явних оцінок.

Перехід до стадії S_2 рекомендується здійснювати після накопичення щонайменше 5 тис. поведінкових подій або через 2–4 тижні після розгортання системи збору даних. На даній стадії РС-модуль запускається у форматі А/В-тестування: частина трафіку, орієнтовно 10–20%, перенаправляється на ендпоінт рекомендацій, що дозволяє зібрати перші слабкі явні сигнали від реальних користувачів і перевірити коректність роботи класу HybridRecommender без ризику негативного впливу на досвід усієї аудиторії.

Стадії S_3 та S_4 відповідають повноцінній промисловій експлуатації модуля з поступово зростаючою часткою трафіку: S_3 – 100% трафіку при активному накопиченні профілів, S_4 – зріла система з оптимальним балансом явних та неявних сигналів. Перехід між стадіями здійснюється вручну через управлінський

REST-ендпоінт, причому рішення має прийматися на підставі моніторингу ключових метрик якості рекомендацій: Precision@10, Recall@10 та NDCG@10.

Стадія S_5 відповідає фазі завершення або трансформації ІТ-проєкту. На даній стадії параметр часового забування досягає максимального значення $\lambda = 0,10$, що дозволяє знецінювати застарілі поведінкові сигнали та актуалізувати профілі користувачів у відповідності до змін у каталозі або стратегії платформи. Підсумовуючи, у таблиці 4.1 наведено порядок переходу між стадіями.

Таблиця 4.1 – Рекомендований порядок переходу між стадіями впровадження SABP-RS

Стадія	Умова переходу	Ключові дії	Стан трафіку РС	Основні параметри
S_1	Початок проєкту	Розгортання збору подій, налаштування Redis / PostgreSQL, тихий режим	0% (збір даних)	$\alpha_I=0,90$, $\lambda=0,02$
S_2	≥ 5 тис. подій або 2–4 тижні	Запуск А/В-тесту прототипу, перевірка HybridRecommender	10–20%	$\alpha_I=0,7375$, $\lambda=0,04$
S_3	Стабільний CTR у А/В-тесті	Розширення на 100% трафіку, моніторинг метрик NDCG@10	100%	$\alpha_I=0,575$, $\lambda=0,06$
S_4	$P@10 \geq 0,14$ або $NDCG@10 \geq 0,17$	Оптимізація коефіцієнтів за результатами А/В, розширення каталогу	100%	$\alpha_E=0,5875$, $\lambda=0,08$
S_5	Стратегічне рішення РМ	Підвищення параметра забування, переорієнтація на нові сегменти	100%	$\alpha_E=0,75$, $\lambda=0,10$

Програмний пакет `recsys` спроектовано за принципом мікросервісної архітектури, що суттєво спрощує його інтеграцію з наявними компонентами IT-проекту. Взаємодія з рекомендаційним модулем здійснюється виключно через REST API, що унеможливорює пряму залежність між кодовою базою платформи та внутрішньою реалізацією РС, тобто команда може оновлювати або переналаштовувати модуль незалежно від основного релізного циклу.

Для інтеграції достатньо реалізувати три точки взаємодії. По-перше, відправлення поведінкових подій на ендпоінт `/api/v1/events` після кожної взаємодії користувача з платформою. По-друге, запит рекомендацій для конкретного користувача через `/api/v1/recommend/{user_id}`. По-третє, передачу управлінського сигналу про зміну стадії через `/api/v1/project/stage`. Усі три ендпоінти задокументовані автоматично через вбудований у FastAPI механізм OpenAPI Swagger, що спрощує роботу команди розробників.

Важливою рекомендацією є забезпечення асинхронної відправки подій із боку платформи, тобто виклик ендпоінту `/api/v1/events` не повинен блокувати основний потік обробки запиту користувача. Оскільки середній час інкрементного оновлення профілю становить менше 5 мс за результатами тестування продуктивності з рисунку 3.3, навантаження на РС-модуль є мінімальним навіть за умов тисяч подій на секунду на стадіях S_3 – S_4 .

Ефективне впровадження SABP-RS неможливе без організації системи моніторингу ключових показників якості. Рекомендується відстежувати щонайменше п'ять метрик у розрізі поточної стадії: Precision@10 , Recall@10 , NDCG@10 , MRR та Hit Rate@10 . Орієнтирами для оцінки є результати експериментального дослідження з таблиці 3.7, тобто досягнення значення $\text{NDCG@10} \geq 0,129$ на стадії S_2 та $\geq 0,176$ на стадіях S_3 – S_4 є індикатором коректної роботи модуля.

