

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук
(повна назва)

Кафедра Системотехніки
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти другий (магістерський)

ГЮИК.506100.008 ПЗ

Дослідження способів врахування уподобань клієнтів при розробці систем
електронної комерції
(тема)

Виконала: здобувач освіти II курсу, групи ITПм-20-1

Москалюк Ю.О.

(прізвище, ініціали)

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

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

Освітня програма Інформаційні технології
проекткування
(повна назва освітньої програми)

Керівник доцент Решетнік В.М
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри системотехніки

(підпис)

проф. Гребеннік І.В.

(прізвище, ініціали)

2021 р

Харківський національний університет радіоелектроніки

Факультет _____ Комп'ютерних наук
(повна назва)

Кафедра _____ Системотехніки

Рівень вищої освіти _____ другий (магістерський)

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

Тип програми _____ Освітньо-професійна
(код і повна назва)

Освітня програма _____ Інформаційні технології проектування
(повна назва)

ЗАТВЕРДЖУЮ

Зав. кафедри _____ СТ

_____ проф. Гребеннік І.В.

« _____ » _____ 2021 р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові _____ Москалюк Юлії Олександрівні
(прізвище, ім'я, по батькові)

1. Тема роботи *Дослідження способів урахування уподобань клієнтів при розробці систем електронної комерції*

затверджена наказом по університету від "8" листопада _____ 2021р. № 1663СТ

2. Термін подання студентом роботи до екзаменаційної комісії _____ 17 грудня 2021 р.

3. Вихідні дані до роботи (проекту) *Компоненти системи електронної комерції. Модуль алгоритму врахування уподобань клієнтів. Функції системи: формування персональних рекомендацій для користувачів. Програмні засоби: ОС Microsoft Windows v.7 або вище. Веб-сервер Apache v.2.4. Середовище розробки PhpStorm 2020.1. СУБД MySQL v.6.3. Мови програмування JavaScript, PHP, фреймворк Laravel.*

4. Перелік питань, що потрібно опрацювати в роботі **4.1** *Аналіз предметної області. 4.1.1* Аналіз досліджуваної предметної області. *4.1.2* Огляд способів врахування уподобань користувачів в системах електронної комерції. *4.1.3* Огляд існуючих систем е-комерції з рекомендаційними системами. *4.1.4* Огляд джерел інформації. *4.1.5* Постановка задачі до кваліфікаційної роботи. **4.2** *Дослідження застосування рекомендаційних систем в системах електронної комерції. 4.2.1* Застосування рекомендаційних систем в системах електронної комерції. *4.2.2* Класифікація та порівняння алгоритмів рекомендаційних систем. **4.3** *Розробка алгоритму врахування уподобань клієнтів для системи електронної комерції. 4.3.1* Обґрунтування вибору математичної моделі. *4.3.2* Опис алгоритму розв'язання задачі. *4.3.3* Оцінка якості роботи алгоритму *4.3.4* Висновки порівняння метрик MAE та RMSE. **4.4** *Інтеграція алгоритму рекомендаційної системи до ІС. 4.4.1* Формування вимог до програмного засобу. *4.4.2* Розробка діаграм варіантів використання. *4.4.3* Обґрунтування вибору технологій створення системи електронної комерції. *4.4.4* Обґрунтування вибору СУБД. *4.4.5* Огляд розробленої бази даних з використанням СУБД MySQL. *4.4.6* Розробка інтерфейсу клієнтської частини системи електронної комерції. *4.4.7* Тестування

розроблених компонентів системи. 4.4.8 Аналіз ефективності функціонування системи електронної комерції з інтегрованим алгоритмом.

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) 5.1 Аналоги методу (1 арк. формату А4). 5.2 Схема роботи системи е-комерції з РС (1 арк. формату А4). 5.3 Досліджуваний метод фільтрації (1 аркуш формату А4). 5.4 Діаграма варіантів використання (1 арк. формату А4). 5.5 Фізична модель даних (1 аркуш формату А4) 5.6 Клієнтський інтерфейс (перегляд та оцінка товарів в каталозі) (1 аркуш формату А4). 5.7 Клієнтський інтерфейс (набір рекомендованих товарів (1 арк. формату А4). 5.8 Аналіз ефективності функціонування алгоритму (1 арк. формату А4).). 5.9 Доповідь результатів дослідження (1 арк. формату А4).

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

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

КАЛЕНДАРНИЙ ПЛАН

Пор. №	Назва етапів кваліфікаційної роботи	Термін виконання етапів роботи	Примітка
1.	Отримання завдання кваліфікаційної роботи	08.11.21	
2.	Аналіз завдання та аналогів з теми кваліфікаційної роботи	09.11 — 12.11.21	
3.	Аналіз літератури	12.11 — 13.11.21	
4.	Дослідження методів врахування уподобань	14.11 – 17.11.21	
5.	Дослідження алгоритмів фільтрації	17.11 – 19.11 21	
6.	Розробка модулів системи електронної комерції	20.11 — 25.11.21	
7.	Розробка модулю алгоритма фільтрації	26.11 — 30.11.21	
8.	Інтеграція модулю алгоритма фільтрації до системи електронної комерції	1.12 — 03.12.21	
9.	Оформлення пояснювальної записки та програмної документації	04.12 — 11.12.21	
10.	Оформлення графічної частини та презентаційних матеріалів комп'ютерного захисту	12.12.21	
11.	Представлення на рецензування	14.12.21	
	Представлення кваліфікаційної роботи в ДЕК	15.12.21	

Дата видачі завдання 8 листопада 2021 р.

Студент _____ Москалюк Ю.О.
(підпис)

Керівник роботи _____
(підпис)

доцент каф. СТ Решетнік В.М.
(посада, прізвище, ініціали)

РЕФЕРАТ

Записка пояснювальна: 58 стор., 22 рис., 9 табл., 3 додатки, 29 джерел.
Графічний матеріал кваліфікаційної роботи містить 9 плакатів.

ЕЛЕКТРОННА КОМЕРЦІЯ, Е-КОМЕРЦІЯ, ВРАХУВАННЯ УПОДОБАНЬ,
РЕКОМЕНДАЦІЙНІ СИСТЕМИ, КОЛАБОРАТИВНА ФІЛЬТРАЦІЯ, ВЕБ-
СИСТЕМИ, ПРОГНОЗУВАННЯ.

Об'єктом дослідження кваліфікаційної роботи є система електронної комерції.

Предметом дослідження кваліфікаційної роботи є способи врахування уподобань клієнтів.

Метою роботи є дослідження способів врахування уподобань клієнтів в системах електронної комерції, що дозволяють підвищити якість продажів та рівень задоволення клієнтів.

Методами дослідження є аналіз теоретичного матеріалу, технічної літератури, способів врахування уподобань, методів фільтрації, практична реалізація модулів веб-системи електронної комерції та експериментальне дослідження ефективності її роботи.

Результатами виконання кваліфікаційної роботи є досліджений та розроблений алгоритм колаборативної фільтрації, інтеграція алгоритму колаборативної фільтрації до системи електронної комерції та дослідження його ефективності.

Галузь застосування – веб-системи електронної комерції з продажу одягу.

ABSTRACT

Explanatory note: 58 p., 22 pics., 9 tables, 3 appendixes, 29 sources, 4 applications.
Graphic material qualification work contains 9 posters.

E-COMMERCE, CONSIDERATION OF PREFERENCES,
RECOMMENDATION SYSTEMS, COLLABORATIVE FILTRATION, WEB
SYSTEMS, FORECASTING.

The object of investigation of the qualification work is the e-commerce system.

The subject of qualification work investigation are ways of consideration of customers preferences.

The goal of the qualification work is to investigate ways of consideration of customers preferences in e-commerce systems that increase sales and improve customer satisfaction.

Investigation methods include analysis of theoretical material, technical literature, methods of consideration of customers preferences, filtering methods, practical implementation of modules of the e-commerce web system and experimental investigation the effectiveness of its performance.

The results of the qualification work include investigated and developed algorithm of collaborative filtering, integration of the algorithm of collaborative filtering into the e-commerce system modules and research of its efficiency.

Scope e-commerce web systems for clothing sales.

ЗМІСТ

ВСТУП	8
1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ	10
1.1 Аналіз досліджуваної предметної області	10
1.2 Огляд способів урахування уподобань користувачів в системах електронної комерції	17
1.3 Огляд існуючих систем е-комерції з рекомендаційними системами	20
1.4 Огляд джерел інформації.....	23
1.5 Постановка задачі до кваліфікаційної роботи.....	25
2 ДОСЛІДЖЕННЯ ЗАСТОСУВАННЯ РЕКОМЕНДАЦІЙНИХ СИСТЕМ ВСИСТЕМАХ ЕЛЕКТРОННОЇ КОМЕРЦІЇ.....	27
2.1 Застосування рекомендаційних систем в системах електронної комерції.....	27
2.2 Класифікація та порівняння алгоритмів рекомендаційних систем.....	31
3 РОЗРОБКА АЛГОРИТМУ ВРАХУВАННЯ УПОДОБАНЬ КЛІЄНТІВ ДЛЯ СИСТЕМИ ЕЛЕКТРОННОЇ КОМЕРЦІЇ.....	39
3.1 Обґрунтування вибору математичної моделі	39
3.2 Опис алгоритму розв'язання задачі.....	44
3.3 Оцінка якості роботи алгоритму.....	47
3.4 Висновки порівняння метрик MAE та RMSE	48
4 ІНТЕГРАЦІЯ АЛГОРИТМУ РЕКОМЕНДАЦІЙНОЇ СИСТЕМИ ДО ІС.....	50
4.1 Формування вимог до програмного засобу	50
4.2 Розробка діаграм варіантів використання	51
4.3 Обґрунтування вибору технологій створення системи електронної комерції.....	54
4.4 Обґрунтування вибору СУБД	55
4.5 Огляд розробленої бази даних з використанням СУБД MySQL.....	55
4.6 Розробка інтерфейсу клієнтської частини системи електронної комерції.....	56
4.7 Тестування розроблених компонентів системи	58
4.8 Аналіз ефективності функціонування системи електронної комерції з інтегрованим алгоритмом.....	61
ВИСНОВКИ.....	66
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ.....	67
Додаток А Графічні матеріали.....	70
Додаток Б Текст програми	80
Додаток В Відомість дипломної роботи	99

ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАК, ОДИНИЦЬ І ТЕРМІНІВ

БД – база даних;

ІС – інформаційна система;

КФ колаборативна фільтрація;

ПЗ – програмне забезпечення;

РС – рекомендаційна система;

СУБД – система управління базами даних;

B2A – Business-to-Administration;

B2B – Business-to-Business;

B2C – Business-to-Customer;

C2A – Customer-to-Administration;

C2B – Customer-to-Business;

C2C – Customer-to-Customer;

CEO – Chief Executive Officer (головний виконавчий директор);

HTTP – HyperText Transfer Protocol (протокол передачі гіпертексту);

MVC Model-View-Controller;

PHP – Hypertext Preprocessor (гіпертекстовий препроцесор).

ВСТУП

Технології у XXI столітті невпинно продовжують розвиватися. Протягом останніх років сфера торгівлі стрімко переміщується в Інтернет в галузь електронної комерції (е-комерції). Інтернет-магазини охоплюють будь-які напрямки торгівлі та значно спрощують взаємодію між продавцями та клієнтами, дозволяють прискорити вихід певної одиниці бізнесу на ринок, скоротити витрати на рекламу тощо.

