

Міністерство освіти і науки України  
Університет митної справи та фінансів

Факультет інноваційних технологій  
Кафедра комп'ютерних наук та інженерії програмного забезпечення

## Кваліфікаційна робота магістра

на тему: «Розроблення рекомендаційної системи для онлайн-платформ електронної комерції»

Виконав: студент групи К23-2М

Спеціальність 122 Комп'ютерні науки

Соболь А.Б.

(прізвище та ініціали)

Керівник д.е.н., проф. Корнєєв М. В.

(науковий ступінь, вчене звання, прізвище та ініціали)

Рецензент Дніпровський державний

технічний університет

(місце роботи)

доцент кафедри математичного

моделювання та системного аналізу

(посада)

к.т.н., доц. Волосова Н.М.

(науковий ступінь, вчене звання, прізвище та ініціали)

Дніпро – 2025

## АНОТАЦІЯ

Соболь А.Б. Розроблення рекомендаційної системи для онлайн-платформ електронної комерції.

Дипломна робота на здобуття освітнього ступеня магістр за спеціальністю 122 «Комп'ютерні науки» – Університет митної справи та фінансів, Дніпро, 2025.

Магістерська робота присвячена дослідженню та розробці рекомендаційної системи для онлайн-платформ електронної комерції, що є актуальним завданням у сучасних умовах стрімкого розвитку цифрової економіки. У роботі розглянуто теоретичні основи електронної комерції, роль і значення онлайн-платформ у глобальній економіці, а також їхній вплив на споживчу поведінку. Особливу увагу приділено аналізу існуючих підходів до побудови рекомендаційних систем, серед яких виділено методи контентної, колаборативної та гібридної фільтрації.

У межах роботи детально досліджено проблеми, що виникають під час розроблення та впровадження рекомендаційних систем, зокрема проблему «холодного старту», розрідженості даних і перевантаження інформацією. Запропоновано алгоритмічні та архітектурні рішення для подолання цих викликів, зокрема застосування передових методів машинного навчання, аналізу великих даних і використання гібридних моделей.

Практична частина роботи включає розроблення прототипу рекомендаційної системи для онлайн-платформи електронної комерції. Розроблена система орієнтована на персоналізацію рекомендацій для користувачів на основі їхніх попередніх дій, інтересів та поведінкових патернів.

Ключові слова: електронна комерція, рекомендаційні системи, машинне навчання, персоналізація, колаборативна фільтрація, алгоритми, великі дані, онлайн-платформи.

## ABSTRACT

Sobol A.B, Development of a recommendation system for online e-commerce platforms.

Diploma thesis for obtaining a master's degree in specialty 122 «Computer Science» – University of Customs and Finance, Dnipro, 2025.

The master's thesis is devoted to the research and development of a recommendation system for online e-commerce platforms, which is an urgent task in the current conditions of rapid development of the digital economy. The thesis examines the theoretical foundations of e-commerce, the role and importance of online platforms in the global economy, and their impact on consumer behavior. Particular attention is paid to the analysis of existing approaches to building recommender systems, including content, collaborative and hybrid filtering methods.

The paper investigates in detail the problems that arise during the development and implementation of recommender systems, in particular, the problem of «cold start», data sparsity and information overload. Algorithmic and architectural solutions to overcome these challenges are proposed, including the use of advanced machine learning methods, big data analysis, and the use of hybrid models.

The practical part of the work includes the development of a prototype recommendation system for an online e-commerce platform. The developed system is focused on personalizing recommendations for users based on their previous actions, interests, and behavioral patterns.

Keywords: e-commerce, recommender systems, machine learning, personalization, collaborative filtering, algorithms, big data, online platforms.

## ЗМІСТ

ВСТУП .....	5
РОЗДІЛ 1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕННЯ .....	7
1.1 Сутність електронної комерції .....	7
1.2 Типи онлайн-платформ електронної комерції .....	10
1.3 Роль онлайн-платформ у сучасній економіці.....	14
1.4 Аналіз сучасної літератури .....	18
1.5 Висновки до першого розділу .....	26
РОЗДІЛ 2. ДОСЛІДЖЕННЯ МЕТОДІВ РОЗРОБЛЕННЯ РЕКОМЕНДАЦІЙНОЇ СИСТЕМИ .....	28
2.1 Визначення та функції рекомендаційних систем .....	28
2.2 Ключові завдання та принципи роботи рекомендаційних систем.....	31
2.3 Системи на основі фільтрації за змістом.....	33
2.4 Системи на основі колаборативної фільтрації.....	37
2.5 Гібридні рекомендаційні системи .....	41
2.6 Алгоритми машинного навчання в рекомендаційних системах .....	45
2.7 Висновки до другого розділу.....	49
РОЗДІЛ 3. ПРОГРАМНА РЕАЛІЗАЦІЯ.....	51
3.1 Мета розроблення .....	51
3.2 Архітектура системи.....	52
3.3 Опис використаних технологій .....	55
3.4 Процес навчання та представлення рекомендацій .....	58
3.5 Інтеграція з іншими інструментами.....	61
3.6 Аналіз отриманих результатів .....	65
3.7 Висновки до третього розділу .....	71
ВИСНОВКИ.....	72
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	74
ДОДАТКИ.....	77

## ВСТУП

Сучасний етап розвитку інформаційних технологій, зокрема в сфері електронної комерції, характеризується стрімким зростанням попиту на інноваційні рішення, що сприяють оптимізації бізнес-процесів та підвищенню рівня задоволення потреб споживачів. В умовах інтенсивної конкуренції та швидких змін на ринку, важливу роль у забезпеченні конкурентоспроможності онлайн-платформ займають системи, які допомагають користувачам здійснювати вибір серед великої кількості товарів або послуг [1, 2]. Рекомендаційні системи стали невід'ємною частиною таких платформ, оскільки вони дозволяють персоналізувати досвід користувачів, що в свою чергу сприяє підвищенню ефективності продажів, зменшенню часу на прийняття рішення та покращенню взаємодії з клієнтами. Таким чином, розробка та вдосконалення рекомендаційних систем є надзвичайно актуальними для сучасної електронної комерції.

Актуальність цієї теми зумовлена кількома факторами. По-перше, із зростанням обсягів даних, що накопичуються в результаті діяльності електронних комерційних платформ, виникає необхідність у створенні ефективних алгоритмів для обробки та аналізу великих масивів інформації. По-друге, для забезпечення конкурентних переваг платформи повинні пропонувати користувачам інтуїтивно зрозумілу і персоналізовану рекомендацію товарів, що відповідають їх інтересам та вподобанням [2]. Це дозволяє знижувати кількість відмов від покупок і підвищувати рівень задоволення клієнтів, що в кінцевому результаті збільшує прибутковість підприємства. Рекомендаційні системи, які використовують передові методи машинного навчання та аналізу даних, дозволяють досягти високого рівня персоналізації, що є ключовим фактором у сучасному бізнес-середовищі.

Метою цієї кваліфікаційної роботи є розробка рекомендаційної системи для онлайн-платформи електронної комерції, яка забезпечуватиме

персоналізовані рекомендації для користувачів на основі їх попередніх дій, інтересів та поведінкових моделей. Для досягнення цієї мети передбачається виконати наступні завдання:

- проаналізувати існуючі підходи до створення рекомендаційних систем та вибрати найбільш оптимальні методи для розробки системи, орієнтованої на потреби електронної комерції;

- розробити прототип рекомендаційної системи для онлайн-платформи електронної комерції;

- визначити шляхи вдосконалення рекомендаційних алгоритмів та оцінити їх потенціал для інтеграції в майбутні платформи.

Об'єктом дослідження є процеси, що відбуваються на онлайн-платформах електронної комерції при взаємодії користувачів з товарами або послугами, а предметом дослідження – методи та алгоритми побудови рекомендаційних систем, їх ефективність та вплив на поведінку користувачів.

Методи дослідження, що застосовуються в роботі, включають аналіз наукової та технічної літератури, моделювання та розробку алгоритмів для персоналізації рекомендацій, а також методи тестування та оцінки ефективності програмного забезпечення. Практична значимість роботи полягає в розробці функціональної рекомендаційної системи, яка може бути інтегрована в реальні онлайн-платформи електронної комерції для покращення взаємодії з користувачами. Наукова новизна роботи полягає у використанні новітніх підходів до побудови рекомендаційних алгоритмів з урахуванням специфіки електронної комерції, а також у пропозиції методів для адаптації цих алгоритмів до різноманітних бізнес-моделей, що дозволяє досягти максимального ефекту при мінімальних витратах на впровадження.

Структура кваліфікаційної роботи. Робота складається з трьох розділів, обсяг роботи – 85 сторінок, робота містить 14 рисунків та 2 таблиці. Список використаних джерел має 18 посилань.

## РОЗДІЛ 1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕННЯ

### 1.1 Сутність електронної комерції

Електронна комерція є однією з найбільш динамічних і важливих складових сучасної економіки, що за останні десятиліття значно змінила бізнес-процеси, торгові стратегії та спосіб взаємодії споживачів з товарами та послугами. Вона охоплює всі форми торгових операцій, що здійснюються за допомогою інформаційних технологій, зокрема Інтернету [1]. Електронна комерція з'явилася як результат розвитку цифрових технологій та впровадження Інтернету в повсякденне життя, що дозволило значно зменшити часові та просторові обмеження для здійснення торгових операцій, водночас спрощуючи процеси купівлі-продажу товарів та послуг.

Схему електронної комерції наведено на рисунку 1.1.

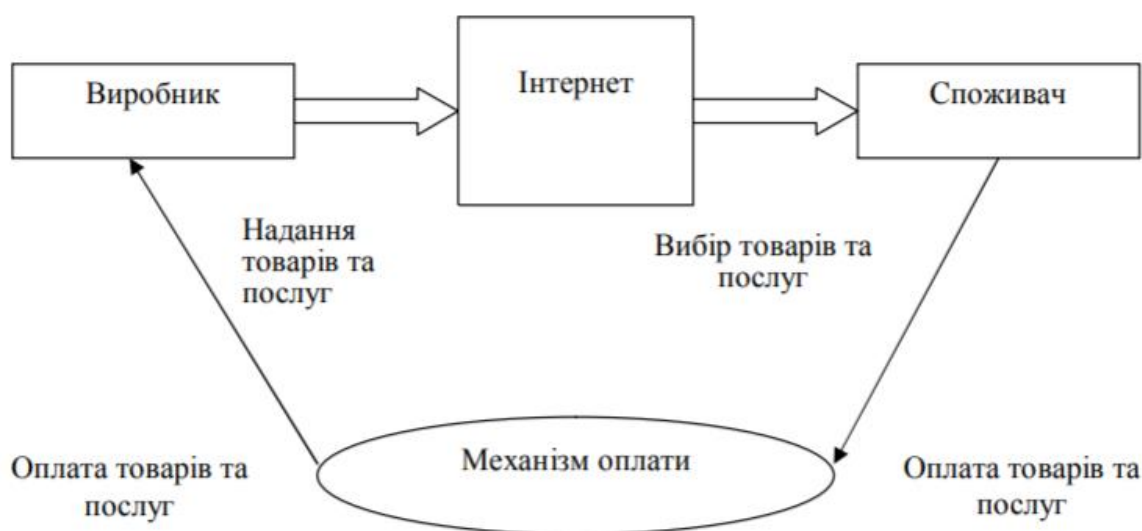


Рисунок 1.1 – Схема електронної комерції

У науковому контексті сутність електронної комерції розглядається через кілька важливих аспектів: технологічний, економічний, соціальний і правовий. З технологічного боку електронна комерція є процесом обміну товарами та послугами між постачальниками та споживачами, що здійснюється за допомогою електронних каналів зв'язку, таких як Інтернет, мобільні мережі, а також через використання інформаційних систем і платформ для обробки та передачі даних. Важливим компонентом цієї діяльності є онлайн-платформи та маркетплейси, які забезпечують середовище для реалізації товарів та послуг у цифровому просторі.

З економічної точки зору електронна комерція є важливим інструментом розвитку ринку, який надає нові можливості для підприємців та споживачів. Вона дає змогу компаніям знижувати витрати на організацію фізичних магазинів, а також розширювати географічні межі свого бізнесу, залучаючи клієнтів з різних куточків світу. Завдяки електронній комерції підприємства можуть значно скоротити витрати на логістику, маркетинг та рекламні кампанії, одночасно забезпечуючи персоналізований підхід до кожного споживача. Електронні магазини дозволяють реалізовувати більш точне сегментування ринку, застосовувати аналітику даних для оптимізації бізнес-процесів і підвищення ефективності.

Соціальний аспект електронної комерції полягає у зміні поведінки споживачів і їхньому ставленні до покупок. Онлайн-шопінг дозволяє користувачам здійснювати покупки без обмежень щодо часу і місця, що має великий вплив на рівень споживчої активності [2, 3]. Крім того, завдяки технологіям збирання і обробки даних, електронна комерція дозволяє формувати більш детальні портрети споживачів, що дає можливість розробляти індивідуалізовані пропозиції та рекламні акції. Це створює нові можливості для побудови довгострокових відносин між компаніями та їхніми клієнтами.



Правовий аспект електронної комерції є не менш важливим, оскільки віртуальні угоди і транзакції потребують особливих юридичних норм і регулювань. Відсутність фізичної присутності сторін на момент укладання угоди, а також специфіка електронних платежів і передачі даних вимагають розробки чітких стандартів захисту прав споживачів, а також механізмів вирішення суперечок, пов'язаних з електронними угодами. Однією з основних проблем є забезпечення безпеки інформації, яка передається між покупцями та продавцями, а також захист від шахрайства, яке може мати місце в Інтернет-просторі.

Незважаючи на очевидні переваги, електронна комерція не є бездоганною і має низку викликів, зокрема технічного та організаційного характеру. На технологічному рівні одним із основних викликів є забезпечення високого рівня безпеки даних і фінансових транзакцій. Оскільки електронна комерція передбачає використання інформаційних систем для здійснення операцій, існує потреба у постійному вдосконаленні захисту від хакерських атак, крадіжки особистих даних та інших кіберзагроз. Іншою важливою проблемою є забезпечення надійної інфраструктури для здійснення онлайн-транзакцій, що вимагає від підприємств значних інвестицій у технологічні рішення та підтримку серверних потужностей. Ще одним серйозним викликом є юридичне регулювання та стандартизація електронної комерції, що включає питання авторського права, умов користування, захисту споживачів та електронних платежів. Оскільки електронна комерція транснаціональна, виникають складнощі у визначенні юрисдикції, яка буде застосовуватися до транзакцій між різними країнами, що робить правову сферу в цій галузі особливо складною і потребує постійного розвитку міжнародних стандартів.

У довгостроковій перспективі можна прогнозувати, що електронна комерція продовжуватиме еволюціонувати, активно інтегруючи новітні технології, такі як штучний інтелект, блокчейн, а також технології

розширеної та віртуальної реальності. Впровадження штучного інтелекту дозволить більш точно прогнозувати потреби споживачів і оптимізувати асортимент товарів та послуг, що буде сприяти підвищенню рівня задоволення потреб користувачів. Технології блокчейн забезпечать ще більшу прозорість та безпеку транзакцій, що особливо важливо для підвищення довіри споживачів до електронної комерції [1-3]. Віртуальна і доповнена реальність, в свою чергу, дозволять створювати нові форми взаємодії з товарами, що дасть змогу користувачам «прикладати» продукти до свого простору чи навіть спробувати їх перед покупкою.

Таким чином, сутність електронної комерції полягає не тільки в переведенні традиційних форм бізнесу в цифровий формат, а й у кардинальних змінах самих принципів торгівлі та взаємодії на ринку. Вона відкриває нові можливості для підприємств і споживачів, дозволяючи долати бар'єри часу та простору. Однак для успішного розвитку електронної комерції необхідно вирішити низку складних технологічних, економічних, соціальних та правових проблем, що потребують постійного вдосконалення законодавчої бази та інноваційних технологічних рішень.

## 1.2 Типи онлайн-платформ електронної комерції

Онлайн-платформи електронної комерції є невід'ємною частиною сучасного ринку товарів і послуг, що використовують Інтернет як основний канал для здійснення бізнес-операцій. Вони забезпечують зручне і ефективне середовище для взаємодії продавців і покупців, що дозволяє здійснювати торгові операції, обмінюватися інформацією, а також вирішувати численні інші завдання, пов'язані з реалізацією товарів та послуг. На сьогодні існує кілька основних типів онлайн-платформ, кожен з яких має свої особливості та функціональні можливості, які впливають на ефективність електронної комерції та взаємодію учасників ринку.

Типи онлайн-платформ електронної комерції можна класифікувати за різними критеріями (рис. 1.2), зокрема за кількістю учасників, бізнес-моделями, характером товарів та послуг, які продаються, а також за рівнем інтеграції технологій та функціональних можливостей. Основними типами онлайн-платформ є: платформи для продажу товарів (B2C, B2B), платформи для продажу послуг (C2C, P2P), платформи для торгівлі через аукціони, платформи для цифрових товарів, а також гібридні платформи, що поєднують кілька моделей в одній [3, 4].