Окрім офлайн-метрик, важливо відстежувати онлайн-показники безпосередньо у продуктивному середовищі, тобто показник клікабельності

рекомендацій, конверсію з рекомендації у покупку та середній час відповіді системи. Відхилення часу відповіді понад 350 мс при повному перерахунку профілю є сигналом для оптимізації стратегії кешування або масштабування апаратних ресурсів.

Рішення про перехід між стадіями рекомендується приймати колективно за участі менеджера проєкту, технічного керівника та аналітика даних на підставі аналізу накопичених метрик. Зміна стадії є управлінською подією, і її доцільно фіксувати у реєстрі ризиків та журналі змін проєкту для забезпечення трасованості усіх рішень.

Практика впровадження рекомендаційних систем свідчить про ряд типових ризиків, до яких необхідно бути готовим заздалегідь. Першим і найпоширенішим ризиком є проблема «холодного старту», яка пом'якшується завдяки стадійно-адаптивним коефіцієнтам SABP-RS, тобто на стадіях S_1 – S_2 , коли явних оцінок немає, система автоматично максимізує вплив неявних сигналів, що дозволяє формувати осмислені рекомендації навіть на основі мінімальних поведінкових даних.

Другим важливим ризиком є деградація якості рекомендацій при зміні поведінки користувачів або реструктуризації каталогу, що мінімізується завдяки механізму стадійного забування, так як параметр $\lambda(\sigma)$ дозволяє гнучко налаштовувати швидкість знецінення застарілих сигналів без необхідності повного перенавчання моделі, що значно підвищує адаптивність системи та знижує обчислювальні витрати.

Третім ризиком є втрата накопичених профілів при перезапуску Redis, що усувається активацією AOF-режиму персистентності на рівні конфігурації Redis. У таблиці 4.2 підсумовано усі можливі ризики, а також наведено відповідні заходи щодо їх мінімізації, що забезпечує цілісне уявлення про управління ризиками в рамках запропонованого підходу.

Таблиця 4.2 – Типові ризики впровадження SABP-RS та рекомендовані заходи з їх мінімізації

Ризик	Ймовірний прояв	Захід з мінімізації
Холодний старт	Нерелевантні рекомендації на S_1-S_2	Використання $\alpha_l(S_l) = 0,90$ для переважання неявних сигналів
Застарілі профілі	Зниження точності рекомендацій при зміні поведінки	Збільшення $\lambda(\sigma)$ при переході до S_4-S_5 , перевірка HR@10
Втрата даних Redis	Скидання профілів та стадії $\sigma(t)$ при перезапуску	Активація AOF-режиму, резервне копіювання Redis RDB
Перевищення часу відповіді	Затримка >350 мс при перерахунку профілю	Оптимізація стратегії кешування, горизонтальне масштабування
Неправильний вибір стадії	Передчасний перехід – погіршення якості рекомендацій	Прийняття рішення на основі порогових значень метрик таблиці 4.1

Впровадження рекомендаційної системи на основі методу SABP-RS у ІТ-проект електронної комерції є керованим і поетапним процесом, що дозволяє досягти якісних рекомендацій вже на початкових стадіях розвитку модуля. Покроковий підхід до зміни стадій, чіткі критерії переходу, мікросервісна архітектура та вбудований механізм моніторингу забезпечують контрольоване впровадження з мінімальними ризиками для основних бізнес-процесів платформи.

4.2 Оцінювання ефекту від впровадження методу персоналізації в контексті проєктного управління

Завершальним етапом дослідження є комплексне оцінювання ефекту від впровадження гібридного методу стадійно-адаптивної персоналізації SABP-RS у розрізі не лише технічних метрик якості рекомендацій, а й показників управління

ІТ-проектом. Такий підхід дозволяє розглянути впровадження РС як повноцінний управлінський захід із вимірюваним ефектом для бізнесу та проектною командою.

Оцінювання здійснюється за трьома взаємопов'язаними вимірами: бізнес-ефектом від персоналізації для платформи електронної комерції, ефектом для процесів управління ІТ-проектом та обмеженнями, що визначають умови застосовності методу.