Виникають нові проблеми, які слід вирішувати шляхом автоматизації. У нагоді стає машинне навчання, вже створено сотні алгоритмів, бібліотек, фреймворків, що дозволяють застосовувати цю технологію для власних досліджень. А для стрімкого розвитку торговельного бізнесу та підвищення прибутків необхідно вдосконалювати процес торгівлі з урахуванням уподобань клієнтів. Користувачам стає все складніше робити вибір серед нескінченної кількості товарів, здійснювати їхній пошук. Тому з'являється необхідність обирати товари з віртуальним «консультантом», який може запропонувати те, що може потенційно сподобатися користувачу. Основні способи урахування уподобань клієнтів базуються на використанні рекомендаційних систем, адже вони спрямовані на взаємодію з користувачем, якому надається певна кількість об'єктів, серед яких він і має зробити свій вибір.

Рекомендаційні системи – це підмножина систем для фільтрації контенту, які допомагають розв'язати проблеми холодного старту, шахрайства, масштабованості, проблеми користувачів щодо складності пошуку товарів або їхнього вибору. Рекомендації підбираються на основі преференцій та поведінки користувача. Система рекомендацій має спрогнозувати реакцію користувача на той чи інший товар і запропонувати інші товари, які також можуть стати до вподоби. Після чого, веб-система інтернет-магазину має показати користувачу схожі товари.

Отже, практично, рекомендаційна система – це сукупність алгоритмів і програм, яка розв'язує задачу передбачення того, що може зацікавити клієнта. Вона може бути використана в інформаційній системі е-комерції.

Об'єктом дослідження кваліфікаційної роботи є процеси функціонування рекомендаційних систем в системах електронної комерції.

Предметом дослідження кваліфікаційної роботи є методи та алгоритми рекомендаційних систем в системах електронної комерції.

Метою кваліфікаційної роботи є дослідження способів урахування вподобань клієнтів при розробці систем електронної комерції. Для досягнення поставленої мети необхідно дослідити та проаналізувати способи урахування уподобань користувачів в системах електронної комерції, інтегрувати модуль рекомендаційної системи на основі обраного алгоритму до розроблених модулів системи електронної комерції, провести аналіз отриманих результатів та оцінити їх.

Актуальність теми обумовлена тим, що завжди є необхідність розширяти та покращувати бізнес за рахунок залучення більшої кількості клієнтів. А це, в свою чергу, потребує покращення досвіду клієнтів у користуванні системами електронної комерції, якого можна досягти шляхом запровадження рекомендаційних систем. Розгортання алгоритмів, які використовуються в рекомендаційних системах, в системах електронної комерції допоможе прискорити вихід певної одиниці бізнесу на ринок, залучити велику кількість задоволених клієнтів, популяризувати товари чи послуги серед цільової аудиторії, підняти продажі та виділитися серед конкурентів, у разі їхньої наявності, а такі кроки стануть на пригоді будь-якому бізнесу у сфері торгівлі.

Публікації: результати проведених досліджень пройшли апробацію на I Міжнародній студентській науковій конференції «Формування сучасної науки: методика та практика» 29 жовтня 2021 року в Кам'янці-Подільському та опубліковані в збірнику [1].

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Аналіз досліджуваної предметної області

Важливою тенденцією сучасного світу є значне зростання кількості інституцій, що займаються накопиченням та обробкою великих обсягів даних. Дані є основними цінностями або фактами, які організовані в єдину базу даних. Багато людей вважають дані синонімом інформації; однак інформація насправді складається з даних, які були організовані, щоб допомогти відповісти на запитання та вирішити проблеми. Інформаційна система визначається як програмне забезпечення, яке допомагає організувати та аналізувати дані. Отже, мета інформаційної системи — перетворити вихідні дані в корисну інформацію, яку можна використовувати для прийняття рішень в організації.

Розглянемо типи інформаційних систем. Системи обробки транзакцій TPS (Transaction Processing System) обробляють всі дані про транзакції клієнтів і співробітників, щоб організація могла оптимізувати робочі процеси та легко отримувати необхідну інформацію [2].

Системи автоматизації офісу керують всіма повсякденними службовими та управлінськими завданнями в бізнесі, щоб допомогти оптимізувати взаємодію та покращити співпрацю.

Системи управління знаннями обробляють досвід з різних тем і допомагають в обміні знаннями, щоб користувачі могли стати більш поінформованими та покращити свою роботу.

Інформаційні системи MIS (Management Information Systems) використовують дані системи обробки TPS (Transaction Processing System), щоб допомогти менеджерам середньої ланки оптимізувати прийняття рішень і контролювати ефективність.

Системи підтримки прийняття рішень DSS (Decision Support System) обробляють інформацію, щоб допомогти менеджерам зробити правильний вибір у потрібний момент.

Інформаційні системи виконавчих систем керують всією необхідною інформацією, необхідною для керівників підприємств для моніторингу конкуренції, відстеження внутрішньої ефективності та визначення можливостей зростання.

Взагалі, на меті всіх бізнес-інформаційних систем стоїть необхідність дозволити організаціям ефективно керувати бізнесом, як правило, за допомогою комп'ютерів. Виокремлюють чотири основних компоненти бізнес-інформаційних систем: апаратне забезпечення, програмне забезпечення, дані та процедури (проектування, розробка та документація).

Звертаючи увагу на те, що торгівля безпосередньо є підмножиною множини бізнесу, пропонується розглянути електронну торгівлю, що потребує впровадження інформаційних систем для успішного її запуску.

Платформи для електронної торгівлі, яка ще має назву «електронна комерція», були б неможливі без інформаційних технологій. Крім того, це дає можливість продавцям і клієнтам легко підключатися через мобільні та веб-платформи, а також відстежувати доставку, повернення та оплату придбаних товарів і керувати ними.

Світовий ринок електронної торгівлі є одним з найбільш динамічних і стабільно зростаючих. Завдяки прогнозам та спостереженням спеціалістів відомо, що темпи зростання обсягів електронної торгівлі випереджатимуть темпи зростання роздрібною торгівлі загалом і надалі. А також відомо, що динаміка показників електронної торгівлі України, як приклад, є набагато кращою за динаміку показників роздрібною торгівлі.

Зростання обороту електронної торгівлі у світі пов'язане з тим, що для все більшої кількості людей Інтернет перетворюється на природне середовище існування, в межах якого задовольняється все більше потреб, у тому числі й покупок та продажу.

Для того, щоб зробити компанії більш успішними та популярними на ринку, необхідно вміти доносити інформацію про свої товари або послуги до клієнта.

Також, невід'ємною складовою є ідеально вибудований процес продажів не тільки той, що реалізується завдяки купівлі приміщень та продажу товарів чи послуг фізично, а й завдяки використанню мережі Інтернет – в електронних магазинах.

Сьогодні велика кількість компаній, що продають певні товари або послуги, стикаються з проблемами збуту. Для того, щоб вирішити подібну проблему необхідно застосовувати більш широкий концептуальний підхід, який допоможе вибудувати електронний бізнес як спосіб ведення підприємницької діяльності, який, в свою чергу, допоможе сягнути стратегічного успіху. Простіше кажучи, такий вид електронного бізнесу і називають електронною комерцією.

Сфера продажів існує вже дуже великий проміжок часу. А розширюватися вона почала ще до появи Інтернету та сучасної форми електронної комерції. Що стосується перших та найпростіших систем комерції, то вони беруть свій початок у 1960-х роках в США, коли вони використовувалися вперше в транспортних компаніях для замовлення різних білетів. Відомо, що довгий час електронна комерція існувала без застосування інтернету та не мала нічого спільного з подібними технологіями та системами. Вона лише характеризувалася використанням протоколів та стандартів, котрі містили правила оформлення електронних ділових документів.

Після того, як в 1980-х роках число користувачів, які обмінювалися між собою електронними діловими документами зростає, то в 1990-х роках сталася поява eBay та Amazon, яка була революційною в індустрії електронної комерції. Відтоді споживачі можуть купувати нескінченну кількість товарів та послуг в Інтернеті, розраховуючись банківською картою, у електронних роздрібних торговців - типових магазинів та вендорів з можливостями електронної комерції. Зараз майже усі підприємства роздрібної торгівлі інтегрують успішну практику онлайн-бізнесу в свої бізнес-моделі.

Людство здебільшого перейшло в епоху електронного обміну даними та залежності від Інтернету. Є декілька загальновідомих причин, що сприяють такому стрімкому розвитку:

- зміщення уваги бізнесу з магазинів до електронної комерції в Інтернеті;
- залучення до Інтернету все більшої кількості користувачів;
- зниження витрат на використання електронної комерції завдяки низькій вартості обміну інформацією.

Електронна комерція (від англ. e-Commerce), далі «е-комерція» – це купівля та продаж товарів і послуг або передача коштів чи даних через електронну мережу, насамперед через Інтернет. Терміни електронна комерція та електронний бізнес часто використовуються як взаємозамінні [3].

За об'єктами і суб'єктами системи електронної комерції поділяються на наступні типи:

- Бізнес-бізнес (B2B – Business-to-Business) – цей термін застосовують у маркетингу, означає товарообмін, обмін інформацією чи послугами (або їхній продаж) між будь-якими компаніями. Такий процес не включає кінцевого фізичного споживача;

- Бізнес-споживач (B2C – Business-to-Customer) – вид е-комерції, який застосовують для позначення процесу взаємодії компанії, яка є юридичною особою (Business), та кінцевого споживача, який є фізичною особою (Customer) в свою чергу. На меті такого виду комерції є продаж товарів, інформації, послуг;

- Споживач-споживач (C2C – Customer-to-Customer) – це вид електронної комерції, де споживачі торгують товарами, послугами та інформацією один з одним в Інтернеті. Зазвичай такі транзакції здійснюються завдяки третій стороні, що надає онлайн -платформу, на якій і здійснюються транзакції.

Споживач-бізнес (C2B – Customer-to-Business) – це тип електронної комерції, в якій споживачі роблять свої продукти та послуги доступними в Інтернеті, щоб компанії могли робити ставки та купувати їх. Це протилежність традиційній комерційній моделі B2C. Популярним прикладом платформи C2B є ринок, що продає фотографії, зображення, засоби масової інформації та елементи дизайну. Іншим прикладом може бути дошка вакансій.

- Бізнес-адміністрація (B2A – Business-to-Administration) – це транзакції, що здійснюються в Інтернеті між компаніями та державною адміністрацією або

державними органами. Багато гілок влади так чи інакше залежать від електронних послуг чи продуктів, особливо це стосується юридичних документів, реєстрів, соціального забезпечення, фіскалів та зайнятості. Підприємства можуть поставляти їх в електронному вигляді. Послуги В2А значно зросли за останні роки, оскільки були зроблені інвестиції у можливості електронного уряду.

– Споживач-адміністрація (С2А – Customer-to-Administration) відноситься до транзакцій, що проводяться в Інтернеті між окремими споживачами та державною адміністрацією чи державними органами.

Пошук товарів зазвичай починається з мережі Інтернет та пошукового рядка браузера. Саме тому, інтернет-магазини – це діюча можливість побудови бізнесу з нуля або розгортання існуючого в незвичайному новому форматі.