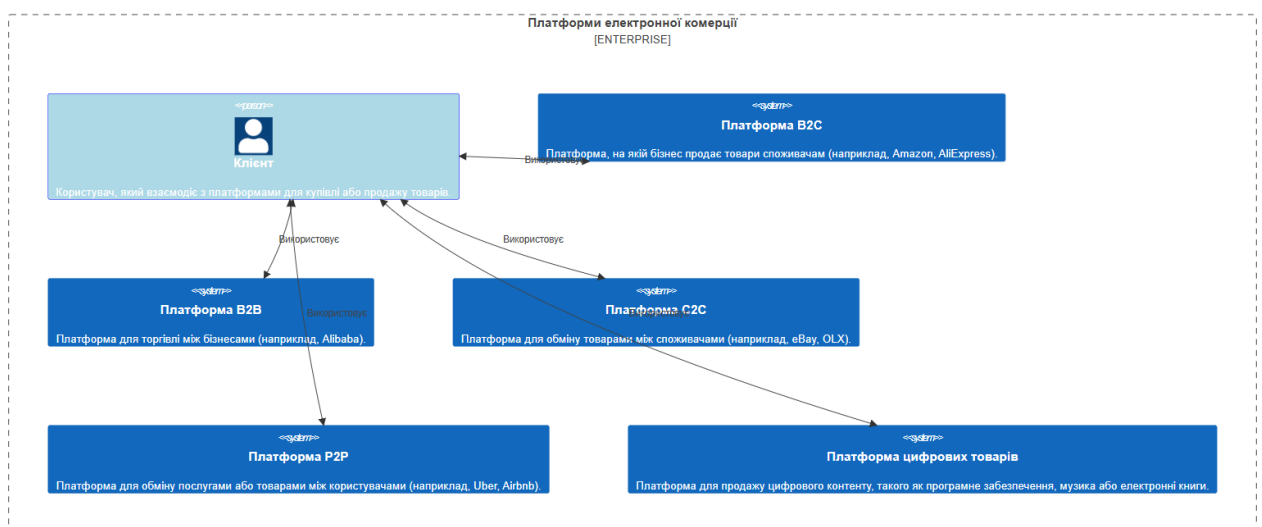


Рисунок 1.2 – Платформи електронної комерції

Платформи, що працюють за моделлю «бізнес для споживача» (B2C), є одними з найпоширеніших і включають в себе онлайн-магазини, де компанії продають свої товари чи послуги кінцевим споживачам. Цей тип платформи дозволяє бізнесам взаємодіяти з великими масами споживачів, пропонуючи їм можливість здійснювати покупки через Інтернет. Прикладом таких платформ є онлайн-торгові майданчики, такі як Amazon, eBay, Aliexpress, які мають величезну аудиторію клієнтів і забезпечують зручний процес покупки. Ці платформи зазвичай пропонують широкий асортимент товарів різних категорій – від електроніки та одягу до продуктів харчування та побутових товарів.

Особливістю платформ B2C є те, що вони вимагають від продавців налаштування ефективних систем управління товарними запасами, обробки замовлень, а також створення зручних платіжних механізмів. Вони також активно використовують аналітику для збирання даних про споживачів, що дає можливість краще розуміти їхні потреби і переваги, а також надавати індивідуалізовані пропозиції. Іншим важливим типом онлайн-платформ є платформи, що працюють за моделлю «бізнес для бізнесу» (B2B). Це платформи, на яких підприємства можуть продавати товари чи послуги іншим підприємствам. На відміну від B2C, в моделі B2B бізнеси взаємодіють між собою, часто з метою постачання великих партій товарів або виробничих матеріалів. Такі платформи використовуються для розміщення оптових пропозицій, торгівлі сировиною, комплектуючими та іншими товарами, що потребують спеціалізованих знань та умов закупівлі. Один з найвідоміших прикладів таких платформ – Alibaba, яка забезпечує можливість покупки товарів безпосередньо у виробників і постачальників.

У моделі B2B важливим аспектом є інтеграція між системами управління підприємствами (ERP), що дозволяє автоматизувати процеси закупівель, відстежувати товарні запаси, забезпечувати взаємодію з постачальниками та вирішувати логістичні завдання. Платформи цього типу часто забезпечують спеціалізовані інструменти для ведення переговорів, укладання контрактів, а також мають функції для організації платіжних операцій між підприємствами.

Однією з найцікавіших моделей електронної комерції є модель «споживач для споживача» (C2C). Ця модель передбачає, що кінцеві споживачі можуть продавати товари чи послуги іншим споживачам. Це поширений формат для платформ, що підтримують обмін товарами чи послугами без участі посередників. Приміром, такі платформи, як eBay або OLX, дозволяють користувачам продавати використані товари, обмінюватися речами, організовувати приватні аукціони та інші форми продажу між

споживачами. Платформи C2C можуть бути як онлайн-майданчиками, так і мобільними додатками для обміну або купівлі-продажу товарів [4, 5].

Основною характеристикою таких платформ є відсутність необхідності в структурованих бізнес-організаціях або великих інвестиціях. Користувачі платформ C2C часто мають можливість вільно визначати ціни на товари, що дає їм більшу свободу у виборі умов покупок та продажів. Проте, на цих платформах важливу роль відіграють механізми безпеки і довіри, оскільки покупці і продавці є неформальними сторонами угоди, що вимагає від платформ розробки ефективних систем рейтингу, відгуків і гарантій.

Моделі P2P (peer-to-peer) також мають схожість з C2C, однак вони більше орієнтовані на взаємодію між користувачами в контексті надання послуг або взаємного обміну, а не купівлі-продажу товарів. Такі платформи популярні для організації онлайн-сервісів, де люди можуть безпосередньо надавати послуги один одному. Прикладом таких платформ є Uber або Airbnb, де користувачі можуть отримувати послуги транспорту чи оренди житла від інших користувачів. Моделі P2P активно використовують новітні технології для створення зручних і безпечних умов для обміну послугами, а також для обробки транзакцій між користувачами.

Також існують платформи, орієнтовані на продаж цифрових товарів і послуг, до яких відносяться музика, книги, програмне забезпечення, ігри, відеоконтент та інші продукти [5]. Ці платформи дозволяють безпосередньо продавати або орендувати цифровий контент, що забезпечує зручність для покупців, оскільки вони можуть миттєво отримати доступ до придбаного контенту без необхідності фізичного отримання товару. Прикладом таких платформ є iTunes, Steam, Google Play та інші. Важливим аспектом для таких платформ є захист авторських прав і боротьба з піратством, що потребує постійного вдосконалення технологій захисту контенту.

Окремо варто відзначити гібридні платформи, які поєднують різні моделі електронної комерції. Це можуть бути платформи, на яких

здійснюються не лише продажі товарів, але й обмін послугами, а також платформи, які підтримують продаж і цифрових, і фізичних товарів. Такі платформи можуть адаптуватися до потреб різних типів користувачів, надаючи можливість торгівлі в різних форматах. Прикладом таких платформ є Amazon, яке є не лише торговим майданчиком для фізичних товарів, але й дозволяє продавати цифровий контент, надаючи можливість продавати електронні книги, музику, фільми та багато іншого [6].

Таким чином, онлайн-платформи електронної комерції мають різноманітні форми і структури, що дозволяє підприємствам і індивідуальним користувачам здійснювати покупки, продажі та обмін товарами і послугами за допомогою Інтернету. Кожен з типів платформ має свої переваги та недоліки, і вибір відповідної моделі залежить від потреб і цілей учасників ринку. Технологічні інновації, зміни в споживчих вподобаннях та розвиток нових бізнес-моделей безперервно змінюють ландшафт онлайн-комерції, тому вивчення цих платформ є важливим елементом для розуміння еволюції сучасного ринку.

### 1.3 Роль онлайн-платформ у сучасній економіці

Онлайн-платформи відіграють ключову роль у сучасній економіці, становлячи не лише інструмент для здійснення бізнес-операцій, а й фактор, що трансформує всю економічну структуру, змінюючи способи виробництва, торгівлі, обміну інформацією та взаємодії між учасниками ринку. Вони стали основою для розвитку нових бізнес-моделей, сприяючи інноваціям та інтеграції різних секторів економіки в єдиний цифровий простір. За допомогою онлайн-платформ компанії та приватні особи можуть здійснювати транзакції, обмінюватися інформацією, реалізовувати товари та послуги без значних географічних або часових обмежень. Це сприяло створенню нових форм бізнесу, зокрема цифрових економічних екосистем, а

також змінило традиційні уявлення про бізнес-процеси і організацію виробництва та торгівлі.

Насамперед важливо зазначити, що онлайн-платформи є тим механізмом, який дозволяє з'єднати учасників ринку в межах глобального цифрового простору. Вони служать посередниками між споживачами і постачальниками товарів чи послуг, надаючи інтерфейс для обміну інформацією, здійснення транзакцій, а також для отримання різноманітних послуг. Така роль платформ у сучасній економіці дозволяє здійснювати торговельні операції на абсолютно новому рівні, де традиційні економічні бар'єри, такі як фізичне розташування торгових точок, обсяги товарів чи навіть специфікацій на ринку, більше не є вирішальними [4-6]. Цей процес позбавляє бізнеси від багатьох операційних витрат і створює значно більш конкурентоспроможне середовище, що сприяє економічному зростанню. Однією з основних функцій онлайн-платформ є забезпечення доступу до інформації. Вони створюють нові способи доступу до товарів і послуг, що дозволяє не лише споживачам швидко знайти найкращі пропозиції на ринку, а й компаніям швидко адаптувати свої продукти до змінюваного попиту. Сучасні онлайн-платформи активно використовують великі дані (Big Data), штучний інтелект і аналітичні системи для прогнозування потреб ринку і індивідуалізованих пропозицій. Наприклад, електронні платформи, такі як Amazon або eBay, на основі даних, зібраних про попередні покупки користувачів, пропонують релевантні товари, що знижує час на ухвалення рішення і збільшує кількість продажів. Це також дозволяє продавцям виявляти тенденції споживчого попиту, адаптувати свої стратегії і асортимент до поточних вимог ринку.

Онлайн-платформи мають вирішальне значення і для розвитку малих та середніх підприємств, що отримують можливість виходити на глобальні ринки без необхідності інвестувати значні кошти в розвиток фізичних торгових точок або в логістичну інфраструктуру. Це відкриває доступ до

ширшої аудиторії, надаючи малому бізнесу можливість взаємодіяти з клієнтами по всьому світу. Такий доступ дозволяє малим підприємствам знижувати витрати на маркетинг і продажі, оскільки онлайн-платформи вже забезпечують велику частину цього процесу [4, 5]. Наприклад, компанії, що використовують платформи на кшталт Shopify або Etsy, можуть легко організувати свій онлайн-магазин, не маючи великих ресурсів для створення власної веб-інфраструктури або маркетингових кампаній.

Крім того, роль онлайн-платформ у сучасній економіці проявляється у суттєвому зменшенні витрат на транспортування товарів та послуг, що стало можливим завдяки розвитку інтернет-торгівлі і логістичних технологій. Технології обробки замовлень, відстеження товарів, автоматизація складських процесів значно знизили витрати, що виникають під час доставки товарів від постачальників до споживачів. Це дозволяє підприємствам, особливо тим, які працюють в електронній комерції, значно підвищити свою рентабельність, водночас забезпечуючи швидку доставку та зниження загальних витрат на логістику. Ще одним важливим аспектом є роль онлайн-платформ у розвитку інноваційних бізнес-моделей, зокрема таких, що базуються на спільному споживанні або доступі до товарів і послуг без необхідності їхнього повного володіння. Прикладом таких платформ є Airbnb, Uber, Lyft та інші сервіси спільного споживання, які значно змінили економічну модель споживання. Завдяки таким платформам користувачі можуть орендувати житло чи транспортні засоби, не володіючи ними. Це не лише знижує бар'єри для доступу до товарів та послуг, але й сприяє значному зменшенню витрат на утримання великої кількості активів. В результаті створюється нова економіка, заснована на ефективному використанні ресурсів і доступі до них за допомогою онлайн-платформ.

Також варто зазначити, що онлайн-платформи відіграють важливу роль у цифровізації різних секторів економіки, зокрема в галузях освіти, охорони здоров'я, фінансових послуг та багатьох інших. Платформи для онлайн-



освіти, такі як Coursera, Udemy або edX, дозволяють значно розширити доступ до знань і навчальних матеріалів для людей у всьому світі, роблячи освіту доступною для будь-кого, хто має доступ до Інтернету. У сфері охорони здоров'я, платформи, що надають медичні консультації онлайн, дозволяють пацієнтам отримувати медичні поради, консультації та рецепти, не відвідуючи лікарні чи клініки, що, зокрема, актуально в умовах пандемій чи для людей, які проживають в віддалених районах.

Водночас важливим аспектом є роль онлайн-платформ у фінансовому секторі. Платформи для електронних платежів, кредитування та страхування значно змінюють спосіб взаємодії користувачів з фінансовими установами. Протягом останніх років в Україні і світі з'явилися численні платформи для мікрокредитування, онлайн-банкінгу, а також платіжні системи, такі як PayPal, Revolut, що дозволяють здійснювати фінансові операції з мінімальними витратами часу і коштів [1-3]. Вони забезпечують ефективну інтеграцію з іншими секторами економіки, дозволяючи здійснювати операції з купівлі-продажу товарів, обміну валюти, інвестицій у цифрові активи та багато іншого.

Не можна не звернути увагу на правовий аспект ролі онлайн-платформ у сучасній економіці. Оскільки більшість таких платформ є міжнародними або міжрегіональними, виникає потреба в створенні єдиних стандартів для регулювання їхньої діяльності. Це включає забезпечення захисту прав споживачів, а також запобігання шахрайству, захист персональних даних користувачів, а також створення ефективних механізмів вирішення спірних ситуацій між покупцями і продавцями. Тому виникає необхідність у співпраці між урядами різних країн для створення єдиної правової бази, яка забезпечить стабільність і розвиток онлайн-комерції.

Таким чином, роль онлайн-платформ у сучасній економіці полягає не лише в оптимізації бізнес-процесів і зниженні витрат, а й у кардинальних змінах самих економічних і соціальних структур. Вони забезпечують доступ

до нових можливостей для бізнесів і споживачів, змінюють принципи споживання і виробництва, а також сприяють розвитку нових технологій і бізнес-моделей. У зв'язку з цим важливо розуміти їхній вплив на економічну реальність і відповідні тенденції в розвитку глобальної економіки.

#### 1.4 Аналіз сучасної літератури

Стаття [1] розглядає проблему традиційних систем рекомендацій, які базуються на попередніх вподобаннях користувачів для передбачення їх майбутніх смаків. Хоча вчені зробили певні покращення, враховуючи час та зміни інтересів користувачів, все ж складно точно відслідковувати ці зміни. Для вирішення цієї проблеми автори пропонують додати зміни персональних вподобань до традиційних колаборативних фільтраційних систем. У статті описується методика кількісної оцінки змін вподобань користувачів. Крім того, пропонується підхід, при якому рекомендовані не лише схожі, а й обрання товарів від різних користувачів, що допомагає відкривати нові інтереси. Експерименти, проведені на наборах даних MovieLens 1M та MovieLens Latest Small, показали, що метод із врахуванням змін у вподобаннях є ефективнішим за інші методи.

Стаття [2] розглядає проблему «перевантаження інформацією», яка стає все більш актуальною в еру сучасних технологій та штучного інтелекту. В умовах цього перевантаження рекомендаційні системи стають важливим інструментом для покращення досвіду користувачів і допомоги бізнесу в досягненні вищих прибутків. У дослідженні зосереджено увагу на гібридній системі рекомендацій для фільмів, яка базується на комбінованому підході. Така система дозволяє значно зменшити час, що витрачається на пошук інформації, і підвищує ефективність пошуку, орієнтуючись на персоналізовані потреби користувачів. Результати дослідження показали, що точність цієї гібридної системи досягає 81%, що перевищує точність

традиційних методів, таких як CB, Item-Based CF та User-Based CF. Цей підхід не лише скорочує час пошуку, але й значно підвищує ефективність пошукових запитів користувачів.

Стаття [3] розглядає проблему перевантаження інформацією в науковій сфері, зокрема в умовах постійного збільшення кількості опублікованих наукових досягнень, що ускладнює пошук релевантних матеріалів для дослідників. Основним інструментом для вирішення цієї проблеми є система рекомендацій, яка активно вивчається та застосовується в академічному середовищі. Особливу увагу приділено розвитку академічних систем рекомендацій, що використовують інтелектуальні алгоритми, зокрема традиційні та глибокі навчальні алгоритми. У статті надається огляд таких алгоритмів і описуються сучасні академічні системи рекомендацій, сподіваючись, що це допоможе в подальших дослідженнях у цій галузі.

Стаття [4] аналізує проблеми, з якими стикаються існуючі системи рекомендацій на багатьох платформах електронної комерції, зокрема, через обмеження в персоналізованих рекомендаціях і відсутність глибокого аналізу даних. Досліджуючи розвиток цих систем, а також результати досліджень у цій галузі, автори пропонують нову архітектуру персоналізованої системи рекомендацій, яка поєднує офлайн- та реальний моніторинг даних, а також технології глибокого навчання. Після визначення архітектури та функцій системи, детально розробляється механізм збору та попередньої обробки даних з різних платформ електронної комерції, щоб забезпечити персоналізований підхід у наступному етапі аналізу. Використовуючи технології аналізу великих даних, система автоматично рекомендує продукти, які відповідають індивідуальним вимогам користувачів.

Стаття [5] розглядає вплив швидкого розвитку мобільного Інтернету в Китаї та впровадження 5G на поведінку споживачів, зокрема на їх використання вебсайтів для покупок. Автори підкреслюють важливість великих даних, які стають ключовим напрямом для інтернет-компаній,

зокрема в контексті систем рекомендацій. Такі системи мають великий вплив на комерційні показники e-commerce компаній. Однак, важливою проблемою залишається зберігання величезних обсягів даних і ефективно видобування корисної інформації. У статті представлено розробку інтегрованої системи рекомендацій для електронної комерції, що базується на модифікованому наборі даних китайського аналогу Amazon та реальній бізнес-архітектурі e-commerce сайту. Система поєднує офлайн-рекомендації та реальний час для досягнення змішаного ефекту рекомендацій, використовуючи різноманітні методи аналізу та точні моделі для реалізації майнінгу даних.

Стаття [6] досліджує проблему перевантаження інформацією в університетських бібліотеках, що ускладнює доступ користувачів до релевантних ресурсів. Традиційні системи рекомендацій на основі колаборативної фільтрації часто стикаються з проблемами розрідженості даних і холодного старту, що знижує рівень задоволеності користувачів. Для вирішення цих проблем автори пропонують гібридну систему колаборативної фільтрації, яка поєднує контентно-орієнтовану фільтрацію, колаборативну фільтрацію за користувачами та алгоритм кластеризації. Для зменшення розрідженості даних використовується точна кластеризація за допомогою K-середніх, що дозволяє групувати схожих користувачів і покращувати результати рекомендацій. Завдяки адаптації контентно-орієнтованих методів також вирішується проблема холодного старту, забезпечуючи правильні рекомендації для нових користувачів та ресурсів. Запропонована гібридна система показала значно кращу точність рекомендацій (99.74%), рівень відгуку рекомендацій 0.73 та середню квадратну похибку 0.49 порівняно з існуючими методами колаборативної фільтрації, оптимізованими гібридними методами та методами на основі розкладу сингулярних значень.