Основним вимірюваним результатом впровадження SABP-RS є підвищення якості рекомендацій, яке безпосередньо транслюється у бізнес-показники платформи. Результати експериментального дослідження з рисунку 3.1 демонструють, що запропонований метод перевершує найближчий аналог у вигляді статичного гібридного підходу за всіма п'ятьма метриками: Precision@10 зростає на 24,6%, Recall@10 – на 31,2%, NDCG@10 – на 23,1%, MRR – на 22,5%, Hit Rate@10 – на 14,6%.

Приріст метрики Hit Rate@10 на 14,6% означає, що частка користувачів, яким система пропонує хоча б один релевантний товар серед перших десяти рекомендацій, збільшується з 48,7% до 55,8%. В умовах платформи електронної комерції це безпосередньо корелює зі зростанням клікабельності рекомендаційного блоку та конверсії з рекомендації у покупку. Галузеві дослідження підтверджують, що підвищення клікабельності рекомендацій на 10-15% здатне забезпечити приріст виручки платформи в межах 5–12% залежно від частки рекомендаційного трафіку в загальному обсязі продажів.

Принципово важливим для бізнесу є ефект на ранніх стадіях впровадження. Як показує таблиця 3.7, на стадії S_1 приріст NDCG@10 відносно статичного гібриду становить 47,3%, а на S_2 – 33,0%, що означає, що SABP-RS забезпечує якісні рекомендації вже від перших тижнів роботи модуля, не вимагаючи накопичення великої бази явних оцінок. Для бізнесу це скорочує «мертвий» період після запуску рекомендаційного функціоналу – час, протягом якого система вже розгорнута, але ще не приносить вимірюваного ефекту.

Додатковим бізнес-ефектом є покращення показника утримання користувачів. Системи з персоналізованими рекомендаціями демонструють вищу залученість користувачів: персоналізований контент збільшує середню тривалість сесії та кількість переглядів товарів, що особливо актуально для стадій S_3 – S_4 , коли профілі користувачів є зрілими, а значення $\text{Hit Rate}@10 = 0,558$ – $0,613$ означає, що більше половини користувачів знаходять релевантні товари в рекомендаційному блоці при кожному відвідуванні платформи.

Таблиця 4.3 – Очікуваний бізнес-ефект від впровадження SABP-RS за усіма п'ятьма стадіями

Стадія	Hit Rate@10 SABP-RS	Приріст NDCG@10 до Static Hybrid	Прогнозований приріст CTR	Ключовий бізнес-ефект
S_1	0,348	+47,3%	помірний (тест. трафік)	Усунення ефекту «порожніх» рекомендацій при запуску
S_2	0,457	+33,0%	+8–12% у А/В-групі	Перші вимірювані конверсії з рекомендацій
S_3	0,567	+18,9%	+12–18%	Зростання виручки, підвищення середнього чека
S_4	0,613	+9,5%	+15–20%	Максимальна ефективність, зниження операційних витрат
S_5	0,573	+10,7%	+12–16%	Утримання релевантності при зміні каталогу

Впровадження SABP-RS впливає на управління IT-проєктом у кількох аспектах, що виходять за межі суто технічних показників якості. Першим і найбільш значущим аспектом є зменшення операційних витрат на підтримку системи рекомендацій. На відміну від методів машинного навчання, що вимагають повного перенавчання моделі при зміні розподілу вхідних даних, SABP-RS адаптується до нової стадії шляхом простої зміни одного параметра $\sigma(t)$ через REST-ендпоінт. Операція займає менше 2 мс і не потребує зупинки сервісу, тобто є нульовим простоем.

Другий аспект стосується планування ресурсів проєкту. Стадійна модель S_1 – S_5 дозволяє менеджеру проєкту чітко визначати контрольні точки у розвитку РС-модуля і прив'язувати до них виділення обчислювальних ресурсів. Зокрема, на стадіях S_1 – S_2 достатньо мінімальної конфігурації, тоді як повноцінне промислове навантаження у 50 тис. активних користувачів і 100 тис. товарів доцільно закладати лише починаючи з S_3 , що дозволяє оптимізувати витрати на інфраструктуру і уникнути передчасного виділення ресурсів, що не використовуються.

Третім аспектом є прозорість і трасованість рішень. Усі ключові параметри методу, тобто $\alpha_E(\sigma)$, $\alpha_I(\sigma)$ і $\lambda(\sigma)$ визначаються через детерміновані лінійні формули (3.1)–(3.3), а не підбираються емпірично. Таким чином, менеджер проєкту може у будь-який момент пояснити, чому система формує певні рекомендації, і які параметри впливають на їх якість. Така прозорість відповідає принципам управління проєктами за РМВОК 8 [4] щодо документованості та обґрунтованості проєктних рішень.