Веб-система як складова інтернет-магазину взагалі є прикладом В2С-механізму, на меті якої стоїть спрощення роботи власників бізнесу-продавців та покупців. Зазвичай термін позначає процес продажу продуктів безпосередньо споживачам, включаючи покупки в магазині або їжу в ресторані. Сьогодні він описує транзакції між інтернет-роздрібними торговцями та їх клієнтами.

Слід зазначити переваги моделі В2С: великий та різноманітний ринок, простіше розширення, простіший маркетинг, скарбниця даних, якими досить легко керувати. Стосовно основних недоліків, які слід вказати – це високий рівень конкуренції, наявність посередників, сегментований ринок.

Е-комерція В2С-типу стає більш популярною останнім часом. Вона спрощує покупку як для споживачів, так і для самих продавців. Запуск інтернет-магазинів не потребує багато часу, а покупців можна залучити за рахунок простоти використання та швидкості обробки замовлень, яка дозволяє купувати товари з дому або робити це з будь-якої частини світу за наявності Інтернету. Всі процеси, що відбуваються в рамках електронної торгівлі, здійснюються відповідно до Закону України №675 Про електронну комерцію. «Закон визначає організаційно-правові засади діяльності у сфері електронної комерції в Україні, встановлює порядок вчинення електронних правочинів із застосуванням інформаційно-

телекомунікаційних систем та визначає права і обов'язки учасників відносин у сфері електронної комерції» [4].

Проте, на законодавчому рівні торгівля в Інтернеті в межах України ще й досі не набула відповідних чинних законів та мало регулюється взагалі. Тому таке питання й досі залишається відкритим та просувається власниками бізнесів, що розгортаються на просторі Інтернету.

До основних переваг систем електронної комерції можна віднести:

- клієнти можуть зберегти зусилля та час;
- широкий асортимент товарів доступний в інтернеті, тому порівняння можна легко зробити перед покупкою;
- клієнт може робити покупки в будь-який час і з будь-якого місця, об'єкт доступний у будь-який час тижня;
- детальна інформація про продукт доступна в інтернеті, що допомагає клієнту прийняти рішення про покупку;
- наявна можливість переглянути схожі товари в тому самому магазині або товари, що є хітами продажів за певний час, а це, в свою чергу, допомагає клієнтам ознайомлюватися з більш широким асортиментом товарів, не роблячи додаткових кроків самостійно;
- електронна торгівля пропонує прості умови оплати, такі як оплата при доставці, які спонукають клієнта робити покупки в інтернеті.

Але існують і певні недоліки електронної комерції:

- клієнти можуть не бути впевнені в якості продуктів, пропонованих в інтернеті;
- відсутність в електронній роздрібній торгівлі тенденції кожної людини торгуватися перед вчиненням остаточної покупки;
- клієнти можуть не довіряти платіжним шлюзам і побоюватися неправильного використання кредитних карт або будь-якого іншого способу оплати;

– кожен клієнт хоче бачити і відчувати продукт, який він купує, але це неможливо в разі електронної роздрібної торгівлі, де клієнт приймає рішення, просто дивлячись на зображення;

– клієнту доводиться чекати деякий час, щоб отримати продукт в свої руки;

– клієнт втрачає емоційну прихильність до продавця, що призводить до меншої віри в пропозиції.

Базуючись на цьому, значна кількість компаній, які є одиницями середнього та малого бізнесу, залучаються до електронної комерції, адже вона відкриває новий простір для діяльності, яку неможливо впровадити в фізично існуючих магазинах. Практична цінність систем електронної комерції полягає у високій чутливості до динаміки змін фінансово-економічних показників підприємств та зовнішнього ринку. Це дозволяє ухвалювати ефективні управлінські рішення та прогнозувати критичні виробничі ситуації.

Не всі існуючі системи електронної комерції виявляються адаптивними та зручними для використання. Щоб успішно запровадити модель електронної комерції B2C, підприємства повинні покладатися на наявність платформи, яку можна швидко налаштувати та адаптувати до нових потреб клієнтів, не викликаючи затримок у роботі.

В офлайн-магазинах можна зустріти продавця, який відповідає за задоволення покупців та за збільшення продажів компанії. Компанії електронної комерції не мають такої переваги у вигляді доброзичливого менеджера з продажів, який допоможе клієнту на кожному кроці його покупки. В інтернет-магазинах цю роль виконують алгоритми штучного інтелекту, які створюють системи рекомендацій для кожного клієнта, тобто враховують уподобання. Тому, для задоволення клієнтів дуже важливо вміти орієнтуватися в продуктах та інформації, які їм потрібні. А для таких цілей бажано застосовувати рекомендаційні системи.

1.2 Огляд способів урахування уподобань користувачів в системах електронної комерції

Головним завданням будь-якої рекомендаційної системи є передбачення відношення користувача до якогось нового для нього продукту.

В загальному випадку виділяють три базові методи, як способи урахування уподобань, які використовуються для прийняття рішень у рекомендаційних системах (рис. 1.1):

- метод фільтрації контенту;
- метод колаборативної фільтрації;
- гібридні методи, які комбінують обидва названі вище підходи і, можливо, використовують ще якісь нововведення.

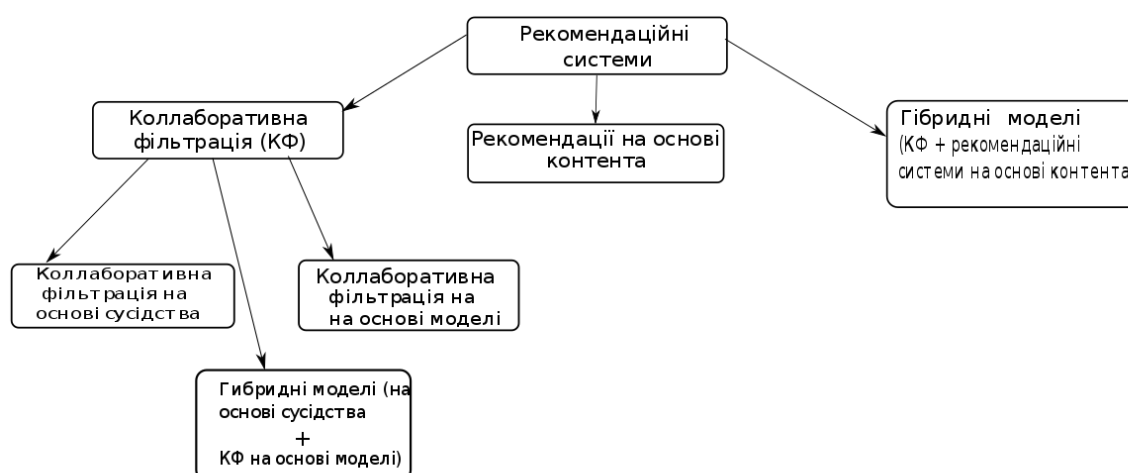


Рисунок 1.1 – Класифікація способів урахування уподобань користувачів

Метод фільтрації контенту враховує переваги лише конкретного користувача та надає рекомендації лише йому. Рекомендовані об'єкти мають описові характеристики ознаки. Ознаки визначають переваги користувача та записані в профілі індивідуальних уподобань. Об'єкти, чії описи мають високий рівень подібності з профілем індивідуальних переваг, рекомендуються користувачеві. Головна перевага: персоналізація рекомендацій під конкретного користувача

системи та очевидність принципів роботи. До недоліків відноситься проблема надання рекомендацій новому користувачеві без історії запитів та взаємодії із системою, обмеженість встановленим числом параметрів об'єкта.

Метод колаборативної фільтрації (КФ), який використовує система, надає рекомендації користувачеві на основі переваг груп користувачів. Коли користувач оцінює об'єкт, система співвідносить користувача до групи, яка поставила обраному об'єкту таку саму оцінку. Надалі користувачеві будуть надаватися об'єкти, високо оцінені користувачами з присвоєної групи. Ставлення користувачів до об'єктів представляється у вигляді матриці.

Метод колаборативної фільтрації, в свою чергу, поділяється на фільтрації за типами:

- колаборативна фільтрація на основі сусідства;
- колаборативна фільтрація на основі моделі;
- гібридні моделі (на основі сусідства та на основі моделі).

КФ на основі сусідства припускає, що нового користувача системи (user-based) зараховують до найближчої за перевагами групи користувачів, або новий об'єкт (item-based) у рекомендаційній системі зараховують до групи об'єктів.

Перевагою підходу, заснованого на пам'яті, стає його простота в розумінні та реалізації, що є достатнім для впровадженні рекомендаційної системи до новоутвореної системи електронної комерції, проте зі збільшенням кількості користувачів та об'єктів знижується продуктивність ІС зі збільшенням потреб у пам'яті.

Підхід, заснований на моделі, створений для зниження обчислювальної складності, що швидко зростає зі збільшенням кількості користувачів та об'єктів у системі. Для груп користувачів будуються статистичні моделі за допомогою методів машинного навчання [5].

Гібридний метод дозволяє позбавити систему більшості недоліків, проте вартість і час створення подібної системи зростає.

Якщо брати до уваги способи, що не базуються на машинному навчанні, то існують способи, які дозволяють обчислити середні значення, моду та ін. для

визначення популярності певних товарів в електронних магазинах. Проте такі способи досить статичні та неточні, адже вони ніяк не зможуть дослідити поведінку користувача та адаптуватися до змін, які можуть відбуватися щохвилини. Зазвичай, це аналіз переглянутих товарів авторизованими користувачами, інформація про які зберігається в базах даних. Або додавання певних маркерів-тегів до кожного товару, які теж мають зберігатися в базах даних та згодом використовуватимуться для вибору товарів користувачами.

Деякі власники бізнесу, які не мають можливості запровадити розроблені алгоритми командою розробників, через різні причини (наприклад, через недостатню кількість грошей), застосовують онлайн-інструменти для відстеження дій користувачів. Потім, на основі зібраних даних, вони проводять аналіз, іноді навіть вручну, та, базуючись на результатах, змінюють вектор руху свого бізнесу. Наприклад, можна знайти найпопулярніші товари, наступним шляхом: визначити кількість замовлень, яка була розміщена за певний період. Це може бути щотижневий, щомісячний, кварталний або річний період. Потім необхідно порівняти аналогічний проміжок часу, щоб побачити, чи збільшився чи зменшився обсяг продажів у відсотках за певними товарами. Це непопулярний та трудомісткий спосіб.

Підсумовуючи: найкраще за все із задачами врахування уподобань можуть впоратися алгоритми, що застосовуються в машинному навчанні. Вони здатні охоплювати велику кількість даних, аналізувати їх та робити прогнози або видавати певні результати. Порівнявши існуючі способи врахування уподобань користувачів, обраний метод колаборативної фільтрації на основі сусідства, адже він є найшвидкішим та найдешевшим для реалізації, а також добре оброблює невеликі об'єми даних, чого достатньо в перший час впровадження системи електронної комерції.

1.3 Огляд існуючих систем е-комерції з рекомендаційними системами

Досліджені декілька найкрупніших світових електронних магазинів, які використовують один або кілька варіантів технології рекомендаційної системи на своїх веб-сайтах.

Першою в списку є веб-платформа продажу товарів Amazon. Amazon, є багатомільярдним гігантом з продажів, який надає кожному клієнту індивідуальний досвід покупок. Amazon беруть за приклад багато інтернет-магазинів, а це робить його ідеальною компанією, дивлячись на яку можна навчатись. Amazon приділяє велику увагу маркетингу на основі даних.