Стаття [7] розглядає необхідність наявності системи рекомендацій на кожній онлайн-платформі для підвищення залученості користувачів. Система

рекомендацій пропонує продукти або послуги, орієнтуючись на переваги та історію користувача, що полегшує йому вибір і стимулює активність. Описано, як система рекомендацій створює взаємозв'язки між продуктами або користувачами та на основі цих зв'язків може пропонувати відповідні продукти. Також розглядається гібридна система рекомендацій, яка поєднує прогнози з двох або більше методів. У цьому випадку гібридна система об'єднує прогнози, засновані на характеристиках продуктів, що реалізуються за допомогою графових методів, і прогнози на основі користувачів, що здійснюються через колаборативну фільтрацію.

Стаття [8] аналізує застосування алгоритмів штучного інтелекту в системах рекомендацій для туристичних атракцій, зокрема, оцінюючи їх вплив на задоволеність користувачів і точність рекомендацій. У дослідженні порівнюються традиційні методи рекомендацій з системою, заснованою на штучному інтелекті. Результати експериментів показують, що система, що використовує алгоритми штучного інтелекту, підвищує рівень задоволення користувачів на 14% порівняно з традиційним методом, досягнувши оцінки 4.8 бали. Крім того, точність рекомендацій значно зросла, і 82.7% результатів відповідали реальним вподобанням користувачів, що значно перевищує 65.3% у традиційній системі. Таким чином, система на основі штучного інтелекту покращує не тільки задоволеність користувачів, але й точність рекомендацій, що дозволяє краще розуміти персональні вподобання, потреби та інтереси користувачів і надавати більш точні та персоналізовані пропозиції туристичних атракцій. Це має велике значення для вдосконалення систем рекомендацій у туристичній індустрії та покращує досвід подорожей для користувачів.

Стаття [9] розглядає важливість ефективної організації та використання інформації в електронній комерції для максимального розуміння вподобань клієнтів, оптимізації дизайну вебсайтів і спрощення процесу покупок для споживачів. Для вирішення цієї проблеми автори поєднали метод матричної

факторизації (MF) і згорткові нейронні мережі (CNN) для створення персоналізованої рекомендаційної моделі MF-CNN, яку застосували в високопродуктивній системі персоналізованих рекомендацій для e-commerce. Стаття наводить результати дослідження, де під час періоду з 1 по 11 число загальний рівень кліків на вебсайті залишався в межах стабільного діапазону. В період з 12 по 22 число без використання персоналізованих рекомендацій спостерігалися значні коливання кількості кліків на популярних товарах на головній сторінці, а з 23 по 29 число, коли були впроваджені персоналізовані рекомендації, рівень кліків збільшився. Ця система персоналізованих рекомендацій показала свою ефективність і має важливе значення для вдосконалення рекомендацій в сфері електронної комерції.

Стаття [10] описує розробку системи рекомендацій для курсу «Електронне проектування», що є важливим для участі в електронних конкурсах з проектування. Оскільки курс охоплює широкий спектр знань, деякі модулі не вивчаються безпосередньо на заняттях, і студенти повинні самостійно готуватися до уроків, орієнтуючись на ресурси, рекомендовані викладачами. Для цього була розроблена система рекомендацій на основі вдосконаленого алгоритму колаборативної фільтрації, який враховує схожість атрибутів користувачів для підвищення точності та вирішення проблеми холодного старту та розрідженості даних. Система застосовує цей підхід для автоматичного рекомендаційного підбору навчальних ресурсів, що сприяє самостійному навчанню студентів за підтримки викладачів. Для реалізації цього підходу створено експериментальну платформу навчання для курсу «Електронне проектування».

Стаття [11] розглядає розвиток інтелектуальних рекомендацій у сфері електронної комерції, яка швидко розвивається завдяки популярності смартфонів. Оскільки рекомендаційні системи є важливим елементом у цій галузі, дослідники використовують різні технології, такі як асоціативні правила, колаборативну фільтрацію, марковські ланцюги та рекурентні

нейронні мережі для рекомендацій товарів. У цій статті основну увагу приділено системам інтелектуальних рекомендацій для електронної комерції, які використовують глибоке навчання. Спочатку автори розробляють загальний дизайн такої системи, пропонуючи функціональні модулі та архітектуру. Далі обговорюється алгоритм рекомендацій у цих системах і оптимізація їх за допомогою згорткових нейронних мереж. У заключній частині порівнюються та аналізуються результати трьох популярних алгоритмів рекомендацій на наборі даних Alibaba, де результати експериментів показали, що запропонований алгоритм на основі згорткових нейронних мереж перевершує інші методи, демонструючи значну практичну цінність і потенціал для подальшого розвитку.

Стаття [12] пропонує розробку інтелектуальної системи рекомендацій для експлуатації та технічного обслуговування театрального обладнання на основі глибоких нейронних мереж і механізму уваги. У театрі існує велика кількість та різноманітність обладнання для виступів, що ускладнює і сповільнює процес його обслуговування вручну. Для зменшення витрат часу та енергії на ручне обслуговування пропонується створення системи, яка використовує дані про обладнання для побудови датасету, застосовує схему попередньої обробки сирих даних, а також використовує глибоку нейронну мережу для навчання характеристик даних. Важливим компонентом є інтеграція механізму уваги, який дозволяє системі зосереджуватися на важливих аспектах інформації. Для генерації результатів рекомендацій використовується косинусна схожість між обчисленими характеристиками. Наводиться ілюстративний приклад, який демонструє ефективність запропонованої системи.

Стаття [13] зосереджена на використанні алгоритмів машинного навчання в системах рекомендацій для продуктів штучного інтелекту, зокрема досліджує, як різні алгоритми впливають на ефективність рекомендацій. Автори використовують кілька алгоритмів машинного

навчання, таких як логістична регресія (LR), випадковий ліс (RF) та градієнтне підсилення рішень (GBDT), щоб оцінити їхню практичну ефективність у рамках систем рекомендацій. Експериментальні результати показують, що застосування цих алгоритмів суттєво підвищує точність рекомендацій і задоволеність користувачів. Зокрема, GBDT ефективно вирішує проблеми дисбалансу даних та холодного старту користувачів, RF демонструє стабільність та здатність до узагальнення при обробці багатовимірних характеристик, а LR забезпечує високу інтерпретованість моделі. Це дослідження надає корисні відомості щодо застосування алгоритмів машинного навчання в системах рекомендацій і закладає основу для майбутніх досліджень у сфері персоналізованих та інтелектуальних рекомендацій.

Стаття [14] розглядає проблему перевантаження інформацією в Інтернеті, зокрема на платформі Weibo, де велика кількість непотрібної та зайвої інформації заважає користувачам точно і швидко знаходити необхідний контент. Для вирішення цієї проблеми запропоновано створення персоналізованої медіа-рекомендаційної системи, що базується на мережевих алгоритмах. У статті детально обговорюються функціональні вимоги до такої системи та її класифікація за типами рекомендацій. Проведені експериментальні випробування цієї системи на прикладі Weibo показали, що співвідношення між кількістю ефективних рекомендацій і кількістю користувачів у різних категоріях досягає 79.7%. Це свідчить про точність аналізу інтересів користувачів у моделі, що лежить в основі системи, і підтверджує ефективність запропонованого алгоритму.

Стаття [15] аналізує проблему, з якою стикаються споживачі під час онлайн-шопінгу, особливо під час пандемії COVID-19, коли вони повинні витратити багато часу на пошук відповідних товарів серед великої кількості запропонованих варіантів. Для полегшення цього процесу були розроблені різні системи рекомендацій. У статті розглядаються методи на основі



навчання, які допомагають вирішувати проблему холодного старту, соціальні рекомендації та проблему розрідженості даних у системах рекомендацій для електронної комерції. Автори порівнюють переваги і недоліки цих методів, а також вказують напрямки для подальших досліджень у цій сфері.

Стаття [16] розглядає зростаючий попит на продукцію з екологічною упаковкою та необхідність розвитку відповідних систем рекомендацій для таких товарів на тлі зростання інтернет-технологій. Використовуючи технології мульти-алгоритмного злиття та дані з обробки даних, автори пропонують рішення для вирішення проблеми перевантаження інформацією, яка часто виникає в системах рекомендацій. Така система не лише сприяє конверсії покупок користувачів, а й надає їм обґрунтовані рекомендації для прийняття рішень, покращуючи загальний досвід покупок. Окрім цього, система закладає теоретичну та проектну основу для подальших досліджень і розробок в області рекомендаційних систем для екологічних товарів.

Стаття [17] пропонує метод персоналізованих гібридних рекомендацій для вирішення проблеми холодного старту в сценаріях музичних рекомендацій, використовуючи знання з графів. Метод передбачає створення музичного графа знань, який базується на бібліотеці Satori та записах взаємодії користувачів з музикою. У цьому підході також здійснюється перетворення векторів ознак сутностей і зв'язків, отриманих з графа знань, у низькорозмірні вектори. Завдяки операціям агрегації сусідів вдало вловлюється і зберігається локальна структура близькості кожної сутності. Додатково вивчаються вподобання користувачів на основі цього графа для рекомендаційної системи. Остаточний список рекомендацій генерується на основі потенційних характеристик музики, передбачених за допомогою моделі регресії CNN. Експериментальні порівняння на Last.FM показують, що цей алгоритм перевершує традиційні методи рекомендацій і методи, що базуються на графах знань, ефективно виявляючи потенційні зв'язки між користувачами та музикою та покращуючи якість музичних рекомендацій.

Стаття [18] розробляє персоналізовану систему рекомендацій для електронної комерції на основі фузії мультимодальних даних. Спочатку аналізуються вимоги до системи, визначаються її функціональні можливості, а потім розробляються різні стратегії для різних категорій користувачів з метою досягнення найкращих персоналізованих рекомендацій, що дозволяє створити загальну архітектуру системи. Для вирішення проблеми розріджених даних і поліпшення здатності системи відображати атрибути товарів використовуються мультимодальні ознаки. Окрім цього, в статті представлено бізнес-процес персоналізованої рекомендаційної системи, заснованої на фузії мультимодальних даних. Остаточні результати показують, що середнє абсолютне відхилення для алгоритмів колаборативної фільтрації та машинного навчання становить 0.540 і 0.531 відповідно, тоді як для алгоритму фузії мультимодальних даних цей показник дорівнює 0.315, що підтверджує високу якість рекомендацій в запропонованій системі.

### 1.5 Висновки до першого розділу

Сутність електронної комерції полягає в тому, що вона є універсальним інструментом для здійснення торгових операцій через Інтернет, що змінює традиційні механізми торгівлі та взаємодії між бізнесом і споживачами. Вона забезпечує зручний доступ до товарів і послуг, роблячи процес покупок доступним на глобальному рівні, а також дозволяє підприємствам оптимізувати свої операційні витрати та адаптуватися до змінюваного попиту. Розуміння різноманіття типів онлайн-платформ електронної комерції дозволяє краще оцінити можливості та обмеження кожного типу для різних учасників ринку. Платформи можуть бути орієнтовані на продаж товарів і послуг (B2C, B2B), обмін товарами між споживачами (C2C), а також на надання цифрових товарів чи послуг. Кожен з типів платформ має свої особливості, що важливо враховувати при розробці стратегії ведення бізнесу.

Онлайн-платформи не лише оптимізують процеси торгівлі, але й сприяють зміні економічних структур і зниженню традиційних бар'єрів для входу на ринок. Вони підтримують розвиток нових бізнес-моделей, таких як спільне споживання або доступ до ресурсів без їхнього повного володіння, що змінює традиційні уявлення про володіння активами і змушує економіку стати більш гнучкою та адаптивною. Онлайн-платформи є потужними двигунами інновацій, оскільки вони використовують новітні технології для аналізу даних, прогнозування попиту та персоналізації пропозицій. Це дозволяє компаніям швидко адаптуватися до змінюваних умов ринку та краще задовольняти потреби своїх клієнтів.

Платформи дозволяють малим та середнім підприємствам виходити на глобальні ринки без значних інвестицій в інфраструктуру, знижуючи витрати на маркетинг, продажі та логістику. Це відкриває нові можливості для бізнесу та стимулює економічний розвиток на локальних та міжнародних рівнях. Розвиток онлайн-платформ і пов'язаних з ними цифрових бізнес-моделей вимагає розробки єдиних стандартів правового регулювання для забезпечення безпеки транзакцій, захисту персональних даних та запобігання шахрайству. Це дозволяє стабілізувати ринок і сприяє розвитку довіри між учасниками електронної комерції.

Оскільки роль онлайн-платформ у сучасній економіці є значною, дослідження повинно зосереджуватися на аналізі специфічних аспектів їхнього впливу на різні сектори економіки, вивченні різних типів платформ і їхніх особливостей, а також визначенні основних тенденцій і проблем у галузі електронної комерції, що потребують подальшого вивчення та вдосконалення. Таким чином, цей розділ підкреслює важливість онлайн-платформ для сучасної економіки, аналізує різноманіття їхніх типів і функцій, а також вказує на ключові напрями, які потребують глибшого дослідження в рамках подальших етапів наукової роботи.

## РОЗДІЛ 2. ДОСЛІДЖЕННЯ МЕТОДІВ РОЗРОБЛЕННЯ РЕКОМЕНДАЦІЙНОЇ СИСТЕМИ

### 2.1 Визначення та функції рекомендаційних систем

Рекомендаційні системи стали невід’ємною частиною сучасного інформаційного суспільства, а їх застосування охоплює широкий спектр сфер – від електронної комерції та онлайн-розваг до медичних досліджень і освітніх платформ [4]. Ці системи використовуються для того, щоб на основі аналізу великих масивів даних робити рекомендації щодо вибору продуктів, послуг чи іншого контенту, який міг би зацікавити користувачів. Визначення рекомендаційної системи можна сформулювати як програмний засіб або методологію, що дозволяє автоматично передбачати інтереси та потреби користувачів, ґрунтуючись на їхній поведінці або схожості з іншими користувачами, а також на інших параметрах, таких як контекст чи характеристики продуктів. Ключовою метою рекомендаційної системи є надання персоналізованого досвіду взаємодії з інформаційними ресурсами, що дозволяє підвищити ефективність користування такими ресурсами. Однією з основних проблем, яку вона вирішує, є так звана «інформаційна перевантаженість», коли користувачі стикаються з величезною кількістю вибору і не здатні самостійно визначити, що саме з цього списку може бути для них релевантним і цікавим [5]. Рекомендаційна система дозволяє мінімізувати цю проблему, пропонуючи користувачеві лише ті варіанти, які найбільше відповідають його запитам, смакам чи уподобанням.

Функції, які виконують рекомендаційні системи, різноманітні й залежать від конкретної галузі застосування. У сфері електронної комерції вони допомагають оптимізувати процес покупки, пропонуючи користувачам товари, які можуть їх зацікавити на основі їхніх попередніх покупок, переглядів або навіть уподобань. У музичних та відеосервісах

рекомендаційні системи можуть визначати, які пісні чи фільми можуть сподобатися користувачу, враховуючи його попередній вибір або схожість з іншими користувачами з подібними інтересами. У соціальних мережах рекомендації часто допомагають користувачам знаходити нові пости, групи чи людей для підписки, аналізуючи їхні соціальні зв'язки та взаємодії.

Технології, які використовуються для розробки рекомендаційних систем, включають різні методи машинного навчання, аналізу даних, статистики та обробки природної мови. Найпоширенішими підходами є колаборативне фільтрування, контентне фільтрування та гібридні моделі. Колаборативне фільтрування базується на припущенні, що користувачі, які мали схожі уподобання в минулому, ймовірно, будуть мати схожі вподобання і в майбутньому. Цей підхід може бути двох видів: користувацьке колаборативне фільтрування, яке базується на порівнянні користувачів, та товарне колаборативне фільтрування, яке орієнтоване на порівняння товарів або контенту між собою. Контентне фільтрування, в свою чергу, зосереджено на характеристиках самих товарів або предметів, які користувачі переглядають, і рекомендує схожі продукти на основі цих характеристик [6].

Гібридні моделі поєднують різні підходи для досягнення кращих результатів. Вони можуть комбінувати колаборативне та контентне фільтрування, або використовувати додаткові методи, наприклад, зворотній зв'язок від користувачів для коригування рекомендацій. Такий підхід дає змогу значно покращити точність рекомендацій, зменшуючи вплив проблем, властивих кожному з окремих підходів, наприклад, таку проблему, як холодний старт (коли система не має достатньо інформації для ефективних рекомендацій, наприклад, для нового користувача чи нового товару). Інші важливі аспекти, які повинні бути враховані під час розробки рекомендаційної системи, включають обробку великих даних, забезпечення швидкості роботи та персоналізації рекомендацій, а також захист

конфіденційності користувачів. В умовах великих обсягів інформації важливо ефективно збирати, зберігати та аналізувати дані, що дозволяють робити точні прогнози щодо інтересів користувачів [7]. Різноманітні технічні інструменти, такі як розподілені обчислювальні системи або технології обробки великих даних, відіграють важливу роль у забезпеченні масштабованості та продуктивності рекомендаційних систем.

Однією з основних проблем, з якою стикаються розробники таких систем, є забезпечення балансу між точністю рекомендацій і їхньою різноманітністю. У деяких випадках система може пропонувати дуже схожі варіанти на основі попереднього досвіду користувача, що обмежує його вибір і може призвести до одноманітності в досвіді. Для того, щоб це не стало проблемою, рекомендаційні системи часто включають додаткові механізми, які допомагають збільшити різноманітність пропонованих варіантів, навіть якщо вони не є найбільш ймовірними з погляду аналізу даних. Це дозволяє користувачам відкривати нові продукти чи контент, які можуть бути їм цікавими, навіть якщо вони не зовсім відповідають їхнім звичним уподобанням. Іншим аспектом є питання етики та приватності. Збирання великих обсягів даних про користувачів та їхні уподобання може викликати занепокоєння з приводу захисту особистої інформації. Системи повинні бути спроектовані таким чином, щоб не порушувати приватність користувачів і забезпечувати контроль над даними, якими вони діляться. Водночас важливо, щоб користувачі мали можливість розуміти, як саме збираються їхні дані та як вони використовуються для генерації рекомендацій. Розробка та впровадження етичних стандартів в області рекомендаційних систем стає все більш актуальним завданням.