Четвертий аспект стосується управління ризиками. Стадійна модель дозволяє локалізувати ризики у часі: якщо на стадії S_2 результати А/В-тестування не задовольняють порогових значень таблиці 4.1, команда може залишатися на стадії S_2 без переходу до S_3 , що унеможливлює вплив неякісних рекомендацій на весь трафік платформи, що принципово відрізняє SABP-RS від підходів, де

система або «запущена», або «не запущена», без проміжних контрольованих станів. У таблиці 4.4 відображено вплив розробленого модуля на всі області управління ІТ-проєктом ЕК.

Таблиця 4.4 – Вплив впровадження SABP-RS на процеси управління ІТ-проєктом

Область управління	Традиційний підхід (статична РС)	Підхід SABP-RS
Введення в дію	Єдина дата запуску, до неї система не працює	Поетапний запуск стадій, корисний ефект з перших тижнів
Адаптація до змін	Повне перенавчання моделі (годин-днів, сервіс на паузі)	Зміна $\sigma(t)$ через API за < 2 мс, без зупинки сервісу
Планування ресурсів	Пікові ресурси з першого дня	Поступове масштабування відповідно до стадії
Управління ризиками	Бінарний ризик: система або коректна, або ні	Стадійна ізоляція ризику, тестовий трафік на S_2
Документованість	Параметри підбираються емпірично, важко пояснити	Детерміновані формули (3.1)–(3.3), повна трасованість
Моніторинг якості	Єдиний набір метрик без урахування фази проєкту	Стадійні орієнтири з таблиці 4.1, динамічний контроль

Оцінювання ефективності впровадження неможливе без урахування витрат, пов'язаних із розгортанням та підтримкою системи. Порівняємо три підходи: класичну колаборативну фільтрацію ItemCF, сучасний метод BPR та розроблений SABP-RS – за сукупністю показників якості та операційних витрат.

Метод BPR демонструє близькі до SABP-RS значення метрик якості рекомендацій на зрілих стадіях S_3 – S_5 , однак має суттєві недоліки з

проектно-управлінської точки зору. По-перше, BPR потребує повного перенавчання при зміні розподілу даних, що є значною операційною витратою. По-друге, метод не має механізму адаптації до ранніх стадій проекту, коли явних оцінок ще немає: на стадії S_I результати BPR суттєво поступаються SABP-RS. По-третє, навчання на повному датасеті займає значно більше часу порівняно з інкрементним оновленням профілів SABP-RS.

SABP-RS натомість поєднує конкурентну якість рекомендацій з мінімальними операційними витратами: відсутністю необхідності перенавчання, інкрементним оновленням профілів за $O(1)$ на подію та часом адаптації до зміни стадії менш ніж 2 мс, що забезпечує вищу питому ефективність, тобто співвідношення якості рекомендацій до витрат на їх підтримку, порівняно з усіма розглянутими базовими методами. У таблиці 4.5 наведено порівняння методів за перерахованими критеріями.

Таблиця 4.5 – Порівняльна оцінка методів рекомендацій за сукупністю критеріїв

Критерій оцінювання	ItemCF	Static Hybrid	BPR	SABP-RS
Precision@10 (усер. S_I-S_5)	0,097	0,118	0,124	0,147
NDCG@10 (усер. S_I-S_5)	0,118	0,143	0,149	0,176
Ефективність на S_I (NDCG@10)	низька	0,055	низька	0,081
Час адаптації до змін	перерахунок	перерахунок	перенавчання	< 2 мс
Потреба у перенавчанні	так	так	так	ні

Кінець таблиці 4.5

Критерій оцінювання	ItemCF	Static Hybrid	BPR	SABP-RS
Трасованість параметрів	часткова	часткова	низька	повна
Стадійна адаптивність	відсутня	відсутня	відсутня	вбудована
Загальна оцінка (1–5)	2	3	3	5

Об’єктивне оцінювання ефекту від впровадження вимагає чіткого визначення меж застосовності запропонованого методу. SABP-RS демонструє найвищу ефективність у контексті ІТ-проектів ЕК з динамічним каталогом, гетерогенними поведінковими даними та чітко вираженим проєктним циклом РС-модуля. Водночас існують умови, за яких застосування методу є менш доцільним або потребує додаткового налаштування.