Не завжди легко змусити користувачів оцінювати товари. Ось чому фільтрація предметів може бути такою ж простою, як натиснути на сукню та побачити більше суконь. Amazon винайшов фільтрацію предметів для своєї системи рекомендацій. Фільтрування елементів найкраще працює, коли на вашій платформі більше користувачів, ніж елементів. Натиснувши посилання «Ваші рекомендації» на Amazon.com, користувачі перейдуть на сторінку з продуктами, рекомендованими саме для вас.

Amazon рекомендує цілий ряд продуктів із різних категорій, які ви переглядали, з метою показати вам продукти, які ви, ймовірно, натиснете, дізнаєтеся більше та купите (рис. 2.2).

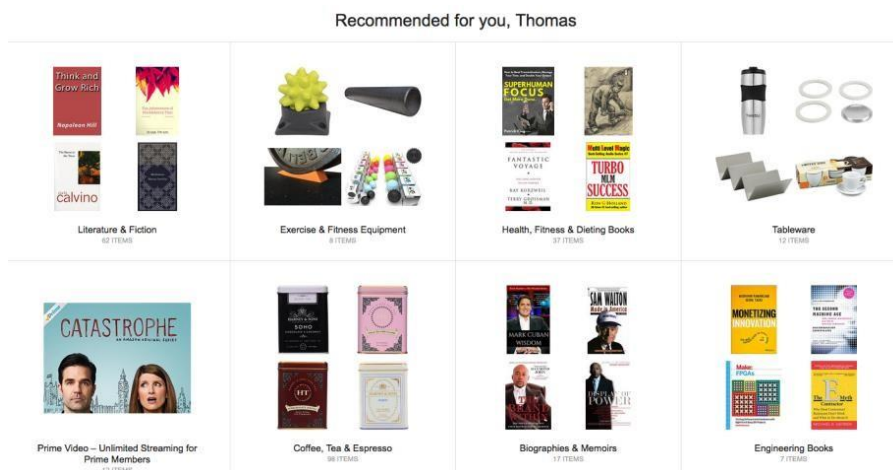


Рисунок 2.2 – Рекомендаційна видача на веб-платформі Amazon

Також, Amazon збирає дані про продукти, які користувачі переглядали, і рекомендує дуже схожі продукти різних форм, розмірів і брендів, щоб допомогти знайти щось дуже схоже на продукт, до якого користувачі вже проявили зацікавленість (рис. 2.3).

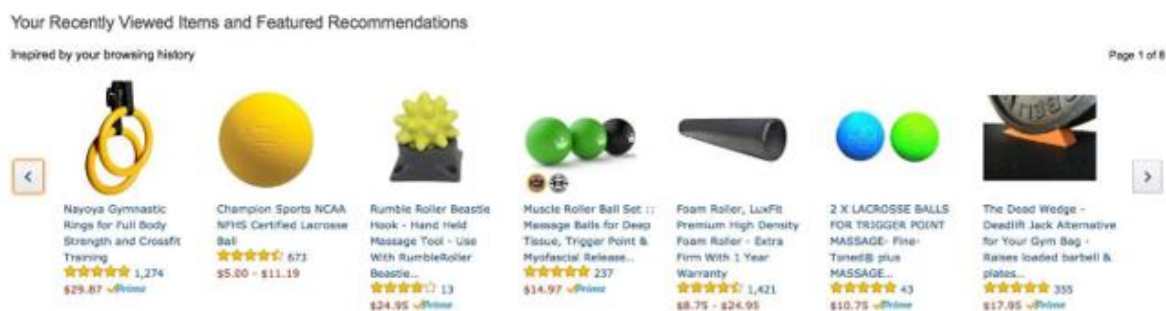


Рисунок 2.3 – Перелік нещодавно переглянутих та рекомендованих товарів на веб- платформі Amazon

В основному, Amazon використовує наступні технології для формування рекомендацій: співвідношення між елементами (Item to Item correlation) та засновані на елементах (Attribute Based), а також колаборативну фільтрацію. Тобто, Amazon застосовує спільну фільтрацію на основі елементів, а також досвід, отриманий від інших користувачів. Він працює в режимі реального часу і може обробляти великі колекції даних і давати рекомендації з чудовим співвідношенням імовірності. Порівнюється вибраний користувачем товар, а також попередня історія придбаних товарів. Потім порівняння оцінює елементи до подібних. Останнім кроком є створення списку рекомендацій на основі рейтингу, наданого користувачеві.

За останні роки Amazon випустила програмне забезпечення глибокого навчання DSSTNE з відкритим кодом – програмне забезпечення є основою для рекомендацій на Amazon [6]. По суті, існує тенденція до більш глибоких систем рекомендацій в електронній комерції. На додаток до можливості відображення

«Популярних речей», все більше компаній покладаються на дуже персоналізовані рекомендації. Як правило, об'єднуються кілька рекомендаційних

стратегій: автоматично включаються інтереси до покупки, популярні товари та інші фактори, такі як доступність товару та зміна ціни.

Другою розглянемо корпорацію електронної комерції eBay, яка сприяє продажам від споживачів до споживачів (C2C) та від бізнесу до споживачів (B2C) через свій веб-сайт. eBay використовує систему, яка дозволяє різним відділам компанії перевіряти дані зі свого набору даних у пісочниці для аналізу. eBay вже досяг значних успіхів у бізнесі завдяки аналітиці даних. Приклад видачі рекомендованих товарів на eBay наведено на рис. 2.4.

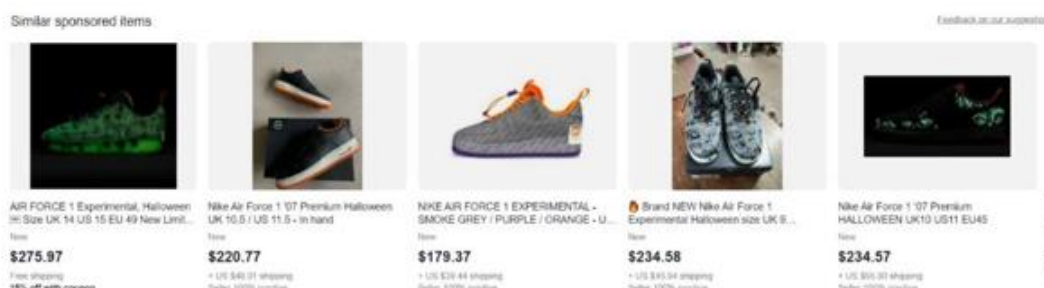


Рисунок 2.4 – Приклад видачі рекомендованих товарів на eBay

eBay використовує профіль зворотного зв'язку. Функція профілю зворотного зв'язку на eBay дозволяє як покупцям, так і продавцям брати участь у профілях зворотного зв'язку інших клієнтів, з якими вони мали справу.

Відгук складається з оцінки задоволеності користувачів (задоволений/нейтральний/незадоволений), а також конкретного коментаря про іншого клієнта. Зворотній зв'язок використовується для забезпечення системи коментарів для покупців, які можуть переглядати профіль продавців. Цей профіль складається з таблиці кількості кожного рейтингу за останні 7 днів, минулого місяця та останніх 6 місяців, а також загального підсумку (наприклад, 867 позитивних відгуків від 776 унікальних клієнтів). За подальшим запитом клієнти можуть переглянути індивідуальні рейтинги та коментарі для продавців

Системи рекомендацій, засновані на відгуках користувачів (які і використовує eBay), оцінюють елементи шляхом узагальнення оцінок

користувачів, щоб вибрати ті, які мають найвищий рейтинг. Рейтинги зазвичай агрегуються за допомогою середньозваженого арифметичного. Однак середнє є досить чутливим до викидів, і, таким чином, може бути не найбільш інформативним агрегатом. Ми порівнюємо точність і надійність трьох різних агрегаторів: середнього, медіани і моди. Результати показують, що медіана часто може бути кращим вибором, ніж середня, і може значно покращити точність і надійність рекомендацій у системах спільної фільтрації [7].

1.4 Огляд джерел інформації

Для того, щоб зрозуміти на скільки важливою є проблема відсутності рекомендаційних систем в системах електронної комерції необхідно звернути увагу на питання, які піднімаються в рамках наукових статей та інших ресурсів.

Перш за все слід зауважити, що проблема урахування вподобань користувачів є досить розповсюдженою. Її висвітлюють у великій кількості робіт, частину з яких перелічено нижче.

Кім Фальк, як спеціаліст з аналізу та обробки даних, приділив значну увагу цій проблемі та присвятив їй роботу «Practical Recommender Systems» [8]. В роботі, спираючись на значну кількість практичних прикладів, автор розкриває сутність рекомендаційних систем та наводить оцінки їх важливості.

К.Фальк вводить основний понятійний апарат зі сфери рекомендаційних систем для його подальшого використання при їх розробці: спеціалізація, контекст тощо.

Спеціалізація – це тип контенту, який рекомендується. Спеціалізація відіграє важливу роль, оскільки служить орієнтиром при організації роботи з рекомендаціями. Спеціалізація дозволяє оцінити, наскільки негативними будуть наслідки помилок. Крім того, від спеціалізації залежить, чи можна рекомендувати те саме по кілька разів.

Задача, частіше за все це – зробити якнайбільше продажів. На яких користувачів слід орієнтуватися насамперед: тих, які звернуться лише одного разу і чекатимуть хороших рекомендацій або тих, хто зареєструється.

Контекст – це умови, в яких користувач отримує рекомендацію. Контекст може також включати погоду і навіть настрої користувача.

Ступінь персоналізації рекомендацій буває різною, від застосування найбільш узагальнених даних до вивчення інформації про конкретних користувачів.

Інтерфейс рекомендаційної системи відображає тип введення та виведення.

Явному введенню даних, при якому користувач вручну вказує інформацію про те, що йому подобається. Інший вид введення – неявний, коли система намагається встановити смаки користувача, спираючись на те, як він з нею взаємодіє. До різновидів виводів відносяться прогнози, рекомендації або фільтрація.

Окрім цього, К. Фальк висвітлює важливість такого поняття як пояснюваність та якість рекомендацій. Ця проблема відома як компроміс між точністю моделі та складністю інтерпретації моделі.

В «Practical Recommender Systems» розглянуті основні алгоритми фільтрації рекомендаційних систем, оцінка та тестування рекомендаційних систем, ранжування, навчання ранжуванню тощо. Пояснюються проблеми масштабованості, холодного старту, обчислення оцінок та ін.

Проте, автор не наводить прикладів або алгоритмів, що стосуються практичного застосування викладених ідей у галузі торгівлі одягом.

В роботі Амелькіна С.А. «Оценка эффективности рекомендательных систем» [9] розглянуті три задачі, які можуть бути розв'язані методами колаборативної фільтрації, а також описано, як постановка задач залежить від спектра шкали оцінок та поняття близькості оцінок, виставлених різними користувачами. Цю роботу можна використовувати як довідник, якщо розглядати колаборативну фільтрацію більш детально. Робота можна стати базою для подальших досліджень, особливо цікавим є приклад задачі вибору послідовності найкращих об'єктів.