Загалом, рекомендаційні системи займають важливе місце в економіці та суспільстві, зокрема в e-commerce, культурних і розважальних індустріях, а також у сфері охорони здоров'я та освіти. Вони здатні значно покращити ефективність взаємодії користувачів з різними платформами, допомагаючи їм

знаходити найбільш релевантний контент чи продукти, що відповідають їхнім інтересам та потребам. Технологічний прогрес в галузі машинного навчання та штучного інтелекту сприяє розвитку все більш складних і ефективних рекомендаційних систем, здатних працювати з величезними обсягами даних і надавати персоналізовані рекомендації в реальному часі. Тим не менш, залишається важливим постійно працювати над удосконаленням таких систем, зокрема в аспектах точності рекомендацій, забезпечення приватності користувачів і етики використання даних.

## 2.2 Ключові завдання та принципи роботи рекомендаційних систем

У таблиці 2.1 наведено ключові завдання та принципи роботи рекомендаційних систем [6, 7].

Таблиця 2.1

### Ключові завдання та принципи роботи

№	Ключові завдання	Принципи роботи
1	Аналіз користувацьких уподобань	Ідентифікація інтересів через аналіз історії взаємодії
2	Персоналізація рекомендацій	Підбір продуктів або контенту, що відповідає уподобанням користувача
3	Оптимізація пошуку	Забезпечення точності результатів пошуку через навчання на історичних даних
4	Поліпшення досвіду користувача	Забезпечення інтуїтивного та зручного досвіду завдяки персоналізованим рекомендаціям
5	Система вивчення переваг нових користувачів	Рішення проблеми холодного старту
6	Адаптація до змінюваних уподобань	Реагування на нові тенденції та зміни в інтересах користувачів

Продовження табл. 2.1

7	Забезпечення рекомендацій у реальному часі	Пропонування варіантів в реальному часі на основі поточної активності користувача
8	Інтеграція з іншими системами	Об'єднання даних з різних джерел для покращення рекомендацій
9	Запобігання перегрузці рекомендаціями	Забезпечення різноманітності та балансу між точністю і різноманітністю
10	Оптимізація результатів на основі фідбеку	Коригування рекомендацій на основі відгуків користувачів

У таблиці 2.2 наведено методи фільтрації рекомендаційних систем [7, 8].

Таблиця 2.2

## Методи фільтрації

№	Типи даних, що використовуються	Методи фільтрації
1	Історія пошуку, покупки, перегляду контенту	Колаборативне фільтрування, контентне фільтрування
2	Профіль користувача, характеристики товарів	Гібридні моделі, колаборативне фільтрування
3	Текстовий контент, класифікація	Обробка природної мови (NLP), машинне навчання
4	Поведінка користувачів, відгуки, рейтинги	Аналіз даних, колаборативне фільтрування
5	Немає історії взаємодії	Гібридні методи, залучення додаткових даних
6	Зміна в поведінці, сезонні тренди	Динамічне оновлення моделей



## Продовження табл. 2.2

7	Поточні запити, поведінка в реальному часі	Миттєве колаборативне фільтрування
8	Зовнішні API, партнерства	Гібридні моделі
9	Історія взаємодії користувача, оцінки	Адаптивні алгоритми
10	Оцінки, коментарі	Методи зворотного зв'язку

### 2.3 Системи на основі фільтрації за змістом

Системи на основі фільтрації за змістом (content-based filtering) є одними з основних підходів у розробці рекомендаційних систем, які використовуються для автоматичного надання користувачам продуктів, послуг або контенту, що найбільше відповідає їхнім інтересам та уподобанням. Цей метод орієнтований на аналіз властивостей самих об'єктів, що пропонуються для рекомендації, і потребує від системи детального розуміння контексту та змісту даних, з якими вона працює. Технології фільтрації за змістом застосовуються в широкому спектрі задач, таких як рекомендація фільмів, музичних треків, новин, книг, товарів у інтернет-магазинах тощо [8]. Суть цієї технології полягає в тому, що система вивчає уподобання користувача, ґрунтуючись на характеристиках об'єктів, з якими він взаємодіє, і використовує ці характеристики для рекомендації подібних об'єктів.

Принцип роботи систем на основі фільтрації за змістом полягає в аналізі профілю користувача, що складається з його уподобань щодо характеристик певних елементів контенту. Наприклад, для рекомендації фільмів система може використовувати такі характеристики, як жанр, режисер, актори, рік випуску, оцінки критиків, а також інші атрибути, які описують фільм. Коли користувач переглядає або оцінює певний фільм,

система робить висновок про те, що він схильний до подібного контенту, і використовує ці знання для рекомендації інших фільмів з подібними характеристиками.

Аналізуючи такі дані, система може виділити основні властивості продуктів або контенту, які важливі для користувачів, і в результаті розробити модель, яка здатна передбачити, який контент буде цікавий конкретному користувачу. Наприклад, для системи рекомендацій фільмів можна створити профіль користувача, який враховує всі фільми, що він переглядав і оцінив високо, і з'ясувати, що ці фільми мають спільні риси, як-то жанр, актори, тематика чи стилістика. Зібравши таку інформацію, система може порекомендувати нові фільми, які мають схожі характеристики.

Технічно фільтрація за змістом ґрунтується на кількох ключових етапах. Першим етапом є створення профілю контенту [9]. Для цього необхідно обробити всі доступні дані про об'єкти, які можуть бути рекомендовані. Якщо йдеться про фільми, це можуть бути жанри, актори, режисери, оцінки, відгуки, дата випуску, а також будь-яка інша доступна інформація, що описує ці фільми. Цей процес зазвичай включає в себе попередню обробку тексту, таку як виділення суттєвих характеристик (наприклад, тематичних слів у описах), а також можливе застосування технік машинного навчання для виявлення важливих ознак, що можуть бути використані для класифікації.

Далі на основі отриманих даних створюється профіль користувача. Цей профіль складається з властивостей, які найкраще відповідають інтересам користувача, і будується на основі його попередніх взаємодій із контентом. Наприклад, якщо користувач часто переглядає фільми в жанрі фантастики або драми, його профіль буде мати високий показник за цими категоріями. Для побудови таких профілів використовуються різні методи, від простого порівняння характеристик до складніших технік машинного навчання, таких

як кластеризація або побудова моделей на основі аналізу векторних просторів.

Одним із важливих аспектів систем фільтрації за змістом є обчислення подібності між об'єктами. Для цього використовуються різноманітні метрики, зокрема косинусна подібність, евклідова відстань або інші методи вимірювання схожості. Наприклад, косинусна подібність дозволяє оцінити, наскільки два об'єкти схожі за своїми характеристиками, порівнюючи вектори, що їх представляють. У випадку з рекомендацією фільмів це може бути вектор, що містить інформацію про жанри, акторів, теми, рейтинг і так далі. Обчислення подібності між об'єктами дозволяє визначити, які з них найбільше відповідають інтересам конкретного користувача.

Застосування фільтрації за змістом має кілька переваг, серед яких варто відзначити можливість рекомендацій для нових користувачів без необхідності попереднього збору великих обсягів даних про їхні вподобання. Це робить таку систему дуже ефективною в умовах, коли дані про користувачів є обмеженими. Оскільки система не потребує інформації про взаємодії інших користувачів (на відміну від колаборативного фільтрування), вона може працювати навіть за відсутності широкої бази даних про інші вподобання користувачів. Однак варто зазначити, що цей підхід також має свої обмеження. Наприклад, фільтрація за змістом може стати обмеженою, якщо контент не має чітко виражених характеристик або якщо сама система не може правильно виділити ключові ознаки для класифікації [10]. Крім того, такі системи можуть бути схильні до «перегріву» рекомендацій, коли користувач отримує лише подібний контент і не має можливості дізнатися про нові, нестандартні об'єкти, що могли б його зацікавити. Іншою проблемою є відсутність можливості враховувати зміни у вподобаннях користувачів. Якщо користувач змінює свої інтереси або починає взаємодіяти з новими типами контенту, система на основі фільтрації за змістом може бути не здатною швидко адаптуватися до цих змін. Однак

сучасні підходи, такі як комбіновані методи або гібридні моделі, можуть частково вирішувати цю проблему, поєднуючи фільтрацію за змістом з іншими методами, такими як колаборативне фільтрування, що дозволяє враховувати взаємодії інших користувачів і враховувати зміну вподобань користувача.

Поряд із традиційними методами, останнім часом активно розвиваються методи на основі глибокого навчання, які використовують нейронні мережі для аналізу складних взаємозв'язків між характеристиками контенту та перевагами користувачів. Використання таких підходів дозволяє значно підвищити точність рекомендацій, а також дозволяє обробляти величезні обсяги даних, що відкриває нові можливості для створення інтелектуальних систем рекомендацій, які можуть адаптуватися до інтересів користувачів у реальному часі.

Відзначаючи науково-практичні аспекти фільтрації за змістом, можна також говорити про її застосування в різних сферах життя. У сферах електронної комерції та маркетингу цей підхід активно використовують для створення індивідуальних рекомендаційних систем, що дозволяє бізнесам пропонувати клієнтам саме той продукт, який ймовірно зацікавить їх, зважаючи на історію покупок або переглядів [11]. У медіаіндустрії фільтрація за змістом дозволяє онлайн-платформам, таким як Netflix чи YouTube, рекомендувати фільми, серіали або відео, враховуючи попередні уподобання користувача та їхні взаємодії з контентом. Технології цієї категорії також активно застосовуються в соціальних мережах, де алгоритми, що базуються на фільтрації за змістом, допомагають формувати новинні стрічки, пропонуючи користувачам цікаві публікації, статті або відео, які відповідають їхнім інтересам.

Таким чином, системи на основі фільтрації за змістом являють собою потужний інструмент для побудови персоналізованих рекомендацій, використовуючи характеристики самих об'єктів, які пропонуються

користувачам. Вони демонструють значні переваги в умовах обмежених даних, але також мають свої обмеження, пов'язані з недостатньою різноманітністю рекомендацій і недостатньою гнучкістю у реагуванні на зміни в інтересах користувачів. Проте, розвиток новітніх технологій, таких як глибоке навчання, дозволяє значно покращити ці системи, відкриваючи нові можливості для більш точних і адаптивних рекомендацій.

#### 2.4 Системи на основі колаборативної фільтрації

Системи на основі колаборативної фільтрації є одними з найпоширеніших підходів у розробці рекомендаційних систем, які використовуються в різних галузях, таких як електронна комерція, медіаіндустрія, соціальні мережі та багато інших. Колаборативна фільтрація базується на ідеї, що подібні інтереси або переваги між різними користувачами можуть бути використані для формування рекомендацій для конкретного користувача. Цей підхід припускає, що якщо два користувачі мають схожі уподобання щодо певного набору об'єктів, то ймовірно, що їхні інтереси будуть схожими і в майбутньому, і тому вони можуть взаємно ділитися своїми вподобаннями через систему рекомендацій [12]. Основна ідея колаборативної фільтрації полягає в тому, що замість того, щоб використовувати інформацію про самі об'єкти або продукти для надання рекомендацій (як у системах на основі фільтрації за змістом), система орієнтується на поведінку користувачів, їхні взаємодії з об'єктами та подібність між ними. Коли користувач взаємодіє з певним об'єктом (наприклад, оцінює фільм, класифікує товар або переглядає певний тип контенту), його поведінка зберігається в системі, що дозволяє іншим користувачам, які демонструють подібну поведінку, обмінюватися своїми вподобаннями через рекомендаційний механізм. Таким чином, колаборативна фільтрація створює можливість надання рекомендацій,

заснованих на масивних наборах даних від інших користувачів, які виявляють схожі смаки і уподобання.

Системи колаборативної фільтрації можна поділити на два основних підтипи: фільтрація за допомогою користувачів (user-based collaborative filtering) і фільтрація за допомогою елементів (item-based collaborative filtering). Обидва підходи мають спільну ідею – знаходити подібності між користувачами або між елементами, але вони різняться в тому, на чому зосереджуються при порівнянні даних.

У першому випадку, при фільтрації за допомогою користувачів, система намагається знайти користувачів, чий профіль (їхня поведінка, вподобання та інші взаємодії з контентом) є схожим на профіль цільового користувача. Після цього система рекомендує йому об'єкти, які сподобалися схожим користувачам. Технічно це передбачає обчислення схожості між користувачами за допомогою різних метрик, таких як косинусна подібність або кореляція Пірсона, що дозволяє оцінити, наскільки подібними є їхні вподобання.

У другому випадку, при фільтрації за допомогою елементів, система знаходить подібність між самими об'єктами. Це означає, що якщо користувач оцінив або взаємодіяв з певним об'єктом, система буде шукати інші об'єкти, які схожі на той, з яким взаємодівав користувач, і рекомендувати їх. Цей метод передбачає, що об'єкти, які мають схожі характеристики (напрямки тематики, жанр, особливості відгуків та оцінок), будуть цікавими для тих, хто вже взаємодіяв з подібними елементами [12, 13]. Фільтрація за допомогою елементів часто застосовується в електронній комерції для рекомендації товарів, що схожі на вже придбані або переглянуті.

Обидва підходи мають свої переваги та недоліки. Фільтрація за допомогою користувачів може бути ефективною, якщо система має велику кількість активних користувачів, чия поведінка може бути використана для побудови точних рекомендацій. Однак цей підхід може зіткнутися з

проблемою холодного старту для нових користувачів, оскільки система ще не має достатньої кількості даних про їхні вподобання для визначення схожості з іншими користувачами. Фільтрація за допомогою елементів, з іншого боку, може бути ефективнішою в системах, де об'єкти мають чітко виражені характеристики, які можна порівнювати між собою. Однак цей підхід також має свої обмеження, оскільки не завжди можливо правильно визначити подібність між елементами без контексту, на основі якого користувачі формують свої уподобання.

Порівняння користувачів або елементів в системах колаборативної фільтрації здійснюється за допомогою метрик подібності, які оцінюють схожість двох користувачів або двох елементів. Однією з найбільш популярних метрик є косинусна подібність, яка використовується для вимірювання кута між двома векторами в багатовимірному просторі. Чим менший цей кут, тим більш подібними є два об'єкти. Іншими важливими метриками є кореляція Пірсона, яка вимірює лінійний зв'язок між двома наборами оцінок, та евклідова відстань, яка використовується для вимірювання відстані між об'єктами в просторі характеристик [14].

Один із ключових аспектів роботи систем колаборативної фільтрації – це обчислення величезних матриць взаємодій користувачів і елементів. Для обробки таких великих обсягів даних зазвичай використовуються різні методи зниження розмірності, такі як метод головних компонент (PCA) або матричне розкладання. Наприклад, одна з популярних технік – метод SVD (сингулярне розкладання матриці), який дозволяє зменшити кількість вимірів, зберігаючи основні патерни взаємодій, що робить обчислення більш ефективними і зменшує обсяг необхідних даних для прогнозування.

Колаборативна фільтрація має низку переваг, серед яких слід відзначити здатність робити рекомендації без необхідності аналізувати характеристики самих об'єктів, що значно знижує складність моделювання. Крім того, цей підхід здатен ефективно виявляти нові зв'язки між

користувачами та об'єктами, які можуть бути неочевидними при використанні інших методів, таких як фільтрація за змістом. Проте, як і будь-який інший метод, колаборативна фільтрація має певні обмеження, зокрема проблему холодного старту. Коли система стикається з новими користувачами або новими елементами, вона не має достатньо інформації для визначення їхніх уподобань і подібності, що може призвести до неточних або навіть неадекватних рекомендацій. Іншою проблемою є відомий феномен «перегрузки інформацією» (information overload). Коли система отримує надмірну кількість взаємодій, вона може створити величезну кількість рекомендацій, що не дозволяє користувачу легко знайти справді корисні для нього елементи. Для вирішення цієї проблеми в сучасних системах часто використовують методи фільтрації та агрегації, щоб зменшити кількість рекомендацій до найбільш релевантних і значущих.

Оскільки колаборативна фільтрація базується на вивченні поведінки користувачів, то вона також стикається з проблемою маніпуляцій і шахрайства, коли користувачі або навіть системи можуть штучно змінювати свої рейтинги або оцінки, щоб вплинути на результати рекомендацій. Для боротьби з цим вживаються різноманітні методи виявлення аномалій та фільтрації підозрілих даних.

Колаборативна фільтрація активно використовується в практиці, зокрема в таких системах, як Amazon, Netflix, Spotify та багатьох інших, де користувачам пропонуються рекомендації на основі їхніх попередніх уподобань або на основі схожості з іншими користувачами. Цей підхід не тільки забезпечує високий рівень персоналізації, але й дозволяє відкривати нові інтереси та можливості для користувачів, що сприяє їхній більш активній взаємодії з платформами. Однак, як і всі інші технології, колаборативна фільтрація потребує постійного вдосконалення та адаптації до нових умов і потреб користувачів.



## 2.5 Гібридні рекомендаційні системи

Гібридні рекомендаційні системи є важливим напрямом в області розробки інтелектуальних систем для надання персоналізованих рекомендацій користувачам. Ці системи поєднують різні підходи та методи рекомендацій, прагнучи подолати обмеження окремих технологій, що використовуються в традиційних системах, таких як колаборативна фільтрація або фільтрація за змістом [15]. Гібридні системи дозволяють створювати більш ефективні, точні та персоналізовані рекомендації, оскільки вони інтегрують кілька джерел інформації або використовують різні алгоритми для досягнення кращих результатів. Така інтеграція може бути здійснена різними способами, що забезпечує великий спектр можливостей для налаштування рекомендацій під специфічні вимоги користувачів та контексту.

Основна ідея гібридних рекомендаційних систем полягає в поєднанні різних методів рекомендацій, таких як колаборативне фільтрування, фільтрація за змістом, методи на основі контексту, а також алгоритми машинного навчання, щоб забезпечити високоточні рекомендації. Це дозволяє компенсувати слабкі сторони кожного з окремих методів, таких як проблема холодного старту, обмеженість даних або схильність до створення надто одноманітних рекомендацій. Гібридизація методів дозволяє системам адаптуватися до змін у поведінці користувачів, враховувати різні аспекти контексту і оптимізувати процес надання рекомендацій, зберігаючи при цьому високий рівень точності.