По-перше, метод розрахований на платформи з каталогом від 1 тис. товарів та аудиторією від 500 активних користувачів, так як при менших значеннях матриця взаємодій є надмірно розрідженою, і колаборативна складова HybridRecommender не формує достатньої кількості сусідів-кандидатів для якісного ранжування.

По-друге, ефективна робота механізму стадійного забування передбачає регулярне надходження нових поведінкових подій, тобто за умов рідкісного відвідування платформи, тобто приблизно менше кількох тисяч подій на тиждень, параметр $\lambda(\sigma)$ може знецінювати частину корисної інформації разом із застарілими сигналами.

По-третє, детерміновані лінійні апроксимації коефіцієнтів (3.1)–(3.3) є оптимальними для типових сценаріїв ЕК, однак для специфічних ніш, наприклад, платформ продажу товарів тривалого попиту, де користувачі здійснюють покупки раз на рік, граничні значення α_E та λ можуть потребувати індивідуального

калібрування, але це не є обмеженням архітектури методу, а є природною вимогою до параметричного налаштування будь-якої рекомендаційної системи.

По-четверте, запропонований метод орієнтований на користувацьку персоналізацію на основі поведінкових даних і не включає компонентів контентної фільтрації на рівні атрибутів товарів. Для платформ із дуже великою часткою нових товарів, у яких понад 30% каталогу оновлюється щомісяця, рекомендується розширити архітектуру РС за рахунок додаткового CBF-модуля, що опрацьовує атрибути нових позицій до накопичення поведінкових даних про них.

Проведений аналіз свідчить, що впровадження гібридного методу SABP-RS забезпечує комплексний ефект, який виявляється на технічному, бізнес-орієнтованому та проєктно-управлінському рівнях.

На технічному рівні метод перевищує всі розглянуті базові підходи за метриками якості рекомендацій, демонструючи стабільну перевагу на всіх стадіях S_1 – S_5 і особливо значущий приріст на ранніх стадіях.

На бізнес-рівні підвищення Hit Rate@10 з 48,7% до 55,8% та зростання Recall@10 на 31,2% транслиуються у прогнозований приріст коефіцієнту клікабельності рекомендаційного блоку в межах 8–20% залежно від стадії, що за оцінками галузевих досліджень відповідає приросту виручки платформи в діапазоні 5–12%. Скорочення «мертвого» періоду після запуску РС-модуля знижує час до першого вимірюваного бізнес-ефекту від інвестицій у персоналізацію.

На рівні проєктного управління розроблена стадійна модель підвищує передбачуваність впровадження, забезпечує можливість поетапного масштабування ресурсів, спрощує управління ризиками та гарантує повну трасованість рішень. Усе перераховане є критичними властивостями для зрілих ІТ-проєктів із вимогами до документованості та обґрунтованості технічних рішень за РМВОК 8 [4].

Результати, отримані в ході виконання даної кваліфікаційної роботи, формують цілісний підхід до стадійно-адаптивної персоналізації рекомендацій у

ІТ-проектах ЕК. Водночас проведене дослідження виявило низку напрямів, розвиток яких дозволить суттєво розширити можливості та практичну застосовність запропонованого методу SABP-RS.

По-перше, доцільною є інтеграція трансформерних архітектур глибокого навчання для виявлення послідовних патернів поведінки користувачів, що особливо актуально на зрілих стадіях S_3 – S_4 .

По-друге, перспективним є розширення набору поведінкових факторів за рахунок контекстуальних даних: типу пристрою, геолокації та сезонності.

По-третє, потребує дослідження масштабування методу до промислового навантаження з використанням технологій потокової обробки даних.

По-четверте, актуальним залишається вирішення проблеми холодного старту нових товарів на основі їх контентних атрибутів.

Реалізація зазначених напрямів дозволить суттєво розширити практичну застосовність розробленого підходу в різноманітних ІТ-проектах ЕК.

ВИСНОВКИ

У результаті виконання кваліфікаційної роботи проведено комплексне дослідження методів персоналізації рекомендацій на основі поведінкових даних користувачів в ІТ-проектах ЕК, що дозволило досягти поставленої мети та вирішити визначені задачі. У роботі здійснено аналіз предметної області ЕК з позиції ІТ-проектів, визначено роль рекомендаційних систем як одного з ключових інструментів підвищення ефективності цифрових платформ, а також досліджено особливості формування та використання поведінкових даних користувачів як основи персоналізації. Проведений огляд сучасних методів побудови рекомендаційних систем показав, що існуючі підходи мають як суттєві переваги, так і обмеження, пов'язані з проблемами холодного старту, розрідженістю даних, обмеженістю рекомендацій або складністю реалізації, що обумовлює доцільність використання гібридних підходів.