У статті М.Косолапова «Введение в рекомендательные системы» [10] досить просто описані основні підходи з проектування контентних рекомендацій, а також наводяться приклади з використання. Автор зрозуміло наводить підхід, що використовує колаборативну фільтрацію, але не повністю розкриває завдання, які розв'язують рекомендаційні системи.

В статті R.Manjula «Content Based Filtering Techniques in Recommendation System using user preferences» [11] розглядається підхід, який призводить до формування рекомендацій, що відрізняються високою релевантністю та точністю, порівняно з традиційними методами пошуку рекомендацій. Детально розглянуті фільтрація на основі контенту та колаборативна. Автор математично обґрунтовує, що зведена функція, яка використовується для розрахунку якості рекомендації елементів залежить від рейтингового розподілу та типу сукупного аналізу, що є джерелом для майбутніх потенційних досліджень. Змістова робота, яка пропонує альтернативний підхід.

В роботі Sitnikov, D., Titova, O., Minukhin, S., Kovalenko, A., Titov, S. «Informativity of Association Rules from the Viewpoint of Information Theory» [12] розглядається метод фільтрації набору асоціативних правил, які отримані у результаті пошуку логічних залежностей. Автори наголошують, що метод дозволяє працювати з «цікавими» правилами, котрі мають такі рівні підтримки та довіри, які значно відрізняються від очікуваних. Тому, в майбутніх покращеннях досліджуваної системи електронної комерції можна застосувати цей метод фільтрації для підвищення ефективності роботи РС, адже знання наборів асоціативних правил дозволяє оптимізувати товарний асортимент системи та підвищити кількість продажів.

1.5 Постановка задачі до кваліфікаційної роботи

Метою кваліфікаційної роботи є дослідження способів урахування уподобань клієнтів при розробці систем електронної комерції.

Для урахування уподобань клієнтів при розробці систем електронної комерції обрані рекомендаційні системи. Основною задачею рекомендаційних систем є прогнозування рекомендацій для користувача таких товарів, котрі найкраще будуть відповідати його потребам.

На основі проведеного аналізу способів урахування уподобань користувачів в системах електронної комерції та алгоритмів рекомендаційних систем для досягнення поставленої мети необхідно:

- провести порівняльний аналіз алгоритмів рекомендаційних систем та обрати алгоритм для подальшого використання;
- обґрунтувати обрані критерії ефективності використання рекомендаційної системи у складі інформаційної системи електронної комерції, а саме інтернет-магазину з продажу одягу;
- провести аналіз та обґрунтувати вибір технологій створення рекомендаційної системи (СУБД, клієнтська частина, серверна частина);
- дослідити реалізовану інформаційну систему та провести її дослідну експлуатацію на обраному бізнес-ресурсі;
- дослідити ефективність роботи спроектованої інформаційної системи електронної комерції з використанням рекомендаційної системи та без її використання;
- провести аналіз отриманих результатів, аналіз обраних показників ефективності.

2 ДОСЛІДЖЕННЯ ЗАСТОСУВАННЯ РЕКОМЕНДАЦІЙНИХ СИСТЕМ В СИСТЕМАХ ЕЛЕКТРОННОЇ КОМЕРЦІЇ

2.1 Застосування рекомендаційних систем в системах електронної комерції

Донедавна люди зазвичай купували продукти, рекомендовані їм їхніми друзями або людьми, яким вони довіряють. Раніше це був основний спосіб покупки, коли виникали сумніви щодо продукту. Але з настанням цифрової ери це коло розширилося і включало онлайн-сайти, які використовують якийсь механізм рекомендацій.

Використовуючи засоби рекомендаційних алгоритмів в системах електронної комерції, компанії можуть отримувати детальну інформацію про індивідуальні запити кожного споживача та автоматично надавати продукти та послуги, відповідні їхнім вимогам. Одним з простих прикладів може бути електронний журнал, що підлаштовується під конкретного читача, пропонуючи йому статті, цікаві саме для нього, і прибираючи вже прочитані матеріали.

Рекомендаційні системи вбудовані в різні галузі, такі як відновлення знань, машинне навчання, системи підтримки прийняття рішень та класифікацію текстів. Вони виявилися корисними інструментами обробки для онлайн-клієнтів та стали одними з найпоширеніших і найпотужніших інструментів в галузі електронної комерції.

Основна мета впровадження рекомендаційних систем в електронну комерцію полягає в підвищенні коефіцієнта конверсії, тобто, можливості продавати більше рекомендованих товарів у порівнянні з тими, які зазвичай продаються без будь-якого методу рекомендаційних систем.

Окрім того, точність підвищує задоволеність користувачів та є однією з основних функцій рекомендаційних систем. Якщо користувач вважає рекомендовані товари схожими на ті, що він шукає, та цікавими, ймовірність вибору рекомендованих товарів буде висока.

Рекомендаційні системи підвищують лояльність користувачів, надаючи персоналізовані послуги. Допомагають завоювати лояльність користувачів, що в свою чергу підвищує коефіцієнт конверсії. Згодом переваги лояльних користувачів буде легше передбачити, тому точність рекомендованих товарів збільшиться. Також, система повинна мати можливість розпізнавати нових користувачів і точно надавати послуги, щоб завоювати лояльність нових користувачів.

Рекомендаційна система або механізм рекомендацій — це інструмент, який використовує низку алгоритмів, аналіз даних і навіть штучний інтелект для створення рекомендацій в Інтернеті. Ці рекомендації можуть бути індивідуальними для кожного користувача або ні, залежно від цілей кожної платформи, обсягу отриманих даних і навіть типу використовуваної технології. Їхнє призначення — збільшувати прибутки бізнесу шляхом підвищення лояльності користувачів і кількості та різноманітності товарів.

Системи рекомендацій розглядаються як програмні засоби та методи для пропонування продуктів клієнтів за допомогою автоматизованого врахування їхніх фаворитів. Надані пропозиції спрямовані на те, щоб надати клієнтам численні можливості приймати рішення.

Коли увагу зосереджено на персоналізації користувацького досвіду, система використовує дані, пов'язані з профілем та навігацією цього користувача, наприклад кліки, рейтинги та пошуки, щоб рекомендувати товари, які йому найбільше відповідають.

Перш за все, рекомендаційна система збирає інформацію і тим самим полегшує процес прийняття рішень під час показу та рекомендації вибору товари. Товар може бути продуктом, формою вмісту або навіть людиною у випадку сайтів соціальних мереж або в пропозиції друзів у соціальній мережі.

Ця зібрана інформація стосується трьох елементів:

- предмети, які слід рекомендувати;
- користувач, на якого вплинуть ці рекомендації;
- та інші користувачів, які вже взаємодіяли з платформою.

Дані користувача можуть бути двох типів: явні та неявні. Явні: складаються з інформації, наданої користувачами, як правило, у відповідь на будь-яке запитання чи запит, наприклад коментарі та оцінки. Неявні: вони створюються користувачами спонтанно і пов'язані з їхньою поведінкою під час навігації, як кліки, пошуки, час, проведений на сторінці тощо.

Якість рекомендацій безпосередньо пов'язана з кількістю та якістю отриманих даних. Отже, чим більше даних про людину та інших подібних до неї людей, тим кращими та персоналізованішими можуть бути рекомендації, що пробуджують інтерес у цього користувача та допомагають у прийнятті рішень.

Механізм рекомендацій фільтрує дані за допомогою різних алгоритмів і рекомендує користувачам найбільш релевантні товари. Спочатку він фіксує минулу поведінку клієнта і на основі цього рекомендує продукти, які користувачі можуть придбати.

Проте, частіше за все доводиться вирішувати проблему холодного старту при впровадженні рекомендаційних систем до систем електронної комерції. Адже багато алгоритмів фільтрації спираються лише на інформацію про рейтинги, що проставляються користувачами, та не аналізують контент ресурсів. І задача видачі релевантних рекомендацій для нових користувачів називається проблемою холодного старту.

Холодний старт для користувачів можливий на основі демографічних даних, які самі користувачі вказують при реєстрації. Зазвичай, ці показники враховуються як базові. Однак, демографічні дані можуть бути отримані і більше. За допомогою API соціальних мереж можна дізнатися рівень освіти, соціальний статус та інші характеристики.

Існують два основних підходи до застосування у рекомендаціях демографічної інформації про користувача:

– експертно складаються стереотипи для різних демографічних категорій. Тобто експерт сам визначає, що на холодному старті показуватиме кожній із категорій. Головним мінусом такого підходу є необхідність роботи експерта, при цьому користувачеві рекомендуватимуться лише популярні сайти суб'єктивно

підібрані екпертом. Обсяг експертної роботи істотно зростає зі зростанням кількості категорій;

– демографічні категорії визначаються автоматично, виявляючи кластери користувачів зі схожими інтересами. Рекомендації будуються на основі того, які рейтинги проставляли користувачі з тієї ж категорії, тобто того ж віку, статі, розташування та ін.

Другий підхід не вимагає залучення експертів та надає можливість створення необмеженої кількості кластерів [13].

І якщо повністю новий користувач відвідує сайт електронної комерції, цей сайт не матиме жодної попередньої історії цього користувача. Тож як же сайт рекомендує продукти користувачам у такому випадку? Одним з можливих рішень може бути рекомендація продуктів, які користуються високим попитом. Іншим можливим рішенням може бути рекомендація продуктів, які принесуть максимальний прибуток бізнесу. Загальний вигляд системи електронної комерції з використанням рекомендаційних систем наведений на рис. 2.1.

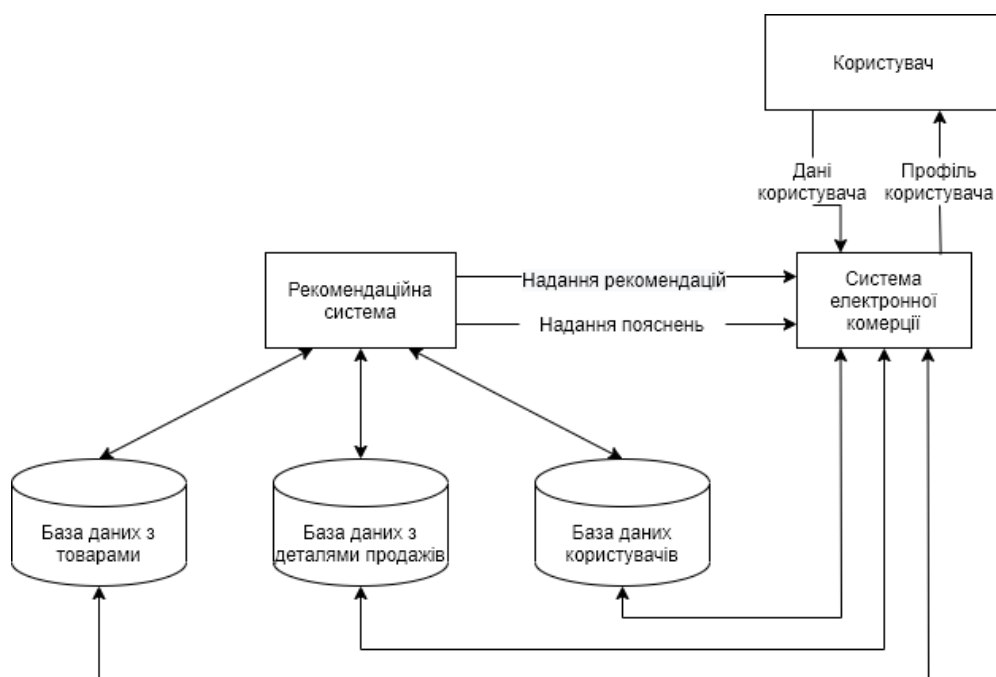


Рисунок 2.1 – Загальна схема роботи системи електронної комерції з використанням рекомендаційних систем

В результаті впровадження рекомендаційних систем до систем електронної комерції збільшуються продажі та зростає кількість клієнтів, що повернулися. Інвестиції на реалізацію та використання зазвичай повертаються через 3-6 місяців.