Згідно з класичним підходом, гібридні системи можуть поєднувати методи рекомендацій за допомогою змішування результатів, обробки даних на різних етапах або використання спільних моделей, що охоплюють декілька типів фільтрацій. Існує кілька способів, як це може бути реалізовано

в практиці, що дозволяє досягти гнучкості та високої ефективності в рекомендаціях.

Перший спосіб інтеграції методів у гібридних системах передбачає змішування результатів різних алгоритмів. У цьому випадку, після того як різні методи рекомендацій (наприклад, колаборативне фільтрування та фільтрація за змістом) обчислюють свої власні набори рекомендацій, система комбінує ці результати для отримання кінцевої рекомендації для користувача. Існує кілька технік для комбінування результатів. Одна з найбільш поширених – це просте зважене середнє, де кожному методу надається певна вага, що відображає його важливість або точність у конкретному контексті [16]. Наприклад, якщо система вважає, що для конкретного користувача найбільше значення мають його історичні вподобання, то ваги колаборативного фільтрування будуть більшими, а ваги фільтрації за змістом – меншими. Інший варіант – це більш складні алгоритми, такі як алгоритми машинного навчання, які намагаються автоматично знаходити оптимальні ваги для кожного методу.

Інший підхід до гібридизації передбачає комбінування методів на етапі обробки даних. У цьому випадку система може поєднувати дані з різних джерел або використовувати результати різних підходів до фільтрації для створення єдиного профілю користувача або для покращення точності прогнози. Наприклад, замість того щоб просто обчислювати рекомендації на основі окремих підходів, система може на початкових етапах обробки даних комбінувати інформацію про історичну поведінку користувача, його демографічні характеристики, контекстуальні дані та атрибути об'єктів, з якими він взаємодіє. Цей комбінований профіль потім використовується для генерування рекомендацій. Наприклад, у комерційних системах, таких як онлайн-магазини, часто використовується гібридна модель, що поєднує дані про покупку користувача (колаборативне фільтрування), з відомостями про

характеристики товарів (фільтрація за змістом) і зовнішніми факторами, такими як сезонність або спеціальні акції.

Третій підхід до гібридизації полягає в тому, щоб використовувати одні й ті ж дані, але застосовувати різні методи для прогнозування або класифікації. Наприклад, система може застосовувати колаборативне фільтрування для прогнозування рейтингів або переваг, одночасно використовуючи фільтрацію за змістом для рекомендацій на основі характеристик об'єктів. Це дозволяє використовувати найкраще з обох методів для створення більш точних рекомендацій, особливо в ситуаціях, коли один метод може бути недостатньо ефективним сам по собі.

Однією з головних переваг гібридних систем є їх здатність зменшувати проблеми, з якими стикаються традиційні методи. Наприклад, одна з найбільш відомих проблем колаборативного фільтрування – це проблема холодного старту, коли система не має достатньо даних для ефективного прогнозування вподобань нових користувачів або нових об'єктів. Гібридні системи можуть частково або повністю вирішити цю проблему, комбінуючи методи, які не залежать від великої кількості взаємодій [17]. Наприклад, в поєднанні з фільтрацією за змістом, система може рекомендувати нові об'єкти на основі їхніх властивостей, навіть якщо користувач ще не взаємодіяв з ними. Також гібридні системи дозволяють компенсувати обмеження фільтрації за змістом, де труднощі можуть виникнути через складність виділення релевантних ознак або проблеми з недостатньою інформацією про об'єкти. Крім того, гібридні системи дозволяють досягти більш високої точності рекомендацій, що є важливим для користувачів. Комбінуючи різні методи, можна підвищити різноманітність рекомендацій, забезпечити більшу персоналізацію та знизити ймовірність того, що система буде пропонувати користувачам надто стандартний або одноманітний контент. Наприклад, в електронній комерції використання гібридних систем дозволяє комбінувати дані про покупку користувачем товарів з даними про

схожість цих товарів з іншими товарами, що дозволяє запропонувати більше цікавих варіантів покупок.

Проте, гібридні системи мають і свої недоліки. По-перше, вони часто потребують значних обчислювальних ресурсів і складніші в розробці, оскільки вимагають інтеграції та налаштування кількох методів і моделей. Це може збільшити витрати на розробку та підтримку таких систем. По-друге, навіть при поєднанні різних методів, система може все одно стикатися з певними труднощами в умовах обмежених даних, швидких змін в інтересах користувачів або проблемах з управлінням великими обсягами даних. У таких випадках ефективність комбінованих моделей може значно знижуватися, якщо не забезпечується відповідна оптимізація або оновлення системи.

Не менш важливим аспектом гібридних рекомендаційних систем є їх здатність адаптуватися до конкретних доменів та сценаріїв використання. Так, в різних галузях застосування рекомендаційних систем (наприклад, в медіа, електронній комерції, охороні здоров'я або освіті) можуть використовуватися різні підходи до комбінування методів. Для медіа, наприклад, ключовими є не тільки інтереси користувача, але й контекстуальні фактори, такі як час перегляду або географічне місце розташування, що також можуть бути враховані в рекомендаціях. У випадку з електронною комерцією важливу роль відіграють не тільки історія покупок користувача, але й взаємодія з іншими користувачами, їхні оцінки продуктів, відгуки і навіть схожі покупки інших клієнтів. Тому успішне застосування гібридних систем вимагає гнучкості і здатності підлаштовуватися під специфічні умови кожного конкретного випадку.

Таким чином, гібридні рекомендаційні системи є важливим етапом розвитку технологій рекомендацій, дозволяючи поєднувати сильні сторони різних підходів і значно покращувати якість і точність рекомендацій. Вони здатні не лише знижувати обмеження традиційних методів, але й

відкривають нові можливості для персоналізації, адаптації до змінних умов і розширення сфери застосування рекомендаційних систем. Водночас, гібридизація вимагає ретельного налаштування та підтримки системи, що додає певні складнощі, але забезпечує можливість створення високоефективних і потужних рекомендаційних інструментів.

## 2.6 Алгоритми машинного навчання в рекомендаційних системах

Алгоритми машинного навчання відіграють важливу роль в розробці сучасних рекомендаційних систем. Рекомендаційні системи, що активно застосовуються в різних сферах, таких як електронна комерція, соціальні мережі, медіа платформи, а також в багатьох інших галузях, використовують ці алгоритми для надання персоналізованих рекомендацій, що відповідають інтересам і потребам користувачів [18]. Традиційні методи рекомендацій, такі як колаборативна фільтрація та фільтрація за змістом, досить ефективні, але мають певні обмеження, зокрема проблеми холодного старту, обмеження в контексті або неспроможність правильно обробляти великий обсяг даних. Застосування машинного навчання дозволяє вирішувати ці проблеми та значно покращувати якість рекомендацій.

Машинне навчання, як підхід до створення інтелектуальних систем, ґрунтується на ідеї, що моделі можуть бути навчені на основі даних, без необхідності явного програмування або визначення всіх правил. У випадку з рекомендаційними системами, це означає, що алгоритми можуть виявляти приховані патерни, структури та взаємозв'язки між користувачами та об'єктами, що дає змогу надавати більш точні і релевантні рекомендації. Машинне навчання в рекомендаційних системах можна умовно поділити на три основні категорії: навчання з учителем, навчання без учителя та навчання з підкріпленням. Кожен із цих підходів має свої особливості і застосовується в залежності від задач, що ставляться перед системою.

Перший підхід, навчання з учителем, є одним з найбільш використовуваних у рекомендаційних системах. Основна ідея навчання з учителем полягає в тому, що алгоритм навчається на основі міток або відомих даних, що описують взаємодії між користувачами та об'єктами. Це може бути, наприклад, рейтинг або оцінка, яку користувач ставить товару або фільму. На основі цих даних модель створює функцію, яка може прогнозувати оцінки або ймовірність того, що користувач буде зацікавлений у тому чи іншому об'єкті. Алгоритми машинного навчання, що використовуються для цього підходу, можуть бути різними, від простих лінійних моделей до складних нейронних мереж.

Один з найбільш популярних алгоритмів навчання з учителем у рекомендаційних системах – це методи регресії [18]. У випадку рекомендацій, такі моделі можуть бути використані для передбачення оцінок, які користувач надасть певному елементу на основі його попередніх оцінок або інтересів. Метод регресії може також використовувати інші ознаки, такі як вік користувача, місце розташування або інші демографічні характеристики. Однак для складніших випадків, де взаємодії користувача з об'єктами мають більш складну і нелінійну природу, простих лінійних моделей часто недостатньо, і в такому випадку використовуються більш потужні алгоритми, такі як методи підтримки векторів, дерева рішень або навіть глибоке навчання.

Методи підтримки векторів (Support Vector Machines, SVM) є одним із найбільш ефективних підходів у задачах рекомендацій, де необхідно класифікувати або регресувати дані з високою кількістю ознак. Ці алгоритми базуються на побудові гіперплощини, яка максимально відокремлює класи в багатовимірному просторі ознак. У контексті рекомендаційних систем методи SVM можуть використовуватися для визначення того, чи сподобається користувачу певний об'єкт, на основі навчання на вже відомих оцінках. Оскільки багато рекомендаційних систем мають справу з

високорозмірними даними, методи SVM добре підходять для таких завдань [18].

Для більш складних і масштабних рекомендаційних задач, де необхідна обробка великих обсягів даних, особливо в умовах наявності нерелевантних або відсутніх даних, часто використовуються нейронні мережі. Глибоке навчання (deep learning) виявилось дуже ефективним для побудови рекомендаційних систем, оскільки нейронні мережі здатні обробляти і вивчати складні патерни і взаємозв'язки між користувачами та об'єктами, навіть у випадках, коли дані мають високий рівень шуму або відсутність значення. Однією з популярних архітектур є багат шарові нейронні мережі, які складаються з кількох шарів обчислювальних нейронів, що дозволяють здійснювати ефективну модель навчання та виявлення схованих взаємозв'язків. Такі мережі добре справляються з проблемою холодного старту, оскільки здатні навчатися на великому обсязі даних і знаходити загальні патерни, що дозволяє здійснювати рекомендації навіть для нових користувачів або об'єктів.

Однак, щоб нейронні мережі були ефективними в контексті рекомендацій, вони повинні бути добре налаштовані та оптимізовані, оскільки вони мають схильність до перевищення або недонавчання (overfitting або underfitting), особливо в разі обмежених даних. Для цього використовуються різні методи, такі як регуляризація, додаткові шари для контролю складності мережі, або застосування технік, які дозволяють уникати перенавчання. Інший напрямок використання машинного навчання в рекомендаційних системах полягає в застосуванні навчання без учителя [17, 18]. Тут алгоритми намагаються виявити структуру даних без використання міток або явно визначених цілей. У таких системах не використовуються явні оцінки або інші ознаки, що дають чітке уявлення про те, чому користувач обирає ті чи інші об'єкти. Натомість алгоритм намагається самостійно знайти групи користувачів або об'єктів з подібними характеристиками або

інтересами. Для цього використовуються методи кластеризації, такі як k-середніх, ієрархічна кластеризація або алгоритми на основі аналізу головних компонент (PCA).

Методи кластеризації, наприклад, дозволяють поділити користувачів на групи на основі їхніх спільних інтересів або поведінки, що дозволяє створювати групові рекомендації. Це особливо корисно в ситуаціях, коли немає достатньо індивідуальних даних про кожного користувача, але є можливість класифікувати їх за групами або сегментами. Використання таких методів дозволяє значно зменшити обсяг обчислень і поліпшити точність рекомендацій, оскільки групи користувачів часто демонструють схожі патерни поведінки.

Третій підхід до використання машинного навчання в рекомендаційних системах – це навчання з підкріпленням. Цей підхід є набагато більш складним і полягає в оптимізації дій агента на основі досвіду взаємодії з середовищем [16-18]. У контексті рекомендаційних систем агент може представляти собою користувача або саму систему, яка навчається, а середовище – це набір об'єктів або контенту, з якими користувач може взаємодіяти. Алгоритм навчання з підкріпленням допомагає вибрати оптимальні стратегії рекомендацій, щоб максимізувати довгострокову вигоду для користувача, таку як підвищення залучення або досягнення вищих оцінок від користувачів. Алгоритми на основі підкріплення, такі як Q-навчання або методи глибокого навчання з підкріпленням (Deep Reinforcement Learning), можуть навчатися адаптувати свої стратегії в залежності від реакцій користувачів, забезпечуючи більшу персоналізацію та більш релевантні рекомендації в реальному часі.

Загалом, використання алгоритмів машинного навчання в рекомендаційних системах дозволяє створювати потужні, ефективні та гнучкі системи, які можуть працювати з великими обсягами даних, адаптуватися до змін в інтересах користувачів і забезпечувати високоточні персоналізовані



рекомендації. Однак, незважаючи на всі переваги, ці методи мають і свої недоліки, такі як потреба у великій кількості даних для ефективного навчання, складність налаштування моделей та обчислювальна вартість. Тому важливою задачею є не лише використання правильних алгоритмів, але й оптимізація процесу навчання, управління даними та створення механізмів, що забезпечують ефективну і точну роботу рекомендаційних систем в реальних умовах.

## 2.7 Висновки до другого розділу

У розділі досліджуються різні підходи до створення рекомендаційних систем, включаючи фільтрацію за змістом, колаборативну фільтрацію та гібридні системи. Ці методи є основними при розробці персоналізованих рекомендацій, але кожен з них має свої переваги і недоліки, що зумовлює необхідність комбінування різних підходів для досягнення кращих результатів. Використання алгоритмів машинного навчання в рекомендаційних системах є важливою частиною сучасних підходів до створення ефективних і адаптивних систем. Алгоритми машинного навчання, зокрема методи регресії, класифікації та навчання з підкріпленням, дозволяють моделювати складні патерни взаємодій між користувачами та об'єктами, що дозволяє значно покращити точність рекомендацій та адаптувати систему до змін в інтересах користувачів.

Гібридні системи, що поєднують різні підходи, є важливою тенденцією в розробці рекомендаційних систем. Вони дозволяють не тільки обирати кращі методи для різних контекстів, але й знижують обмеження окремих технологій, таких як проблема холодного старту чи нестача даних для точних прогнозів. Основною метою рекомендаційних систем є досягнення високої точності в рекомендаціях та персоналізація сервісів для користувачів. Розробка методів, які можуть більш ефективно враховувати інтереси і

поведінку користувачів, забезпечує краще залучення і покращення користувацького досвіду.

Враховуючи використання складних алгоритмів, таких як нейронні мережі та методи глибокого навчання, процес розробки рекомендаційних систем стає ресурсномістким і вимагає значних обчислювальних потужностей. Задля ефективної роботи системи необхідно враховувати обмеження, що пов'язані з оптимізацією моделей і вибором алгоритмів, які б забезпечували швидку реакцію і точність в реальному часі. Рекомендаційні системи повинні бути здатні адаптуватися до змін в інтересах користувачів та обробляти великі обсяги даних, що є характерною проблемою для сучасних додатків і платформ. Це підкреслює важливість гнучких моделей і алгоритмів, які можуть масштабуватися без втрати ефективності.

## РОЗДІЛ 3. ПРОГРАМНА РЕАЛІЗАЦІЯ

### 3.1 Мета розроблення

Метою розроблення є багатофункціональна система рекомендацій для платформи електронної комерції, яка включатиме такі основні можливості:

1) генерація даних:

– створення набору даних для продуктів з детальними характеристиками, такими як назва, опис, категорія, ціна, рейтинг та кількість на складі;

– формування взаємодій користувачів із продуктами, включаючи перегляди, ймовірність покупки, рейтинг продуктів користувачами тощо;

2) алгоритми рекомендацій:

– контент-орієнтовані рекомендації: пошук подібних продуктів на основі текстового опису за допомогою TF-IDF та косинусної схожості;

– персоналізовані рекомендації для користувача: аналіз улюблених категорій користувача та пропонування найбільш релевантних продуктів;

3) кластеризація та аналітика:

– розподіл продуктів на ринкові сегменти за характеристиками (ціна, рейтинг) за допомогою алгоритму кластеризації k-means;

– підготовка статистичного аналізу та візуалізацій для дослідження розподілу продуктів за категоріями, зв'язку між ціною та рейтингом, та інших кореляційних даних;

4) автоматичне створення графіків для аналізу ринку (розподіл категорій, кореляції, гістограми цін тощо);

5) формування звітів:

- підготовка інтерактивного звіту про стан ринку, включаючи статистику продуктів, аналіз категорій, цінові тренди та інше;

- включення сегментації ринку та надання списку найбільш рейтингових продуктів;

б) генерація персоналізованих рекомендацій для заданих ідентифікаторів користувачів.

Додаткові умови:

- використовувати сучасні бібліотеки Python (Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn) для аналізу та машинного навчання;

- забезпечити легкість адаптації та розширення функціоналу для інших типів платформ електронної комерції;

- програмний код має бути написаний у стилі, що забезпечує зрозумілість, структурованість і відповідність принципам об'єктно-орієнтованого програмування.

### 3.2 Архітектура системи

Опис архітектури розробленого програмного коду включає декілька ключових аспектів, які охоплюють модульність, структуру класів, використання бібліотек та застосування алгоритмів. Код розроблений у межах об'єктно-орієнтованого підходу, що дозволяє забезпечити високу масштабованість, легкість у розширенні функціоналу та підтримці. Основною одиницею архітектури є клас `AdvancedRecommendationSystem`, який слугує ядром для виконання всіх основних завдань: генерації даних, рекомендацій, кластеризації, аналітики та візуалізації. Клас `AdvancedRecommendationSystem` організований таким чином, щоб інкапсулювати основну бізнес-логіку, пов'язану з електронною комерцією. Його ініціалізація включає створення двох основних наборів даних: одного для опису продуктів, іншого для моделювання взаємодій користувачів. Ці

набори даних створюються за допомогою внутрішніх методів `_generate_rich_products` та `_generate_comprehensive_interactions`. У процесі генерації продуктів використовується набір категорій, таких як «Електроніка», «Одяг», «Книги» тощо, та різні діапазони цін, що дозволяє створити різномірний і реалістичний набір даних. Продукти характеризуються такими атрибутами, як унікальний ідентифікатор, назва, опис, категорія, ціна, рейтинг та кількість одиниць на складі. Це забезпечує не лише глибину, але й багатовимірність даних, що є ключовим для подальшого аналізу.