Основним результатом роботи є розробка гібридного методу персоналізації рекомендацій, який базується на формуванні поведінкового профілю користувача та поєднує різні підходи до обробки даних. Запропонований метод враховує як явні, так і неявні поведінкові сигнали, а також особливості їх використання на різних етапах життєвого циклу ІТ-проекту. На відміну від існуючих рішень, він дозволяє більш гнучко адаптувати рекомендаційний механізм до динамічних змін поведінки користувачів та умов функціонування системи. У межах роботи розроблено модель персоналізації рекомендацій, яка інтегрується в архітектуру ІТ-проекту ЕК та забезпечує формування релевантних пропозицій на основі аналізу накопичених даних.

Проведене експериментальне дослідження та апробація запропонованого методу підтвердили його ефективність, зокрема за рахунок підвищення релевантності рекомендацій та покращення показників взаємодії користувачів із

системою. Отримані результати відповідають сучасному рівню наукових і технічних знань у сфері РС та узгоджуються з актуальними тенденціями їх розвитку, які передбачають використання гібридних моделей і методів аналізу великих даних.

Практичне значення роботи полягає в можливості застосування розробленого методу в ІТ-проєктах ЕК, зокрема в системах, орієнтованих на підвищення якості обслуговування користувачів, зростання конверсії та оптимізацію бізнес-процесів. Результати дослідження можуть бути використані при проєктуванні та модернізації РС у веб- та мобільних платформах, а також у рамках систем підтримки прийняття рішень. Ступінь впровадження результатів можна оцінити як такий, що передбачає можливість їх інтеграції у реальні ІТ-проєкти після відповідної адаптації до конкретних умов і вимог.

Науково-технічна значущість роботи полягає в розвитку підходів до персоналізації рекомендацій за рахунок урахування поведінкових даних у поєднанні з контекстом ІТ-проєктного середовища. Соціально-економічна значущість проявляється у підвищенні якості користувацького досвіду, більш ефективному задоволенні потреб споживачів та зростанні ефективності електронної комерції в цілому.

Подальші дослідження доцільно спрямувати на вдосконалення запропонованого методу шляхом використання більш складних моделей машинного та глибокого навчання, розширення набору врахованих поведінкових факторів, а також дослідження питань масштабування та обробки даних у реальному часі. Графічний матеріал даної роботи надано в додатку А.

Роботу виконано за допомогою методичних вказівок до науково-дослідної практики для здобувачів вищої освіти усіх форм навчання другого (магістерського) рівня вищої освіти спеціальності 122 «Комп'ютерні науки» освітньо-наукової програми «Управління проєктами в галузі інформаційних технологій» [39] згідно з вимогами державного стандарту України 3008:2015 [40].

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Синьова В.О., Білова Т.Г. Аналіз методів персоналізації контенту в ІТ-проектах електронної комерції на базі поведінкових патернів користувачів // 29-й Міжнародний молодіжний форум «Радіoeлектроніка та молодь у ХХІ столітті». Зб. матеріалів форуму. Т. 6. Харків: ХНУРЕ. 2025. С. 202–204.
2. Синьова В.О., Білова Т.Г. Гібридний метод персоналізації рекомендацій на основі поведінкових даних користувачів // 30-й Міжнародний молодіжний форум «Радіoeлектроніка та молодь у ХХІ столітті». Зб. матеріалів форуму. Т. 6. Харків: ХНУРЕ. 2026. С. 231–233.
3. Laudon K. C., Traver C. G. E-Commerce 2020-2021: Business, Technology and Society. 16th ed. Pearson, 2020. 912 p.
4. A Guide to the Project Management Body of Knowledge (PMBOK Guide). 8th ed. Project Management Institute, 2026. 408 p.
5. Aliev R. Recommender Algorithms in 2026: A Practitioner's Guide: Structured and practical overview of this algorithmic landscape. Mathematical Foundations and code samples. Independently published, 2025. 403 p.
6. Abtahi A. T. Exploring Consumer Preferences: The Significance Of Personalization In E-Commerce. Malaysian E Commerce Journal. 2023. Vol. 8, No. 1. P. 1–7. DOI: 10.26480/mecj.01.2024.01.07.
7. Chalyi S., Leshchynskyi V., Leshchynska I. Multilevel personalization of explanations in recommender systems. Advanced Information Systems. 2020. Vol. 4, No. 2. P. 170–175. DOI: 10.20998/2522-9052.2020.2.25.
8. Ricci F., Rokach L., Shapira B. Recommender Systems Handbook. 3rd ed. New York: Springer, 2022. 1071 p.
9. Модель та метод оцінки та вибору інформаційної системи управління проектами / В. М. Левикін, О. В. Петриченко, І. О. Юр'єв, М. І. Купенко // АСУ та