Якщо ми зможемо порекомендувати кілька товарів клієнту на основі його потреб та інтересів, це позитивно вплине на користувацький досвід і призведе до частих відвідувань. Отже, нині компанії створюють розумні та інтелектуальні механізми рекомендацій, вивчаючи минулу поведінку своїх користувачів.

2.2 Класифікація та порівняння алгоритмів рекомендаційних систем

Інакше системи рекомендацій можна описати як певні методи, основним призначенням яких є виконання фільтрації інформації, що надається користувачеві, з метою відображення саме тих даних, що найбільше зацікавлять клієнта. Їх підбір має здійснюватися на основі дій та вподобань клієнта, в результаті чого система рекомендацій, спрогнозувавши реакцію, запропонує кінцевому користувачеві кілька варіантів, що будуть співпадати з його інтересами.

Для розробки програмних рішень у системах рекомендацій можна використовувати декілька різних методів. До найбільш класичних відносяться алгоритми Summary-based (неперсональні), Content-based (моделі, засновані на описі товару), Collaborative Filtering (колаборативна фільтрація), Matrix Factorization (методи засновані на матричному розкладанні) та деякі інші [13].

Метод фільтрації на основі контенту – це алгоритм, що залежить від домену та більше акцентує увагу на аналізі атрибутів товарів з метою створення прогнозів.

У цьому методі використовуються формування результатів на основі атрибутів, що має кожен елемент. Такими атрибутами, наприклад для фільму, можуть слугувати його жанр, акторський склад, дата виходу до прокату та ін. Основним недоліком методу є необхідність глибокого знання та опису особливостей товарів [14].

Метод фільтрації, заснований на знаннях, розв'язує проблему «холодного старту» коли система не має достатньої кількості даних для аналізу. Цей метод

фільтрації дає рекомендації не на основі історії рейтингу користувача, а на основі конкретних запитів.

Колаборативний метод фільтрації для створення рекомендацій порівнює користувачів з відповідними інтересами та уподобаннями, розраховуючи подібність між їхніми профілями, тобто базується на моделі попередньої поведінки користувача, яка може бути побудована виключно на основі поведінки цього користувача або з урахуванням поведінки інших користувачів з подібними характеристиками. За основу для вибору результатів цей метод використовує розрахунок за рейтингами, які виставляє користувач, переглядами, діями [15].

Наприклад, якщо кілька користувачів високо оцінили товари N – Perfect та X – Good, то третій користувач, що також поставив високу оцінку для першого отримає у результатах рекомендації другий товар.

Для початку слід зазначити, що існує кілька видів колаборативної фільтрації [16]:

а) фільтрація на основі сусідства – історично вважається першим підходом і використовується у багатьох системах рекомендацій. Основним принципом такого підходу підбір групи користувачів або елементів, що схожі за певними характеристиками до шуканого користувача чи елемента, для якого буде виконуватися підбір об'єктів. Далі для прогнозу оцінок буде використовуватися комбінація ваги (тобто, наскільки близьким є збіг) та оцінок групи користувачів або об'єктів.

б) фільтрація, що заснована на моделі. Підбір рекомендацій базується на вимірах параметрів статичних моделей для оцінок користувачів, що будуються з використанням методу байесовських мереж, кластеризації, латентної семантичної моделі.

в) гібридна – поєднує у собі обидва вищезгаданих підходи. Саме гібридну фільтрацію зазвичай використовують для комерційних сайтів, оскільки він допомагає позбутися недоліків, що викликаються іншими методами та, як наслідок, значно покращує точність фінальної вибірки. Проте такий підхід достатньо важкий у реалізації [17].

На першому етапі розглянемо колаборативну фільтрацію на основі сусідства. Для аналізу візьмемо матрицю оцінок, що була виставлена користувачами для певних елементів одягу (товарів) (табл. 2.1).

Таблиця 2.1 – Матриця оцінок товарів користувачами

	Плаття «Альфа»	Плаття «Літо»	Плаття «Колібрі»	Плаття «Мрія»	Блуза «Клео»	Костюм бавовняний	Спідниця «Єва»
Юлія	4.0000	5.0000		1.0000			
Ганна		2.0000		4.0000		5.0000	4.0000
Олена	5.0000		3.0000	2.0000		1.0000	
Марта		3.0000			5.0000	4.0000	

Інтуїтивно можна зрозуміти, що користувачу Юлія варто виконувати підбір результатів серед товарів, які подобаються іншим користувачам, що є найближчими до оцінок з Юлією. Для цього необхідно отримати числовий вираз «схожості» користувачів. Наприклад, є k товарів. Оцінки кожного користувача є вектором у k -мірному просторі, виходячи з цього можна порівняти вектори для знаходження «схожих» користувачів, використовуючи один з методів, що є найбільш вживаними:

а) косинус подібності – обчислення косинусу кута між двома векторами. Найбільш подібні вектори мають значення косинуса 1, а найменш подібні – 0 (2.1)

$$\text{sim} = \cos(\Theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} \quad (2.1)$$

де, A_i та B_i є координатами векторів A та B ;

б) коефіцієнт кореляції Пірсона – результат кореляції слугує показником «схожості» об'єктів. Наприклад, якщо кореляція дорівнює -1, то об'єкти абсолютно не схожі, а якщо 1 – то максимально схожі між собою (2.2)

$$r_{xy} = \frac{\sum_{i=1}^m (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^m (x_i - \bar{x})^2 \sum_{i=1}^m (y_i - \bar{y})^2}} = \frac{\cos(x, y)}{\sqrt{s_x^2 s_y^2}} \quad (2.2)$$

де, \bar{x} , \bar{y} – вибіркові середні,

x^m , y^m , s_x^2 , s_y^2 – вибіркові дисперсії,

$r_{xy} \in [-1; 1]$;

в) Коефіцієнт Танимото – приймає значення від 0 до 1, розраховується за наступною формулою (2.3):

$$T(A, B) = \frac{N_c}{N_a + N_b - N_c} \quad (2.3)$$

де N_a – кількість оцінок для товарів у множині А,

N_b – кількість оцінок для товарів у множині В,

N_c – кількість спільних для множин А та В оцінок для товарів.

г) Манхеттенська відстань або відстань міських кварталів (2.4)

$$d(A, B) = \sum_{i=1}^n |A_i - B_i| \quad (2.4)$$

де А, В – користувачі та їх оцінки,

n – кількість предметів у матриці

г) Евклідова відстань (2.5)

$$d(A, B) = \sqrt{\sum_{i=1}^n (|A_i - B_i|)^2} \quad (2.5)$$

де А, В – користувачі та їх оцінки,

n – кількість предметів у матриці

Для наведеного прикладу першим кроком розрахуємо косинус подібності для Юлії з іншими користувачами, (табл. 2.2).

Таблиця 2.2 – Результати розрахунків

	Ганна	Олена	Марта	Сума
Юлія	0.2766	0.5436	0.3273	1.1475

Другим кроком необхідно помножити оцінки користувачів на їх коефіцієнт подібності, задля того, щоб більш подібні користувачі суттєвіше впливали на результат, аніж ті, в кого коефіцієнт є меншим, після чого знайдену суму розділити на суму знайдених збігів у всіх користувачів. Результат надано у таблиці 2.3, де напівжирним виділено товари, які Юлія ще не обирала.

Таблиця 2.3 – Результат множення оцінок на коефіцієнт подібності

	Плаття «Альфа»	Плаття «Літо»	Плаття «Колібрі»	Плаття «Мрія»	Блуза «Клео»	Костюм двійка	Спідниця «Єва»
Ганна	0.0000	0.5532	0.0000	1.1064	0.0000	1.3830	1.1064
Олена	2.7180	0.0000	1.6308	1.0872	0.0000	0.5436	0.0000
Марта	0.0000	0.9819	0.0000	0.0000	1.6365	1.3092	0.0000
Сума	2.7180	1.5351	1.6308	2.1936	1.6365	3.2358	1.1064
Резуль- тат	2.3686	1.3378	1.4212	1.9116	1.4261	2.8199	0.0642

Базуючись на отриманих результатах, найбільш високою рекомендацією можна вважати одиницю товару Костюм двійка.

Стосовно колаборативної фільтрації на основі моделі, цей вид фільтрації має декілька методів реалізації:

Колаборативна фільтрація на основі моделі має декілька методів реалізації:

а) метод Байесових мереж, в основі якого лежить ймовірнісна модель теореми Байеса. Для того, щоб обробити інформацію, необхідно створити модель Байеса

для кожного з користувачів, що хоча б раз оцінював об'єкти, базуючись на їх характеристиках. Для знаходження найближчої категорії об'єктів необхідно прорахувати належність об'єкта до кожної з категорій та обрати саме ту, що має найбільшу ймовірність;

б) методи кластерного аналізу. Для використання цього методу можна використовувати один з найвідоміших його алгоритмів – k-means. Суть цього алгоритму у розділенні об'єктів чи користувачів на певні сектори (кластери) за певними ознаками, що визначаються заздалегідь. Для виконання алгоритму необхідно вибрати k центрів кластерів, що, власне, відповідатиме кількості кластерів та зменшити сумарне квадратичне відхилення об'єктів чи користувачів від центру кластера (2.6).

$$d = \sum_i^k \sum_{x_j \in K} (x_j - u_i)^2 \quad (2.6)$$

де, k – кількість векторів

u – центр мас векторів із множини кластерів K;

в) метод на основі Марковських моделей полягає у тому, що попередній стан об'єкта напряму впливає на ймовірність його переходу до іншого стану, тобто те, що, наприклад, клієнт дивився минулого разу має прямий вплив на те, що він дивитиметься у цей раз. Марковська модель побудована на основі Марковського ланцюга, що має три компоненти: стани об'єкта, функція переходу об'єкту між станами та початковий стан.

г) методи латентного семантичного аналізу. Такий аналіз зазвичай застосовується для класифікації текстів, статей та ін. Суть такого методу у тому, що спочатку з тексту виключаються слова, що є у будь-яких текстах та не мають сенсового навантаження. Далі, в залежності від мови тексту можна використати стеммінг, тобто позбавитися закінчень для знаходження початкової форми слова. Після цього необхідно виключити слова, що зустрічаються лише 1 раз. Для отриманих в результаті даних необхідно провести сингулярне розкладання (2.7).

$$M = U \cdot W \cdot V^t \quad (2.7)$$

де, U , V^t – ортогональні матриці,

W – діагональна матриця

Аналізуючи матриці, що були отримані внаслідок сингулярного розкладання можна виділити перелік ключових слів, згідно з яким потім можна виконати групування контенту за тематикою. У випадках з системами рекомендації таким чином можна знайти які саме рекомендації знаходяться у сфері інтересів користувача.