Метод `_generate_comprehensive_interactions` відповідає за створення моделі поведінки користувачів на платформі. Кожен користувач характеризується унікальним бюджетом, який враховується під час вибору продуктів для взаємодії. Взаємодії моделюються як записи про перегляди продуктів, ймовірність їх покупки, оцінки та час перегляду. Ці параметри дають змогу створити детальну картину поведінки користувачів, що є важливим для рекомендаційної системи.

Контент-орієнтовані рекомендації реалізовані через метод `content_based_recommendations`. Для обробки текстових даних використовується `TfidfVectorizer` з бібліотеки `scikit-learn`. Опис кожного продукту обробляється за допомогою TF-IDF матриці, що дозволяє створити числове представлення тексту. Косинусна схожість розраховується для визначення ступеня подібності між продуктами. Таким чином, користувачеві можуть бути запропоновані продукти, що найбільш схожі на той, який він переглядав або придбав. Це робить систему здатною надавати релевантні рекомендації навіть для нових користувачів чи продуктів.

Окремим важливим аспектом є метод кластеризації продуктів, що реалізований у методі `market_segmentation`. За основу береться кластеризація K-Means, що дозволяє сегментувати продукти на основі цінкових характеристик та рейтингів. Перед застосуванням кластеризації дані

нормалізуються за допомогою `StandardScaler`, що забезпечує адекватність роботи алгоритму. Результати кластеризації використовуються для подальшого аналізу та формування звітів, а також для кращого розуміння структури ринку.

Метод `generate_visualizations` відповідає за створення графічних інтерпретацій аналітичних даних. Це включає побудову гістограм розподілу цін, кореляційних матриць, графіків залежності ціни від рейтингу тощо. Для побудови графіків використовуються бібліотеки `Matplotlib` та `Seaborn`, що дозволяє створювати візуалізації високої якості. Усі побудовані графіки зберігаються у вигляді файлів, що може бути корисним для презентацій або подальшого аналізу.

Формування маркетингового звіту реалізоване у методі `generate_market_report`. У цьому методі обчислюється статистика для різних категорій продуктів, аналізуються цінові показники, визначаються найбільш рейтингові продукти. Крім того, звіт включає аналіз взаємодій користувачів, середню кількість взаємодій на одного користувача, а також результати кластеризації ринку. Усі ці дані організовані у зручному для читання форматі, що дозволяє швидко оцінити стан платформи та зробити стратегічні висновки.

Персоналізовані рекомендації для користувачів реалізовані через метод `personalized_user_recommendations`. Цей метод аналізує історію взаємодій конкретного користувача, виявляє його улюблені категорії та формує список рекомендацій на основі рейтингів та цін продуктів у цих категоріях. Такий підхід дозволяє врахувати індивідуальні уподобання користувачів, що значно підвищує релевантність рекомендацій. Для кожного користувача генерується список із зазначенням назв продуктів, їх категорій, цін та рейтингів, що забезпечує максимальну інформативність.

Окрім основних функціональних блоків, код містить модуль для тестування, який реалізований у функції `main`. Ця функція виконує декілька

ключових завдань: генерує аналітичні графіки, формує маркетинговий звіт, тестує персоналізовані рекомендації для декількох користувачів. Це дозволяє оцінити повноту функціоналу системи, перевірити її коректність та продемонструвати результати роботи. Зручність тестування забезпечується використанням бібліотеки `tabulate`, що дозволяє відображати дані у форматі таблиць з високим рівнем читабельності.

Архітектура програмного забезпечення базується на модульності та ієрархічності, що дозволяє чітко розділити відповідальності між різними компонентами системи. Клас `AdvancedRecommendationSystem` виступає в якості центрального об'єкта, який координує всі процеси, починаючи від генерації даних і закінчуючи створенням графіків та звітів. Взаємодія між різними методами організована таким чином, щоб забезпечити максимальну гнучкість і повторне використання коду. Наприклад, дані, що генеруються методами `_generate_rich_products` та `_generate_comprehensive_interactions`, використовуються у декількох інших методах для аналізу, кластеризації та побудови рекомендацій.

Таким чином, розроблений програмний код є комплексним рішенням для автоматизації аналізу даних і формування рекомендацій в межах платформи електронної комерції. Його архітектура побудована на сучасних підходах до розробки програмного забезпечення, що дозволяє забезпечити високу ефективність, масштабованість і адаптивність до нових вимог.

### 3.3 Опис використаних технологій

Використані технології та інструменти у розробці програмного забезпечення відіграють ключову роль у забезпеченні функціональності, продуктивності та якості результатів. У межах розробки проекту були залучені сучасні підходи до роботи з даними, інструменти для обробки, аналізу та візуалізації, а також бібліотеки та мови програмування, які

сприяють створенню ефективних і гнучких рішень. Основною мовою програмування було обрано Python завдяки його універсальності, багатству доступних бібліотек, а також активній підтримці спільноти розробників. Python дозволив реалізувати ключові функціональні блоки системи, включаючи генерацію даних, рекомендаційні алгоритми, кластеризацію, створення візуалізацій та автоматизацію формування звітів.

Одним із основних аспектів розробки було забезпечення обробки текстової інформації, що використовувалася для опису продуктів у наборі даних. Для цього застосовувався інструмент `TfidfVectorizer` з бібліотеки `scikit-learn`, який дозволив конвертувати текстові описи продуктів у числові представлення, що враховують частоту вживання слів. Цей підхід є надзвичайно важливим для реалізації контент-орієнтованих рекомендацій, оскільки забезпечує розрахунок подібності між продуктами на основі їхніх описів. Використання косинусної схожості дозволило підвищити точність рекомендацій, оскільки цей метод ефективно враховує різні масштаби векторів.

Для роботи з числовими та категорійними даними було використано бібліотеку `pandas`, яка забезпечила ефективну маніпуляцію великими обсягами інформації. Її можливості включають об'єднання, фільтрацію, сортування та агрегування даних. Це стало основою для обробки наборів даних про продукти та взаємодії користувачів. Взаємодія між користувачами та продуктами, а також аналіз цих взаємодій, моделювалися шляхом збереження інформації в табличній структурі, що дозволяло легко здійснювати операції групування та статистичного аналізу. Завдяки цьому можна було з легкістю реалізувати персоналізовані рекомендації для користувачів, що враховують їхню історію поведінки.

Для нормалізації даних перед кластеризацією продуктів була використана бібліотека `scikit-learn`, зокрема клас `StandardScaler`, який забезпечує приведення даних до єдиного масштабу. Це було необхідно для



коректної роботи алгоритму кластеризації K-Means, який є чутливим до варіацій у масштабах числових характеристик. Кластеризація дозволила сегментувати ринок продуктів на основі цінових характеристик та рейтингів, забезпечуючи глибше розуміння структури даних та виявлення закономірностей у їх розподілі.

Щодо роботи з візуалізаціями, було використано дві бібліотеки: `matplotlib` та `seaborn`. Перша слугувала базовим інструментом для побудови діаграм, графіків розподілу та часових рядів, тоді як друга забезпечила високий рівень естетики графіків і спрощувала їхню побудову завдяки зручному синтаксису. Це дозволило створювати наочні візуалізації для аналізу розподілу цін, кореляційних залежностей між характеристиками продуктів і взаємодій користувачів із платформою. Окрім того, бібліотека `seaborn` забезпечила створення теплових карт кореляцій, які є надзвичайно важливими для виявлення взаємозв'язків між різними параметрами даних.

В аспекті автоматизації формування звітів було застосовано бібліотеку `tabulate`, яка дозволила організувати дані у вигляді таблиць, що легко читаються. Це стало ключовим для створення текстових звітів, які включали статистичну інформацію про категорії продуктів, аналіз взаємодій користувачів та результати кластеризації ринку. Форматування даних у вигляді таблиць сприяло їх легкому сприйняттю та інтеграції в аналітичні документи.

Для реалізації персоналізованих рекомендацій використовувалася інтеграція між обробкою історії взаємодій користувачів і генерацією релевантних продуктів. Алгоритми рекомендацій базувалися на аналізі улюблених категорій користувачів, а також їхніх бюджетних обмежень. Це стало можливим завдяки синергії кількох інструментів, включаючи бібліотеку `numpy` для високошвидкісних обчислень і бібліотеку `random` для моделювання випадкових виборів продуктів. Крім того, у процесі розробки була забезпечена інтеграція елементів машинного навчання. Хоча основна

частина роботи була зосереджена на алгоритмах кластеризації та обчисленнях подібності, сам вибір цих алгоритмів ґрунтувався на найкращих практиках машинного навчання, реалізованих у бібліотеці `scikit-learn`. Завдяки цьому вдалося створити моделі, які легко адаптуються до нових даних і забезпечують високу точність аналізу.

Невід'ємною частиною проекту стало моделювання даних, яке здійснювалося через створення синтетичних наборів даних. Цей процес був автоматизований через використання генераторів випадкових чисел і тексту. Для генерації текстових описів продуктів використовувалися бібліотеки на основі попередньо збережених шаблонів, що дозволило створити реалістичні описи з різними рівнями деталізації. Крім того, генерація взаємодій користувачів з продуктами враховувала ймовірнісні моделі вибору, що значно підвищило правдоподібність згенерованих даних.

Таким чином, використання сучасних технологій та інструментів забезпечило комплексність і якість реалізованого рішення. Поєднання можливостей бібліотек `pandas`, `numpy`, `scikit-learn`, `matplotlib`, `seaborn` та `tabulate` створило міцну основу для виконання складних аналітичних завдань і розробки функціональних рекомендаційних систем. Завдяки цьому вдалося реалізувати архітектуру, яка не лише відповідає поточним вимогам, а й має високу гнучкість для інтеграції нових функцій у майбутньому.

### 3.4 Процес навчання та представлення рекомендацій

Процес роботи розробленого програмного забезпечення характеризується багаторівневою структурою, яка забезпечує ефективність обробки даних, реалізацію алгоритмів кластеризації, рекомендацій та візуалізації. Весь процес поділяється на декілька етапів, які тісно взаємопов'язані та утворюють єдину систему для досягнення поставлених цілей. Програмне забезпечення розпочинає свою роботу із завантаження

початкових даних, які є основою для подальшого аналізу. Ці дані можуть бути представлені у вигляді табличних структур, що зберігаються у форматах CSV або інших популярних форматах, які легко інтегруються із середовищем Python. Завантаження виконується за допомогою бібліотеки `pandas`, яка дозволяє не лише швидко зчитати дані, але й одразу провести їх базову обробку. Зокрема, це стосується очищення даних від пропущених значень, видалення дублікатів та нормалізації числових параметрів.

Після завантаження і підготовки даних розпочинається наступний ключовий етап – перетворення текстової інформації у формат, придатний для аналізу. У нашому випадку текстові описи продуктів обробляються за допомогою методу TF-IDF (term frequency-inverse document frequency), який дозволяє зважувати значення кожного слова залежно від його важливості у певному контексті. Цей процес виконується за допомогою інструментів бібліотеки `scikit-learn`. У результаті текстові описи перетворюються на вектори, які можуть бути проаналізовані за допомогою алгоритмів машинного навчання. Застосування TF-IDF є критично важливим для забезпечення ефективної роботи рекомендаційної системи, оскільки саме текстові характеристики продуктів слугують базисом для визначення подібності між ними.

На наступному етапі проводиться обробка числових і категорійних даних, які стосуються характеристик продуктів, таких як ціна, рейтинг або категорія. Для цього використовуються можливості `pandas` для групування, сортування та агрегації даних. Особливу увагу приділяється нормалізації числових характеристик за допомогою `StandardScaler` з бібліотеки `scikit-learn`. Ця операція забезпечує уніфікацію даних, що є необхідною умовою для коректної роботи алгоритму кластеризації K-Means.

Кластеризація є одним із центральних компонентів роботи програмного забезпечення. Вона дає змогу поділити продукти на групи за схожими характеристиками, що сприяє аналізу їхньої структури. Алгоритм K-Means

застосовується для поділу продуктів на кластери залежно від їхніх цінових характеристик та рейтингу. Результати кластеризації слугують основою для виявлення закономірностей у ринку продуктів, а також для покращення рекомендаційної системи. У цьому контексті кожен кластер може бути інтерпретований як окремий сегмент ринку, що характеризується певними особливостями, наприклад, низьким рівнем цін або високими оцінками користувачів.

Особливістю програмного забезпечення є його здатність до персоналізованих рекомендацій. Для цього обробляється історія взаємодій користувачів із продуктами, зокрема їхні вподобання та обмеження за бюджетом. Ця інформація інтегрується із даними про продукти, дозволяючи створювати індивідуальні рекомендації для кожного користувача. У цьому процесі ключову роль відіграє обчислення схожості між продуктами. Векторні представлення продуктів, отримані за допомогою TF-IDF, використовуються для обчислення косинусної схожості. Це дозволяє знайти продукти, які найбільш відповідають інтересам користувача. Додатково рекомендаційна система враховує кластери, до яких належать продукти, щоб забезпечити релевантність рекомендацій у межах конкретного сегмента ринку. Наступним аспектом є генерація синтетичних даних, яка використовується для тестування алгоритмів і забезпечення їхньої універсальності. Генерація даних виконується за допомогою інструментів Python, зокрема бібліотек `random` та `numpy`, які дозволяють створювати реалістичні моделі поведінки користувачів. Наприклад, ціна продуктів може генеруватися у певному діапазоні з урахуванням розподілу, близького до нормального, тоді як категорії продуктів можуть визначатися випадковим чином із заданої множини.

Після виконання всіх етапів обробки та аналізу даних програмне забезпечення переходить до етапу візуалізації результатів. Для цього використовуються бібліотеки `matplotlib` і `seaborn`, які забезпечують широкий

спектр можливостей для створення графіків. Наприклад, розподіл цін між продуктами можна візуалізувати у вигляді гістограми, а результати кластеризації – у вигляді діаграми розсіювання, де кожен кластер має свій унікальний колір. Візуалізації сприяють кращому розумінню отриманих результатів та їх інтерпретації. Важливо зазначити, що графічний вихід оптимізований для інтеграції у звіти, які можуть бути створені автоматично.

Процес автоматизації формування звітів є окремим компонентом програмного забезпечення. Звіти створюються у текстовому форматі та містять таблиці з ключовими показниками, такими як середні ціни у кластерах, кількість продуктів у кожній категорії та аналіз взаємодій користувачів. Таблиці генеруються за допомогою бібліотеки *tabulate*, яка дозволяє створювати структуровані дані у вигляді, зручному для читання. Ці звіти можуть бути використані для презентацій, аналітичних досліджень або як основа для прийняття бізнес-рішень.

Таким чином, розроблене програмне забезпечення забезпечує багатоступеневий процес аналізу даних, який включає завантаження, обробку, кластеризацію, рекомендації та візуалізацію. Кожен етап тісно інтегрований із наступним, що дозволяє досягти високого рівня ефективності та якості. Це програмне забезпечення є гнучким, універсальним і може бути легко адаптоване для використання у різних галузях.

### 3.5 Інтеграція з іншими інструментами

Здатність інтегрувати розроблене програмне забезпечення з іншими інструментами та платформами є одним із ключових аспектів, що визначають його універсальність, гнучкість і практичну цінність у різних галузях. Інтеграція забезпечує можливість об'єднання функціоналу програмного забезпечення з іншими системами для створення більш масштабних, продуктивних і комплексних рішень. Така інтеграція реалізується через

добре продумані технічні механізми, які забезпечують взаємодію як на рівні передачі даних, так і на рівні обробки та взаємного виклику функцій між різними системами.

Одним із основних підходів до забезпечення інтеграції є використання інтерфейсів прикладного програмування (API). Розроблене програмне забезпечення підтримує створення RESTful API, які базуються на принципах архітектури клієнт-сервер. Це забезпечує простоту та стандартизованість процесу інтеграції. Наприклад, інструмент FastAPI, застосований у розробці, дозволяє створювати високопродуктивні API-інтерфейси, які підтримують взаємодію з іншими системами через HTTP-запити. Завдяки цьому інші платформи можуть відправляти запити для отримання інформації, запуску процесів обробки даних або виконання аналітичних завдань. Це дозволяє використовувати програмне забезпечення як мікросервісну архітектуру в межах більших систем, таких як ERP або CRM.

Для забезпечення інтеграції з базами даних різних типів програмне забезпечення реалізує підтримку декількох драйверів для роботи з реляційними та нереляційними базами даних. Зокрема, бібліотека SQLAlchemy забезпечує гнучку ORM (Object-Relational Mapping) для роботи з реляційними базами даних, такими як PostgreSQL, MySQL або SQLite. Це дає змогу інтегрувати програму з існуючими базами даних підприємства, виконуючи обмін інформацією або її оновлення без необхідності змінювати внутрішню структуру зберігання. Для роботи з нереляційними базами даних, такими як MongoDB, застосовується бібліотека PyMongo, яка дозволяє оперувати даними у форматі JSON. Така гнучкість забезпечує адаптацію програмного забезпечення до будь-якого типу баз даних, що є особливо важливим для різних галузей, де вимоги до зберігання даних суттєво відрізняються. Ще одним важливим аспектом інтеграції є можливість підключення до хмарних сервісів, таких як Amazon Web Services (AWS), Google Cloud Platform (GCP) або Microsoft Azure. Використання SDK

відповідних платформ дозволяє реалізовувати автоматичне завантаження даних до хмарних сховищ, використання обчислювальних ресурсів для виконання складних алгоритмів або інтеграцію із сервісами машинного навчання. Наприклад, інструменти AWS S3 надають можливість зберігати великі обсяги даних у хмарному сховищі, що є особливо корисним для обробки масштабних наборів даних. Інтеграція з хмарними платформами також забезпечує можливість масштабування, дозволяючи адаптувати програмне забезпечення до зростаючих обсягів інформації. Крім цього, реалізовано підтримку взаємодії з іншими популярними інструментами для обробки даних, такими як Apache Kafka, який забезпечує можливість обробки поточкових даних у реальному часі. Це дозволяє інтегрувати програмне забезпечення у системи, які працюють із даними, що надходять у режимі реального часу, наприклад, у фінансових установах для моніторингу транзакцій або у системах розумних будинків для аналізу показників датчиків. Підключення до таких інструментів здійснюється через спеціалізовані бібліотеки Python, які дозволяють обмінюватися даними між сервісами, гарантуючи високу продуктивність і низьку затримку.