прилади автоматики : всеукр. міжвід. наук.-техн. зб. – Харків : ХНУРЕ, 2024. – Вип. 181. – С. 25–41. – DOI: 10.30837/0135-1710.2024.181.025.

10. Latha Y. M., Rao B. S. Amazon product recommendation system based on a modified convolutional neural network. *ETRI Journal*. 2024. Vol. 46, Iss. 4. P. 633–647. DOI: 10.4218/etrij.2023-0162.

11. Deep Learning for Recommender Systems: A Netflix Case Study / H. Steck et al. *AI Magazine*. 2021. Vol. 42, No. 3. P. 7–18. DOI: 10.1609/aimag.v42i3.18140.

12. Zhang Y., Chen X. Explainable Recommendation: A Survey and New Perspectives. *Foundations and Trends in Information Retrieval*. 2020. Vol. 14, No. 1. P. 1–101. DOI: 10.1561/15000000066.

13. Chalyi S., Leshchynskyi V. Construction of patterns of user preferences dynamics for explanations in the recommender system. *Advanced Information Systems*. 2021. Vol. 5, No. 1. P. 107–112. DOI: 10.20998/2522-9052.2021.1.15.

14. Glisovic N., Kragic D., Tegner M. Item Cold Start in E-Commerce Recommender Systems: A Survey. *IEEE Access*. 2025. Vol. 13. P. 164702–164722. DOI: 10.1109/ACCESS.2025.3611026.

15. Про захист персональних даних : Закон України від 01.06.2010 р. № 2297-VI. Дата оновлення: 12.02.2025. URL: <https://zakon.rada.gov.ua/laws/show/2297-17> (дата звернення: 10.02.2026).

16. Петров К. Е., Кобзев І. В. Метод визначення переважності споживчих характеристик товару на основі інформації про його вибір покупцем. *Системи управління, навігації та зв'язку*. 2022. № 2 (68). С. 68–74. DOI: 10.26906/SUNZ.2022.2.068.

17. Recent Developments in Recommender Systems: A Survey [Review Article] / Y. Li et al. *IEEE Computational Intelligence Magazine*. 2024. Vol. 19, no. 2. P. 78–95. DOI: 10.1109/MCI.2024.3363984.

18. Чала О. В., Євдокімов Б. С. Розробка методу уточнення рекомендацій з використанням темпоральних знань в задачах індивідуального страхування. *АСУ*

та прилади автоматики. 2024. № 1 (181). С. 17–24. DOI: 10.30837/0135-1710.2024.181.017.

19. Recommender System: A Comprehensive Overview of Technical Challenges and Social Implications / Y. An et al. ICCCK Transactions on Sensing, Communication, and Control. 2024. Vol. 1, no. 1. P. 30–51. DOI: 10.62762/TSCC.2024.898503.

20. Recommender Systems / C. Lucchese et al. Wirtschaftsinformatik. 2021. Vol. 39. P. 401–404. DOI: 10.1007/978-0-387-30164-8_705.

21. Recommender Systems with Generative Retrieval / S. Rajput et al. ArXiv. 2023. abs/2305.05065. DOI: 10.48550/arXiv.2305.05065.

22. Content-Based Collaborative Generation for Recommender Systems / Y. Wang et al. Proceedings of the 33rd ACM International Conference on Information and Knowledge Management. 2024. DOI: 10.1145/3627673.3679692.

23. Understanding customer loyalty-aware recommender systems in E-commerce: an analytical perspective / R. Esmeli et al. Electronic Commerce Research. 2025. DOI: 10.1007/s10660-025-09954-6.

24. Nam L. N. A Robust Approach for Hybrid Personalized Recommender Systems. International Conference on Theory and Practice of Digital Libraries. 2023. DOI: 10.1007/978-3-031-43849-3_14.