Проте у методу колаборативної фільтрації є суттєві недоліки, про які не варто забувати, коли планується його використання. Серед таких:

а) проблема холодного старту. Найбільша проблема цього методу, пов'язана з тим, що нові користувачі ще не мають історії оцінок, покупок, переглядів і через це система рекомендації не може показати дійсно необхідні результати;

б) масштабованість. При великій кількості користувачів алгоритм колаборативної фільтрації є занадто важким для виконання;

в) синонімія. Цей метод не може розпізнати певні категорії об'єктів, якщо вони матимуть синонімічні характеристики. Наприклад, категорії фільмів «фільми для навчання» та «навчальні фільми» будуть сприйняті як різні, а не однакові;

г) шахрайство. Недобросовісні користувачі або ж компанії можуть навмисно підвищувати свій рейтинг, або ж занижувати чужий рейтинг, використовуючи ботів, підставні акаунти та ін.

г) проблема білих ворон обумовлена користувачами, що мають унікальний, відмінний від усіх смак. Таким користувачам важко що-небудь порекомендувати;

д) зменшення різноманіття, що пов'язано з тим, що нові товари не мають можливості потрапити до рекомендацій, оскільки мають значно нижчу оцінку, товари, що вже давно на ринку. Як наслідок такої логіки отримується суттєве зменшення різноманіття товарів.

Алгоритми колаборативної фільтрації і фільтрації, заснованої на контенті більш адаптивні, адже їм необхідна тільки базова інформація про предмет. Вони використовуються для рекомендації низького ступеня участі, наприклад, при виборі одягу. Це означає, що користувачі прикладають менше зусиль для оцінки предмета [18].

В результаті проведеного аналізу можна зробити висновок, що найбільш ефективними методами фільтрації є: фільтрації на основі контенту (item-based) та колаборативної фільтрації (user-based), з точки зору адміністратора системи інтернет-магазину, для використання в системах електронної комерції з метою урахування уподобань клієнтів.

3 РОЗРОБКА АЛГОРИТМУ ВРАХУВАННЯ УПОДОБАНЬ КЛІЄНТІВ ДЛЯ СИСТЕМИ ЕЛЕКТРОННОЇ КОМЕРЦІЇ

3.1 Обґрунтування вибору математичної моделі

В деякій системі електронної комерції є зареєстровані користувачі (множина $U = \{u_i\}, (i = 1, n)$), в конкретній ситуації вибору користувач u_v здійснює пошук певного товару. На його попередніх виборах визначена множина елементів системи, які можуть бути потенційно йому рекомендовані.

Позначимо цю множину елементів як $S = \{sj\}, (j = 1, m)$. Для створення рекомендації можуть бути застосовані різні методи, які відрізняються чутливістю до повноти вхідних даних.

Для розв'язання поставленої задачі до кваліфікаційної роботи необхідно знайти найбільш релевантний об'єкт для користувача uj , за наведеною формулою (3.1), використовуючи різні методи створення рекомендацій, оцінити їх якість на основі отриманих результатів за критеріями.

Задача знаходження найбільш релевантного об'єкта:

$$\forall u \in U, s'_u = \operatorname{argmax}_{s \in S} h(u, s), \quad (3.1)$$

де U – множина користувачів, u – вибраний користувач, S – множина товарів системи, які потенційно можуть бути рекомендовані користувачу, s – вибраний товар, h - функція, яка показує ступінь зіставлення деякого товару s з деяким користувачем u .

Тому, задача зводиться до знаходження такого товару $s' \in S$, для якого ступінь зіставлення з користувачем $u \in U$ є максимальним (у випадку рекомендації одного товару).

На точність рекомендації впливає не лише інформація всередині системи, а й вибір функції h [19].

Найчастіше рекомендаційні системи застосовуються для створень рекомендацій користувачам або групам користувачів за описаними схемами та використовують ряд різних технологій. Порівняно два основні типи фільтрації рекомендаційної системи:

– системи фільтрації на основі контенту (item-based) досліджують властивості рекомендованих елементів. Наприклад, якщо користувач комерційної платформи переглянув багато товарів у розділі «Плаття», то рекомендуємо товар, який класифікується в базі даних як «Плаття»;

– системи колаборативної фільтрації (user-based) рекомендують товари на основі показників подібності між користувачами та/або елементами. Рекомендовані користувачеві елементи ті, яким віддають перевагу подібні користувачі. Така система рекомендацій може використовувати за основу пошук подібності та кластеризацію.

Фільтрація на основі контенту, перш за все, опирається на інформацію про контент в системі, а не про користувачів. Рекомендації базуються на знаходженні схожих товарів до тих, які користувач вже оцінив раніше. Для цього необхідно створити профіль користувача та профіль товару. Після цього, на основі параметрів товарів системи, можна зробити висновок щодо відповідності конкретного товару конкретному користувачу. Для опису товарів системи та створення їх профілю рекомендаційні системи ставлять у відповідність кожному елементу певний набір ключових слів.

Алгоритми фільтрації на основі контенту мають рекомендувати товари на основі накопичених знань користувачів. Ця техніка пов'язана з порівнянням інтересів користувачів із характеристиками товару, тому важливо надати значну характеристику товарів у системі. Першим пріоритетом має бути вибір улюблених товарів для кожного покупця [20].

Обидві стратегії можна застосовувати у можливій комбінації. По-перше, користувачеві надається список функцій, щоб вибрати найбільш цікаві товари.

По-друге, алгоритми ведуть облік усіх товарів, обраних користувачем у минулому, і формують дані про поведінку клієнта. Профіль покупця обертається

навколо вибору, смаків і переваг покупця, і формує рейтинг покупця. Він включає в себе кількість разів, які один покупець натискає на зацікавлені товари або скільки разів ці товари з'являлися в списках бажань. Роботу рекомендаційної системи, яка використовує тип фільтрації на основі контенту наведено на рис. 3.1.

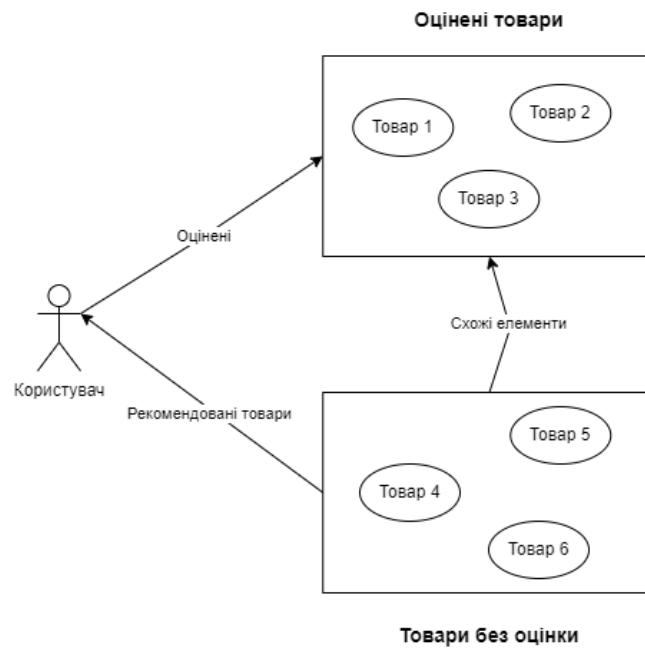


Рисунок 3.1 – Схематичне зображення роботи фільтрації товарів на основі контенту в рекомендаційній системі

Фільтрація на основі контенту полягає у подібності між елементами. Близькість і схожість товару вимірюються на основі схожого вмісту товару. Коли говориться про вміст, він включає жанр, категорію предмета тощо.

До переваг фільтрації товарів на основі контенту належать:

- незалежність користувачів. Метод, заснований на контенті, повинен лише проаналізувати елементи та профіль користувача для рекомендації;
- прозорість. Метод на основі вмісту може показати рекомендації товарів на основі певних функцій;
- відсутність холодного старту. Нові елементи можуть бути запропоновані, перш ніж їх оцінить значна кількість користувачів [21].

Рекомендаційні системи, які використовують колаборативну фільтрацію, на відміну від фільтрації на основі контенту аналізують дані про користувачів, а не про товари. Кожному користувачу у відповідність ставиться деяка група користувачів зі схожими вподобаннями. На основі цього формулюється гіпотеза, що користувачі, які однаково оцінили деякі об'єкти раніше, швидше за все, однаково оцінюватимуть інші об'єкти в майбутньому. Схематичне зображення роботи рекомендаційної системи на основі колаборативної фільтрації наведено на рис. 3.2.

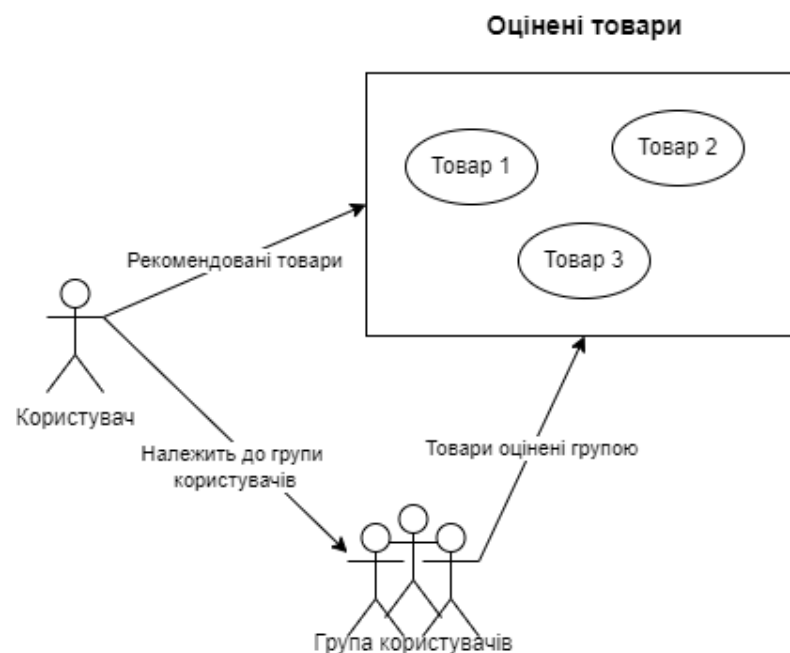


Рисунок 3.2 – Схематичне зображення роботи колаборативної фільтрації в рекомендаційній системі

В такій системі враховується лише поведінка користувачів. Недостатньо тільки вмісту та інформації профілю. Користувач, який дає позитивну оцінку продуктам, буде пов'язаний з поведінкою інших користувачів, які дають подібну оцінку.

Основна ідея цього підходу — пропонувати нові товари на основі близькості в поведінці подібних користувачів.

Загальні проблеми колаборативної фільтрації:

– матриця рейтингів, як правило, дуже розріджена (sparse) – зазвичай і користувачів, і товарів багато, а рейтингів набагато менше, адже середній користувач оцінює зовсім небагато продуктів; інші ж елементи матриці невідомі, і їх треба передбачувати;

– проблема холодного старту. Для користувачів – як рекомендувати товари новим користувачам або тим, хто вже щось мінімально оцінив. Для продуктів – скільки потрібно рейтингів нового продукту, перш ніж його можна буде впевнено рекомендувати;

– проблема масштабованості. Чим більше користувачів в системі, тим складніше виконувати розрахунки системою та надавати точні рекомендації користувачам.