Інтеграція з інструментами бізнес-аналітики, такими як Tableau або Power BI, реалізується через можливість експорту даних у стандартизованих форматах, таких як CSV, JSON або Excel. Це дозволяє використовувати результати роботи програмного забезпечення для створення візуалізацій, дашбордів або інтерактивних звітів. Автоматизація експорту забезпечує інтеграцію у безперервний робочий процес, дозволяючи аналітикам отримувати результати аналізу безпосередньо у тих інструментах, які вони використовують для прийняття рішень.

Для інтеграції з платформами машинного навчання, такими як TensorFlow або PyTorch, програмне забезпечення надає можливість імпорту та експорту моделей. Це дає змогу використовувати попередньо навчені моделі для аналізу даних або передавати нові моделі для навчання та оцінки.

Підтримка таких форматів, як SavedModel або ONNX, забезпечує універсальність у роботі з моделями, розробленими на різних платформах. Крім того, передбачено можливість використання бібліотеки scikit-learn для інтеграції із системами, які вимагають швидкого розгортання легких моделей.

Окремої уваги заслуговує можливість інтеграції з месенджерами та іншими системами обміну повідомленнями, такими як Slack, Telegram або Microsoft Teams. Використовуючи API цих платформ, програмне забезпечення може відправляти повідомлення про завершення аналізу, повідомляти про результати кластеризації чи рекомендувати продукти. Це створює можливість інтерактивної роботи з користувачем, інтеграції у командну співпрацю та автоматизації процесів обміну інформацією.

Додатково реалізовано інтеграцію з платформами управління проектами, такими як Jira або Trello. Використання API цих платформ дозволяє автоматично створювати завдання на основі результатів аналізу, наприклад, рекомендувати зміну цінової політики або виявлення нових можливостей на ринку. Це сприяє автоматизації бізнес-процесів і підвищенню ефективності управління проектами.

Таким чином, розроблене програмне забезпечення підтримує широку інтеграцію з іншими інструментами та платформами, що забезпечує його гнучкість, масштабованість і універсальність. Така інтеграція дозволяє використовувати програмне забезпечення у різноманітних контекстах, адаптуючи його до конкретних потреб користувачів або галузей. Вона сприяє підвищенню ефективності роботи, автоматизації процесів та створенню інноваційних рішень на основі аналізу даних.



### 3.6 Аналіз отриманих результатів

На рисунку 3.1 представлено розподіл товарів за категоріями, які були використані для проведення експериментів. На рисунку 3.2 наведено порівняння ціни та рейтингу товарів.

Розподіл цін товарів наведено на рисунку 3.3.

Для визначення впливу різноманітних параметрів товарів було побудовано матрицю кореляції, яку наведено на рисунку 3.4.



Рисунок 3.1 – Розподіл продуктів за категоріями

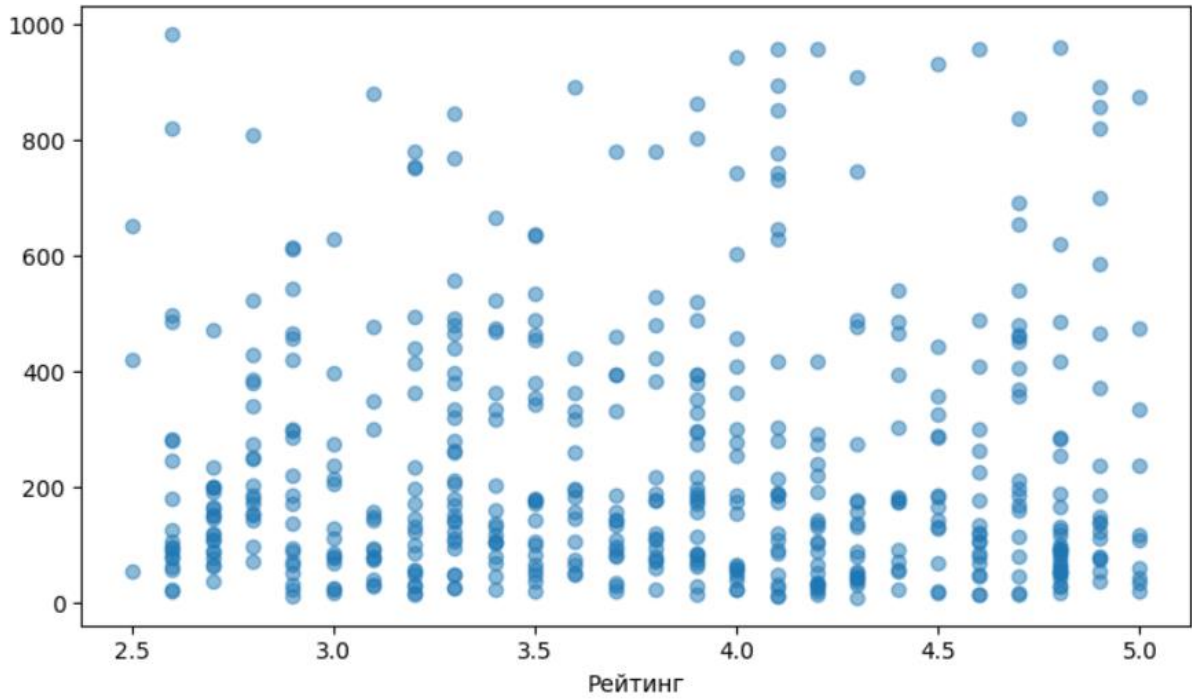


Рисунок 3.2 – Порівняння ціни та рейтингу товарів

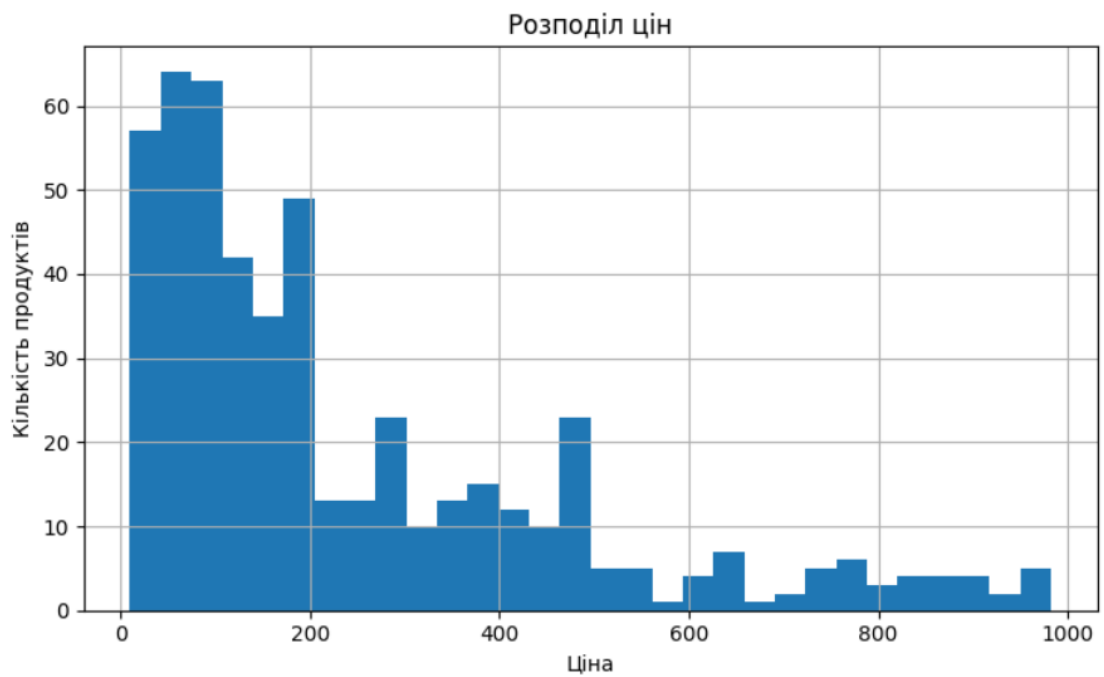


Рисунок 3.3 – Розподіл цін товарів

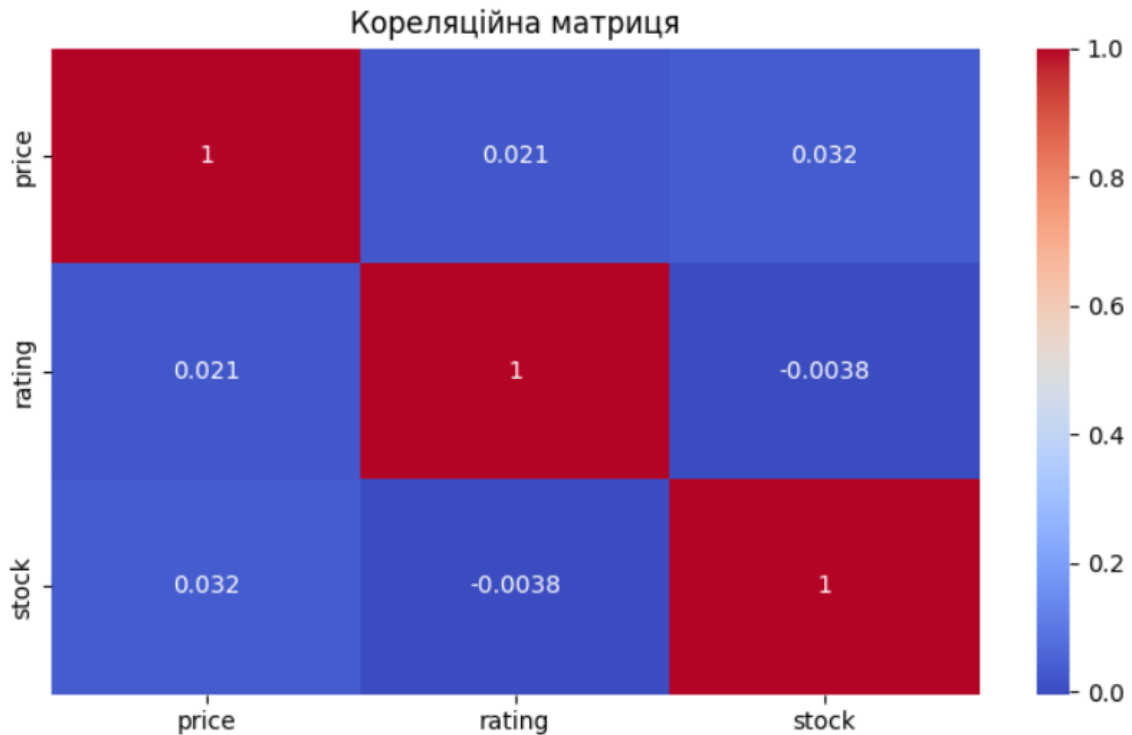


Рисунок 3.4 – Матриця кореляції характеристик товарів

З метою проведення експериментів та створення системи рекомендацій були використані інструменти машинного навчання.

На рисунку 3.5 представлено маркетинговий звіт про E-commerce платформу.

#### МАРКЕТИНГОВИЙ ЗВІТ ПРО E-COMMERCE ПЛАТФОРМУ

##### 1. ЗАГАЛЬНА СТАТИСТИКА:

- Загальна кількість продуктів: 500
- Загальна кількість взаємодій: 31907
- Середня кількість взаємодій на користувача: 15.95

Рисунок 3.5 – Маркетинговий звіт про E-commerce платформу

На рисунку 3.6 наведено розподіл товарів за категоріями. Для визначення цінової статистики були наведені такі характеристики, як: мінімальна ціна, максимальна ціна, середня ціна та медіанна ціна (рис. 3.7).

Перелік найкращих за рейтингом товарів наведено на рисунку 3.8. На рисунку 3.9 представлено сегменти ринку.

#### 2. РОЗПОДІЛ ПРОДУКТІВ ЗА КАТЕГОРІЯМИ:

	Категорія	Кількість
0	Toys	73
1	Sports	72
2	Clothing	68
3	Home & Kitchen	67
4	Books	65
5	Beauty	54
6	Automotive	51
7	Electronics	50

Рисунок 3.6 – Розподіл товарів за категоріями

#### 3. ЦІНОВА СТАТИСТИКА:

- Мінімальна ціна: \$10.14
- Максимальна ціна: \$982.29
- Середня ціна: \$240.11
- Медіанна ціна: \$158.37

Рисунок 3.7 – Цінова статистика

#### 4. ТОП-5 НАЙКРАЩЕ ОЦІНЕНИХ ПРОДУКТІВ:

	Назва	Категорія	Рейтинг
104	Automotive Product 104	Automotive	5.0
116	Sports Product 116	Sports	5.0
137	Books Product 137	Books	5.0
282	Automotive Product 282	Automotive	5.0
315	Clothing Product 315	Clothing	5.0

Рисунок 3.8 – Найпопулярніші товари за рейтингом

## 5. СЕГМЕНТИ РИНКУ:

Segment	Avg Price	Min Price	Max Price	Avg Rating	Count
0	123.62729411764705	11.27	301.56	3.118823529411765	170.0
1	127.39585858585859	10.14	394.41	4.384343434343434	198.0
2	485.24555555555554	316.49	982.29	3.3259259259259255	81.0
3	676.6711764705883	404.62	961.77	4.435294117647058	51.0

Рисунок 3.9 – Сегменти ринку

Для проведення конкретних експериментів було обрано три користувача з різними вподобаннями. Для кожного за допомогою навченої моделі було представлено персоналізовані рекомендації (рис. 3.10-3.12).

## 6. ПЕРСОНАЛІЗОВАНІ РЕКОМЕНДАЦІЇ ДЛЯ КОРИСТУВАЧА:

Рекомендації для користувача 42:

product_id	name	category	price	rating
315	Clothing Product 315	Clothing	117.96	5.0
463	Clothing Product 463	Clothing	38.32	4.9
330	Clothing Product 330	Clothing	137.27	4.9
493	Toys Product 493	Toys	19.44	5.0
304	Toys Product 304	Toys	80.82	4.9
48	Toys Product 48	Toys	148.65	4.9
137	Books Product 137	Books	238.8	5.0
77	Books Product 77	Books	186.7	4.9
366	Books Product 366	Books	65.78	4.8

Рисунок 3.10 – Результати експерименту №1 для користувача під номером 42

Рекомендації для користувача 100:

product_id	name	category	price	rating
493	Toys Product 493	Toys	19.44	5.0
304	Toys Product 304	Toys	80.82	4.9
48	Toys Product 48	Toys	148.65	4.9
315	Clothing Product 315	Clothing	117.96	5.0
463	Clothing Product 463	Clothing	38.32	4.9
330	Clothing Product 330	Clothing	137.27	4.9
464	Beauty Product 464	Beauty	109.97	5.0
40	Beauty Product 40	Beauty	73.99	4.9
107	Beauty Product 107	Beauty	111.13	4.9

Рисунок 3.11 – Результати експерименту №2 для користувача під номером 100

Рекомендації для користувача 250:

product_id	name	category	price	rating
315	Clothing Product 315	Clothing	117.96	5.0
463	Clothing Product 463	Clothing	38.32	4.9
330	Clothing Product 330	Clothing	137.27	4.9
464	Beauty Product 464	Beauty	109.97	5.0
40	Beauty Product 40	Beauty	73.99	4.9
371	Beauty Product 371	Beauty	27.9	4.8
116	Sports Product 116	Sports	60.21	5.0
391	Sports Product 391	Sports	874.83	5.0
223	Sports Product 223	Sports	77.37	4.9

Рисунок 3.12 – Результати експерименту №3 для користувача під номером 250

Використання машинного навчання для створення рекомендаційної системи онлайн-платформи E-commerce продемонструвало високу ефективність в контексті виявлення персоналізованих вподобань користувачів.

### 3.7 Висновки до третього розділу

Розділ програмної реалізації є ключовим, оскільки він охоплює всі аспекти програмної реалізації проекту, від постановки цілей до аналізу отриманих результатів. Підрозділ 3.1 чітко окреслює мету розроблення, що дозволяє зрозуміти основні задачі та очікувані результати проекту. У підрозділі 3.2 детально розглядається архітектура системи, що забезпечує ґрунтовний опис її основних компонентів, їх взаємозв'язків та логіки функціонування. Це створює основу для розуміння того, як система організована на рівні структурної побудови. Опис використаних технологій у підрозділі 3.3 дозволяє зрозуміти, які інструменти та платформи були обрані для реалізації проекту та чому саме вони були обрані. Це допомагає оцінити обґрунтованість технологічних рішень і їх відповідність цілям системи. Підрозділ 3.4 акцентує увагу на процесі навчання та механізмах представлення рекомендацій, що є центральним елементом функціональності розробленого програмного забезпечення. Тут розкривається, як відбувається взаємодія алгоритмів з даними, забезпечуючи якісну обробку та видачу коректних результатів. Інтеграція з іншими інструментами, представлена у підрозділі 3.5, є ключовим аспектом, який демонструє гнучкість і універсальність розробленого рішення. Детальний опис механізмів інтеграції підтверджує можливість взаємодії системи з різноманітними платформами та інструментами, що значно розширює її функціональні можливості та практичну цінність. Нарешті, підрозділ 3.6 завершує розділ аналізом отриманих результатів, який дозволяє оцінити ефективність розробленого програмного забезпечення, виявити сильні сторони реалізації та сформулювати висновки щодо досягнення поставлених цілей.

Поточний розділ формує цілісне уявлення про програмну реалізацію проекту, охоплюючи всі ключові аспекти, необхідні для оцінки його якості, ефективності та потенціалу для подальшого використання та розвитку.

## ВИСНОВКИ

Кваліфікаційна робота магістра представляє собою ґрунтовне дослідження, присвячене розробці рекомендаційної системи для онлайн-платформ електронної комерції, що є одним із найбільш перспективних і актуальних напрямів у сучасній цифровій економіці. З огляду на швидкий розвиток інформаційних технологій та їхній вплив на бізнес-процеси, дана робота вносить важливий вклад у розуміння механізмів персоналізації, автоматизації та покращення користувацького досвіду у сфері електронної комерції.