25. Bootstrapped Personalized Popularity for Cold Start Recommender Systems / I. Chaimalas et al. Proceedings of the 17th ACM Conference on Recommender Systems. 2023. DOI: 10.1145/3604915.3608820.

26. Recommender Systems in E-Commerce: A Deep Dive into AI-Driven Optimization / M. K. Mishra et al. 2025 First International Conference on Advances in Computer Science, Electrical, Electronics, and Communication Technologies (CE2CT). 2025. P. 1392–1398. DOI: 10.1109/CE2CT64011.2025.10939560.

27. Collaborative filtering recommender systems taxonomy / H. Papadakis et al. Knowledge and Information Systems. 2022. Vol. 64. P. 35–74. DOI: 10.1007/s10115-021-01628-7.

28. Recommender Systems in E-Commerce: Algorithmic Taxonomy, Comparative Evaluation, and a Framework for Next-Generation Personalization / H. Bachri et al. 2025 4th International Conference on Embedded Systems and Artificial Intelligence (ESAI). 2025. P. 1–8. DOI: 10.1109/ESAI67033.2025.11438290.

29. Wu Z. Optimizing E-commerce Recommender Systems: A Comprehensive Review of Techniques and Future Directions. *Applied and Computational Engineering*. 2024. DOI: 10.54254/2755-2721/97/20241175.

30. Nasir M., Ezeife C. I. A Survey and Taxonomy of Sequential Recommender Systems for E-commerce Product Recommendation. *SN Computer Science*. 2023. Vol. 4. DOI: 10.1007/s42979-023-02166-5.

31. Hariyale I., Raghuwanshi D. M. Advanced Hybrid Recommender Systems: Enhancing Precision and Addressing Cold-Start Challenges in E-Commerce and Content Streaming. *International Journal of Computing and Digital Systems*. 2024. Vol. 16, no. 1. DOI: 10.12785/ijcds/1601129.

32. Tang H., Zhong J., Liu G. Deep learning techniques in e-commerce recommender systems and their impact on business marketing strategies. *Proceedings of the 5th International Conference on E-Commerce and Internet Technology (ECIT)*. 2024. DOI: 10.4108/eai.15-3-2024.2346153.

33. Albalawi A., Alharbi S. E-Commerce Recommender Systems: A Comprehensive Survey. 2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA). 2023. P. 503–509. DOI: 10.1109/ICIRCA57980.2023.10220872.

34. Recommender Systems in E-Commerce / L. Mohanty et al. 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT). 2022. P. 114–119. DOI: 10.1109/CCiCT56684.2022.00032.

35. Decomposed Collaborative Filtering: Modeling Explicit and Implicit Factors For Recommender Systems / H. Chen et al. *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*. 2021. DOI: 10.1145/3437963.3441826.

36. Shrivastava V., Kumar S. Comparative Analysis of Functionality and Aspects for Hybrid Recommender Systems. *International Journal on Recent and Innovation Trends in Computing and Communication*. 2023. Vol. 11, no. 8s. DOI: 10.17762/ijritcc.v11i8s.7236.

37. Unified Representation Learning for Multi-Intent Diversity and Behavioral Uncertainty in Recommender Systems / W. Xu et al. 2025 10th International Conference on Computer and Information Processing Technology (ISCIPT). 2025. P. 114–117. DOI: 10.1109/ISCIPT67144.2025.11265095.

38. Torkashvand A., Jameii S. M., Reza A. Deep learning-based collaborative filtering recommender systems: a comprehensive and systematic review. *Neural Computing and Applications*. 2023. Vol. 35. P. 24783–24827. DOI: 10.1007/s00521-023-08958-3.

39. Методичні вказівки до науково-дослідної практики для здобувачів вищої освіти усіх форм навчання другого (магістерського) рівня вищої освіти спеціальності 122 Комп'ютерні науки освітньо-наукової програми «Управління проектами в галузі інформаційних технологій». / Упоряд.: К.Е. Петров, Євланов М.В., Цвіркун О.А.. – Харків: ХНУРЕ, 2025. – 19 с.

40. Інформація та документація. Звіти у сфері науки і техніки. Структура та правила оформлювання: ДСТУ 3008:2015 / В. Земцева, Ю. Поліщук, Р. Санченко та ін.; ред. І. Дьячкова; Київ, 2016. 26 с.