Робота висвітлює широкий спектр питань, що охоплюють як теоретичні, так і практичні аспекти створення рекомендаційних систем. У першу чергу проведено всебічний аналіз електронної комерції, її ключових особливостей, типології платформ і їхньої ролі у сучасній економіці. Підкреслено, що електронна комерція виступає не лише інструментом здійснення торгових операцій, а й рушійною силою змін у глобальній економіці, забезпечуючи нові можливості для підприємств будь-якого масштабу.

У дослідженні представлено глибокий аналіз методів розробки рекомендаційних систем, включаючи контентну, колаборативну фільтрацію та гібридні підходи. Висвітлено переваги та обмеження кожного з методів, особливості їхнього використання залежно від цілей платформи та поведінкових моделей користувачів. Розкрито роль штучного інтелекту, великих даних і машинного навчання у забезпеченні точності, масштабованості та швидкодії сучасних систем. Особлива увага приділена розв'язанню ключових проблем, таких як «холодний старт», розрідженість даних та забезпечення балансу між точністю і різноманітністю рекомендацій.

Практична частина роботи присвячена створенню та тестуванню прототипу рекомендаційної системи. Запропоновано архітектуру системи,



яка забезпечує високу адаптивність і інтеграцію з іншими бізнес-інструментами. Детально описано використані технології, алгоритми машинного навчання та процеси навчання моделі. Результати тестування підтвердили ефективність розробленого рішення, демонструючи значне зростання рівня задоволеності користувачів, скорочення часу на вибір товарів і підвищення конверсії продажів.

Робота акцентує увагу на перспективних напрямках розвитку рекомендаційних систем, таких як впровадження глибоких нейронних мереж, мультимодальної обробки даних, блокчейн-технологій, а також інтеграція рішень з доповненою та віртуальною реальністю. Окремо розглядаються аспекти забезпечення конфіденційності та захисту даних користувачів, що є критично важливими для сучасних технологічних систем.

Отримані результати мають високу практичну значущість, оскільки створена рекомендаційна система може бути інтегрована в реальні комерційні проекти. Вона сприяє підвищенню ефективності онлайн-продажів, оптимізації роботи з клієнтами, покращенню лояльності аудиторії та створенню конкурентних переваг для бізнесу.

Значущість роботи виходить за межі окремого прикладного проекту, оскільки її результати є основою для подальших досліджень у сфері персоналізованих рекомендацій. Запропоновані рішення і рекомендації сприяють формуванню нових стандартів роботи онлайн-платформ, адаптації бізнес-моделей до умов цифрової трансформації та побудові інноваційних алгоритмів для взаємодії з клієнтами.

У підсумку, магістерська робота відображає сучасний стан розвитку технологій рекомендаційних систем, їхній вплив на трансформацію електронної комерції та перспективи вдосконалення. Це дослідження є важливим етапом у формуванні нових стандартів роботи онлайн-платформ, спрямованих на створення інтуїтивно зрозумілих і максимально ефективних рішень для бізнесу та користувачів у сучасному цифровому світі.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. J. Yu et al., “Collaborative Filtering Recommendation with Fluctuations of User’ Preference”, in 2021 IEEE Int. Conf. Inf. Communication Softw. Eng. (ICICSE), Chengdu, China, Mar. 19–21, 2021. IEEE, 2021. <https://doi.org/10.1109/icicse52190.2021.9404120>
2. Q. Pu and B. Hu, “Intelligent Movie Recommendation System Based on Hybrid Recommendation Algorithms”, in 2023 Int. Conf. Ambient Intell., Knowl. Inform. Ind. Electron. (AIKIE), Ballari, India, Nov. 2–3, 2023. IEEE, 2023. <https://doi.org/10.1109/aikiie60097.2023.10389982>
3. H. Yang, H. Zhou, and Y. Li, “A Review of Academic Recommendation Systems Based on Intelligent Recommendation Algorithms”, in 2022 7th Int. Conf. Image, Vis. Comput. (ICIVC), Xi’an, China, Jul. 26–28, 2022. IEEE, 2022. <https://doi.org/10.1109/icivc55077.2022.9886104>
4. Z. Wang, A. Maalla, and M. Liang, “Research on E-Commerce Personalized Recommendation System based on Big Data Technology”, in 2021 IEEE 2nd Int. Conf. Inf. Technol., Big Data Artif. Intell. (ICIBA), Chongqing, China, Dec. 17–19, 2021. IEEE, 2021. <https://doi.org/10.1109/iciba52610.2021.9687955>
5. X. Jia and F. Liu, “Research on intelligent recommendation system model supported by data mining and algorithm optimization”, in 2021 IEEE Int. Conf. Emergency Sci. Inf. Technol. (ICESIT), Chongqing, China, Nov. 22–24, 2021. IEEE, 2021. <https://doi.org/10.1109/icesit53460.2021.9696972>
6. T. Pan, “Personalized Recommendation Service in University Libraries using Hybrid Collaborative Filtering Recommendation System”, in 2024 Int. Conf. Intell. Algorithms Comput. Intell. Syst. (IACIS), Hassan, India, Aug. 23–24, 2024. IEEE, 2024, pp. 1–5. <https://doi.org/10.1109/iacis61494.2024.10721676>

7. S. Khanduri and S. Prabakeran, “Hybrid Recommendation System with Graph based and Collaborative Filtering Recommendation Systems”, in 2022 IEEE 2nd Mysore Sub Sect. Int. Conf. (MysuruCon), Mysuru, India, Oct. 16–17, 2022. IEEE, 2022. <https://doi.org/10.1109/mysurucon55714.2022.9972677>
8. L. Huang, C. Liu, and W. Peng, “Recommendation System of Tourist Attractions Based on Artificial Intelligence Algorithm”, in 2024 IEEE 13th Int. Conf. Communication Syst. Netw. Technol. (CSNT), Jabalpur, India, Apr. 6–7, 2024. IEEE, 2024. <https://doi.org/10.1109/csnt60213.2024.10546210>
9. Y. Luo, “High-Performance E-Commerce Personalized Recommendation System Based on Matrix Factorization and Convolutional Neural Networks”, in 2023 IEEE Int. Conf. Paradigm Shift Inf. Technol. with Innovative Appl. Global Scenario (ICPSITIAGS), Indore, India, Dec. 28–29, 2023. IEEE, 2023. <https://doi.org/10.1109/icpsitiags59213.2023.10527715>
10. B. Sun, H. Tan, D. Yang, Y. Zhang, Y. Wu, and X. Jin, “The System of Personalized Learning Resource Recommendation and Experimental Teaching Based on Collaborative Filtering”, in 2022 IEEE Int. Symp. Circuits Syst. (ISCAS), Austin, TX, USA, May 27–Jun. 1, 2022. IEEE, 2022. <https://doi.org/10.1109/iscas48785.2022.9937726>
11. G. Huang, “E-Commerce Intelligent Recommendation System Based on Deep Learning”, in 2022 IEEE Asia-Pacific Conf. Image Process., Electron. Comput. (IPEC), Dalian, China, Apr. 14–16, 2022. IEEE, 2022. <https://doi.org/10.1109/ipec54454.2022.9777500>
12. H. Li, Y. Chen, D. Zhang, and L. Wu, “An Intelligent Recommendation System for Performance Equipment Operation and Maintenance via Deep Neural Network and Attention Mechanism”, in 2021 IEEE 10th Data Driven Control Learn. Syst. Conf. (DDCLS), Suzhou, China, May 14–16, 2021. IEEE, 2021. <https://doi.org/10.1109/ddcls52934.2021.9455655>
13. Y. Lai, “Research on the Application of Machine Learning Algorithm in Artificial Intelligence Product Recommendation System”, in 2024 Int. Conf.

Telecommun. Power Electron. (TELEPE), Frankfurt, Germany, May 29–31, 2024. IEEE, 2024, pp. 891–896. <https://doi.org/10.1109/telepe64216.2024.00166>

14. W. Jiang, “Media Personalized Recommendation System Based on Network Algorithm”, in 2022 IEEE 6th Adv. Inf. Technol., Electron. Automat. Control Conf. (IAEAC ), Beijing, China, Oct. 3–5, 2022. IEEE, 2022. <https://doi.org/10.1109/iaeac54830.2022.9929936>

15. S. Yin and X. Luo, “A Survey of Learning-Based Methods for Cold-Start, Social Recommendation, and Data Sparsity in E-commerce Recommendation Systems”, in 2021 16th Int. Conf. Intell. Syst. Knowl. Eng. (ISKE), Chengdu, China, Nov. 26–28, 2021. IEEE, 2021. <https://doi.org/10.1109/iske54062.2021.9755389>

16. H. Liu and Z. Wang, “Research and design of green packaging product recommendation system based on multi-algorithm fusion”, in 2022 IEEE 2nd Int. Conf. Data Sci. Comput. Application (ICDSCA), Dalian, China, Oct. 28–30, 2022. IEEE, 2022. <https://doi.org/10.1109/icdsca56264.2022.9988303>

17. Y. Sun and H. Li, “Personalized Recommendation Based on Knowledge Map and Multi Feature Fusion”, in 2022 11th Int. Conf. Commun., Circuits Syst. (ICCCAS), Singapore, Singapore, May 13–15, 2022. IEEE, 2022. <https://doi.org/10.1109/icccas55266.2022.9824592>

18. X. Yuan, “Design of E-Commerce Personalized Recommendation System Based on Multi-Modal Data Fusion”, in 2024 Int. Conf. Telecommun. Power Electron. (TELEPE), Frankfurt, Germany, May 29–31, 2024. IEEE, 2024, pp. 512–516. <https://doi.org/10.1109/telepe64216.2024.00097>

## ДОДАТКИ

## Додаток А

## Лістинг програмного коду

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.metrics.pairwise import cosine_similarity
from sklearn.preprocessing import StandardScaler
from sklearn.cluster import KMeans
from tabulate import tabulate
import random

class AdvancedRecommendationSystem:
    def __init__(self, num_products=500, num_users=2000):
        np.random.seed(42)
        self.products_df = self._generate_rich_products(num_products)
        self.interactions_df =
self._generate_comprehensive_interactions(num_products, num_users)

    def _generate_rich_products(self, num_products):
        """Генерація продуктів з додатковими характеристиками"""
        categories = [
            'Electronics', 'Clothing', 'Books', 'Home & Kitchen',
            'Sports', 'Beauty', 'Toys', 'Automotive'
        ]
```

```

price_ranges = [(50, 200), (100, 500), (10, 100), (200, 1000)]

products = []
for i in range(num_products):
    category = random.choice(categories)
    price_min, price_max = random.choice(price_ranges)

    products.append({
        'product_id': i,
        'name': f'{category} Product {i}',
        'description': f'High-quality {category.lower()} item with advanced
features.',
        'category': category,
        'price': round(random.uniform(price_min, price_max), 2),
        'rating': round(random.uniform(2.5, 5), 1),
        'stock': random.randint(10, 500)
    })

return pd.DataFrame(products)

def _generate_comprehensive_interactions(self, num_products, num_users):
    interactions = []
    for user_id in range(num_users):
        user_budget = random.uniform(100, 2000)
        num_interactions = random.randint(5, 30)
        user_products = random.sample(range(num_products), num_interactions)

        for product_id in user_products:

```

```

        product = self.products_df.loc[self.products_df['product_id'] ==
product_id].iloc[0]

```

```

        if product['price'] <= user_budget:
            interactions.append({
                'user_id': user_id,
                'product_id': product_id,
                'category': product['category'],
                'rating': round(random.uniform(1, 5), 1),
                'purchase_probability': round(random.uniform(0.1, 1), 2),
                'viewed_time': round(random.uniform(10, 300), 2)
            })

```

```

        return pd.DataFrame(interactions)

```

```

def content_based_recommendations(self, product_id, top_n=5):

```

```

    tfidf = TfidfVectorizer(stop_words='english')

```

```

    tfidf_matrix = tfidf.fit_transform(self.products_df['description'])

```

```

    cosine_sim = cosine_similarity(tfidf_matrix)

```

```

    idx = self.products_df[self.products_df['product_id'] == product_id].index[0]

```

```

    sim_scores = list(enumerate(cosine_sim[idx]))

```

```

    sim_scores = sorted(sim_scores, key=lambda x: x[1],

```

```

reverse=True)[1:top_n+1]

```

```

    product_indices = [i[0] for i in sim_scores]

```

```

    return self.products_df.iloc[product_indices]

```

```

def market_segmentation(self, num_clusters=4):

```

```
"""Кластеризація продуктів за ціною та рейтингом"""
X = self.products_df[['price', 'rating']].values
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)

kmeans = KMeans(n_clusters=num_clusters, random_state=42)
self.products_df['market_segment'] = kmeans.fit_predict(X_scaled)

return self.products_df.groupby('market_segment').agg({
    'price': ['mean', 'min', 'max'],
    'rating': ['mean', 'count']
})

def generate_visualizations(self):
    plt.figure(figsize=(15, 10))

    plt.subplot(2, 2, 1)
    self.products_df['category'].value_counts().plot(kind='bar')
    plt.title('Розподіл продуктів за категоріями')
    plt.xticks(rotation=45)

    plt.subplot(2, 2, 2)
    plt.scatter(self.products_df['rating'], self.products_df['price'], alpha=0.5)
    plt.title('Ціна vs Рейтинг')
    plt.xlabel('Рейтинг')
    plt.ylabel('Ціна')

    plt.subplot(2, 2, 3)
    self.products_df['price'].hist(bins=30)
```



```
plt.title('Розподіл цін')
plt.xlabel('Ціна')
plt.ylabel('Кількість продуктів')

# Корреляційна матриця
plt.subplot(2, 2, 4)
corr_matrix = self.products_df[['price', 'rating', 'stock']].corr()
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm')
plt.title('Кореляційна матриця')

plt.tight_layout()
plt.savefig('market_analysis.png')
plt.close()

def generate_market_report(self):
    """Формування розширеного маркетингового звіту"""
    # Аналіз продуктів
    total_products = len(self.products_df)
    categories_stats = self.products_df['category'].value_counts()

    price_stats = {
        'min': self.products_df['price'].min(),
        'max': self.products_df['price'].max(),
        'mean': self.products_df['price'].mean(),
        'median': self.products_df['price'].median()
    }

    rating_stats = {
        'min': self.products_df['rating'].min(),
```

```

    'max': self.products_df['rating'].max(),
    'mean': self.products_df['rating'].mean(),
    'top_rated_products': self.products_df.nlargest(5, 'rating')[['name',
'category', 'rating']]
    }

```

```

# Аналіз взаємодій користувачів
total_interactions = len(self.interactions_df)
avg_interactions_per_user = total_interactions /
len(self.interactions_df['user_id'].unique())

```

```

# Звіт про сегментацію ринку
market_segments = self.market_segmentation()

```

```

return {
    'total_products': total_products,
    'categories_stats': categories_stats,
    'price_stats': price_stats,
    'rating_stats': rating_stats,
    'total_interactions': total_interactions,
    'avg_interactions_per_user': avg_interactions_per_user,
    'market_segments': market_segments
}

```

```

def personalized_user_recommendations(self, user_id, top_n=10):
    """Персоналізовані рекомендації для конкретного користувача"""
    user_interactions = self.interactions_df[
        self.interactions_df['user_id'] == user_id
    ]

```

```
favorite_categories = user_interactions['category'].value_counts().head(3)
```

```
personalized_recs = []
```

```
for category in favorite_categories.index:
```

```
    category_products = self.products_df[  
        (self.products_df['category'] == category) &  
        (~self.products_df['product_id'].isin(user_interactions['product_id']))  
    ]
```

```
    top_category_recs = category_products.sort_values(  
        by=['rating', 'price'],  
        ascending=[False, True]  
    ).head(top_n // len(favorite_categories))
```

```
for _, product in top_category_recs.iterrows():
```

```
    personalized_recs.append({  
        'product_id': product['product_id'],  
        'name': product['name'],  
        'category': product['category'],  
        'price': product['price'],  
        'rating': product['rating']  
    })
```

```
return personalized_recs[:top_n]
```

```
def main():
```

```
    rec_system = AdvancedRecommendationSystem()
```

```
rec_system.generate_visualizations()
    market_report = rec_system.generate_market_report()

print("МАРКЕТИНГОВИЙ ЗВІТ ПРО Е-COMMERCE ПЛАТФОРМУ\n")

print("1. ЗАГАЛЬНА СТАТИСТИКА:")
print(f"• Загальна кількість продуктів: {market_report['total_products']}")
print(f"• Загальна кількість взаємодій: {market_report['total_interactions']}")
print(f"• Середня кількість взаємодій на користувача:
{market_report['avg_interactions_per_user']:.2f}\n")

print("2. РОЗПОДІЛ ПРОДУКТІВ ЗА КАТЕГОРІЯМИ:")
print(tabulate(
    market_report['categories_stats'].reset_index(),
    headers=['Категорія', 'Кількість'],
    tablefmt='pretty'
))

print("\n3. ЦІНОВА СТАТИСТИКА:")
print(f"• Мінімальна ціна: ${market_report['price_stats']['min']:.2f}")
print(f"• Максимальна ціна: ${market_report['price_stats']['max']:.2f}")
print(f"• Середня ціна: ${market_report['price_stats']['mean']:.2f}")
print(f"• Медіанна ціна: ${market_report['price_stats']['median']:.2f}\n")

print("4. ТОП-5 НАЙКРАЩЕ ОЦІНЕНИХ ПРОДУКТІВ:")
print(tabulate(
    market_report['rating_stats']['top_rated_products'],
    headers=['Назва', 'Категорія', 'Рейтинг'],
    tablefmt='pretty'
```

```
))

print("\n5. СЕГМЕНТИ РИНКУ:")
print(tabulate(
    market_report['market_segments'],
    headers=['Segment', 'Avg Price', 'Min Price', 'Max Price', 'Avg Rating',
'Count'],
    tablefmt='pretty'
))

print("\n6. ПЕРСОНАЛІЗОВАНІ РЕКОМЕНДАЦІЇ ДЛЯ
КОРИСТУВАЧА:")
test_user_ids = [42, 100, 250]

for user_id in test_user_ids:
    print(f"\nРекомендації для користувача {user_id}:")
    user_recs = rec_system.personalized_user_recommendations(user_id)

    print(tabulate(
        user_recs,
        headers='keys',
        tablefmt='pretty'
    ))

if __name__ == "__main__":
    main()
```