Що до переваг колаборативної фільтрації можна, то можна відзначити наступні:

– випадковість. Модель може допомогти користувачам відкрити нові інтереси. Система може не знати, що користувач зацікавлений в певному товарі, але модель всеодно може рекомендувати його, оскільки подібні користувачі цікавляться цим товаром;

– відсутність необхідності знань предметної області, адже вбудовування нових товарів вивчаються автоматично;

– проблема холодного старту. Навіть якщо інформація про товар недоступна, всеодно можна передбачити рейтинг товару, не чекаючи, поки користувач його придбає;

– гнучкість. Відображає зміни в інтересах користувачів з часом: фокусування виключно на вмісті не забезпечує гнучкості щодо точки зору користувача та його уподобань;

– легкий старт. До певної міри системі потрібна лише матриця зворотного зв'язку для навчання моделі факторизації матриці. Зокрема, системі не потрібні контекстні функції [22].

Також діє принцип накопичення знань: чим більше товарів оглянути та оцінені користувачем, тим точніше формуються рекомендації. Тому, все частіше системи пропонують користувачам залишати оцінки для підвищення якості роботи системи.

Точність рекомендацій залежить від відношення між кількістю користувачів та об'єктів. Якщо кількість користувачів більше, ніж кількість об'єктів, краще використовувати алгоритм, заснований на порівнянні об'єктів (item-based).

В інших випадках використовують метод, що базується на порівнянні користувачів (user-based).

Можна підсумувати, що колаборативна фільтрація забезпечує потужну прогнозну силу для рекомендаційних систем і водночас вимагає найменшої інформації, тому в роботі досліджено впровадження колаборативної фільтрації до рекомендаційної системи, шляхом знаходження найбільш релевантного об'єкта для користувача за методом колаборативної фільтрації на основі сусідства, для врахування уподобань користувачів та інтеграція РС до ІС.

3.2 Опис алгоритму розв'язання задачі

Для коректної роботи колаборативного методу фільтрації та, власне, застосування його у системах рекомендації в ІС має бути зареєстровано більше за 1 користувача, що вже виставляв оцінки хоча б 2 товарам. Вхідними даними є характеристики товарів та їхні оцінки, що виставляються користувачами системи явно (оцінювання проводиться прямим введенням значень в систему) або неявно (тобто аналіз кількості переглядів, часу перебування на сторінці та ін.). Ці оцінки можуть бути внесені з використанням будь-якого числового або символічного ранжування. Наприклад, це можуть бути цілі числа від 1 до 5 або букви англійського алфавіту від А до Е. Вихідними даними є рекомендації, що представлені системою після аналізу.

Можна уявити ці дані як матрицю, кожен рядок якої відповідає користувачеві, а стовпець – товару.

Існує сильно розряджена матриця, що складається з рейтингів (рядки матриці вектор уподобань для кожного користувача), які користувачі надали продуктам (стовпці матриці вектор оцінок користувачів для кожного товару):

$$R = (r_{u,i})_{u=1,i=1}^{U,I}, \quad (3.2)$$

де R матриця елементів, $r_{u,i}$ невідомі оцінки, $r'_{u,i}$ вже існуючі оцінки в матриці.

Задача: передбачити оцінки $r_{u,i}$, знаючи вже існуючі оцінки в матриці $r'_{u,i}$; а також ті, які з невідомих елементів будуть максимальними у своєму рядку, тобто, які товари найбільше сподобаються тому чи іншому користувачеві.

Алгоритм передбачення оцінок:

а) знайти, наскільки інші користувачі в базі даних ІС схожі на певного користувача;

б) враховуючи оцінки інших користувачів, передбачити, яку оцінку поставить обраний користувач певному товару, враховуючи з найбільшою вагою тих користувачів, які найбільше схожі на обраного користувача.

По-перше, залишимо в векторах уподобань для кожного користувача та оцінок користувачів для кожного товару ті елементи, для обох з яких відомі значення, тобто лише ті продукти, які оцінили обидва користувачі, або тільки тих користувачів, які обидва оцінили обраний продукт.

В результаті необхідно визначити, наскільки схожі два вектори дійсних чисел. Класичне рішення задачі обчислити коефіцієнт кореляції: для двох векторів переваг користувачів u та j коефіцієнт кореляції Пірсона

$$w_{u,j} = \frac{\sum_a (r_{u,a} - \bar{r}_u)(r_{j,a} - \bar{r}_j)}{\sqrt{\sum_a (r_{u,a} - \bar{r}_u)^2} \sqrt{\sum_a (r_{j,a} - \bar{r}_j)^2}} \quad (3.3)$$

де \bar{r}_u середній рейтинг, що виставлений користувачем u , а $r_{u,j} \in [-1;1]$.

Матриця оцінки товарів в ІС інтернет-магазин жіночого одягу.

Таблиця 3.1 Матриця оцінки товарів в ІС інтернет-магазин жіночого одягу

	Плаття «Літо»	Сорочка «№1»	Капелюх «Захід»	Штани «Кюлоти»	Футболка біла
Юлія	?	3	4	5	2
Саша	3	5	2	2	5
Наталя	5	3		4	3
Ганна	5	5	5		4
Марія	2	3		2	2

Результати обчислення кореляції для user-based рекомендацій між вектором уподобань Юлії та інших користувачів системи наведено в табл. 3.1.

Таблиця 3.1 Кореляція для user-based рекомендацій

Саша	-0.8944
Наталя	0.9449
Ганна	0.8660
Марія	0.9049

Для того щоб скористатися обчисленими оцінками схожості товарів необхідно приблизити новий рейтинг обраного користувача як середній рейтинг цього користувача, а також додати відхилення від середнього рейтингів інших користувачів:

$$\hat{r}_{u,i} = \bar{r}_u + \frac{\sum_j (r_{j,i} - \bar{r}_j) * w_{u,j}}{\sum_j |w_{u,j}|} \quad (3.4)$$

де u товар, $\hat{r}_{u,i}$ накреслене передбачення, \bar{r}_u середній рейтинг, що виставлений користувачем u , $r_{u,j} \in [-1;1]$; $w_{u,j}$ коефіцієнт кореляції.

Очікувана оцінка (найкраще передбачення) що до Плаття «Літо», оціненого користувачем Юлія в ІС інтернет-магазин з продажу жіночого одягу, згідно проведених обчислень, дорівнює 3,9.

Але під час обчислень використано середнє зважене замість простого середнього та середні оцінки відняті від кожного рейтингу користувачів. Проблема полягає в типах користувачів, з якими працює ІС. Починаючи з того, що люди часто оцінюють товари за різними шкалами. Користувач може бути оптимістично налаштованим, де оцінить товар, як 4 з 5, але інший користувач, який є менш оптимістичним або має високі стандарти, може оцінити товар як 2 з 5. Перший оцінив як 4, а другий як 2.

Для покращення алгоритму підвищено його ефективність шляхом нормалізації рейтинга користувачів. Один із способів, що застосовуються обчислення $\hat{r}_{u,i}$, тобто оцінки як середньої, яку користувач дає кожному елементу, а також врахування деякого відхилення, котре показує, наскільки цей елемент кращий чи гірший за середній.

Застосована подібність косинуса для обчислення ваги (3.4) та використане поняття сусідства.

3.3 Оцінка якості роботи алгоритму

Окрім цього, використання такого алгоритму має сенс, коли він є ефективним. Тому, важливим кроком для впровадження або розробки систем рекомендацій є оцінка якості роботи. Існує багато метрик для оцінки точності роботи рекомендаційних систем. Порівняно основні з них.

Метрика MAE, яка вимірює середню величину помилок в наборі прогнозів, без урахування їх напрямлення. MAE вимірює точність для безперервних змінних та означає, що всі індивідуальні відмінності однаково зважуються в середньому (3.5).

$$MAE = \frac{1}{n} \sum_{j=1}^n |r'_{u,o} - r_{u,o}| \quad (3.5)$$

де u – користувач, n – множина пар користувачів u та об'єктів o , $r'_{u,o}$ та c – відповідно оцінений системою та реальний рейтинг товару o користувача u .

Метрика root mean square error (RMSE) використовується для обчислення ефективності та є оцінкою точності передбачення інтересу або ж середньоквадратичне відхилення (3.6).

$$RMSE = \sqrt{\frac{1}{|n|} \sum_{(u,o) \in n} (r'_{u,o} - r_{u,o})^2} \quad (3.6)$$

де, n – множина пар користувачів u та товарів (об'єктів) a ;

$r'_{u,o}$ – оцінка рейтингу товару o для користувача;

$r_{u,o}$ – оцінка поведінки РС.

Формула описує різницю між прогнозом і відповідними спостережуваними значеннями, які зводять у квадрат, а потім усереднюють за вибіркою. Оскільки помилки зводяться в квадрат до їх усереднення, RMSE надає відносно велику вагу великим похибкам. Це означає, що RMSE є найбільш корисним, коли великі помилки особливо небажані.

3.4 Висновки порівняння метрик MAE та RMSE

MAE та RMSE метрики виражають середню помилку прогнозування моделі в одиницях змінної. Обидві метрики можуть варіюватися від 0 до нескінченності і байдужі до напрямку помилок. Це негативно-орієнтовані оцінки, які означають, що чим нижчі значення, тим краще.

Також, MAE та RMSE можна використовувати разом для діагностики коливань помилок у наборі прогнозів. RMSE завжди буде більшою або

дорівнюватиме MAE; чим більша різниця між ними, тим більша дисперсія окремих помилок у вибірці. Якщо $RMSE=MAE$, то всі помилки мають однакову величину.

Що до відмінностей між наведеними метриками, то взяття квадратного кореня із середньої квадратної похибки має деякі наслідки для RMSE. Оскільки помилки зводяться в квадрат до їх усереднення, RMSE надає відносно велику вагу великим похибкам. Це означає, що RMSE має бути більш корисним, коли великі похибки особливо небажані.

Підсумовуючи порівняні метрики, RMSE має низку переваг, адже більше приділяє увагу великим похибкам, тому може бути більш доречним у випадках, коли необхідно мінімізувати великі похибки.

З точки зору інтерпретації, MAE є більш доцільною. RMSE не описує лише середню похибку і має інші наслідки, які важче розібрати та зрозуміти. Але з іншого боку, ще одна явна перевага RMSE над MAE полягає в тому, що RMSE уникає використання абсолютного значення, яке небажано в багатьох математичних розрахунках [23].

4 ІНТЕГРАЦІЯ АЛГОРИТМУ РЕКОМЕНДАЦІЙНОЇ СИСТЕМИ ДО ІС

4.1 Формування вимог до програмного засобу

Для реалізації системи електронної комерції, розробки та інтеграції алгоритму рекомендаційної системи сформовані вимоги до програмного засобу.

За способом організації реалізована клієнт-серверна архітектура, в якій мережне навантаження розподілено між постачальниками послуг, які називаються серверами, і замовниками послуг, які називаються клієнтами.

Принцип роботи клієнт-серверної архітектури: комп'ютер-клієнт за допомогою мережі Інтернет надсилає запит на дані сервера, а сервер приймає процес запиту та доставляє пакети даних клієнту назад [24].

Для впровадження системи електронної комерції та її подальшої інтеграції з РС триланкова клієнт-серверна архітектура є найбільш ефективною. Схематичне зображення архітектури клієнт-сервер наведено на рис. 4.1.

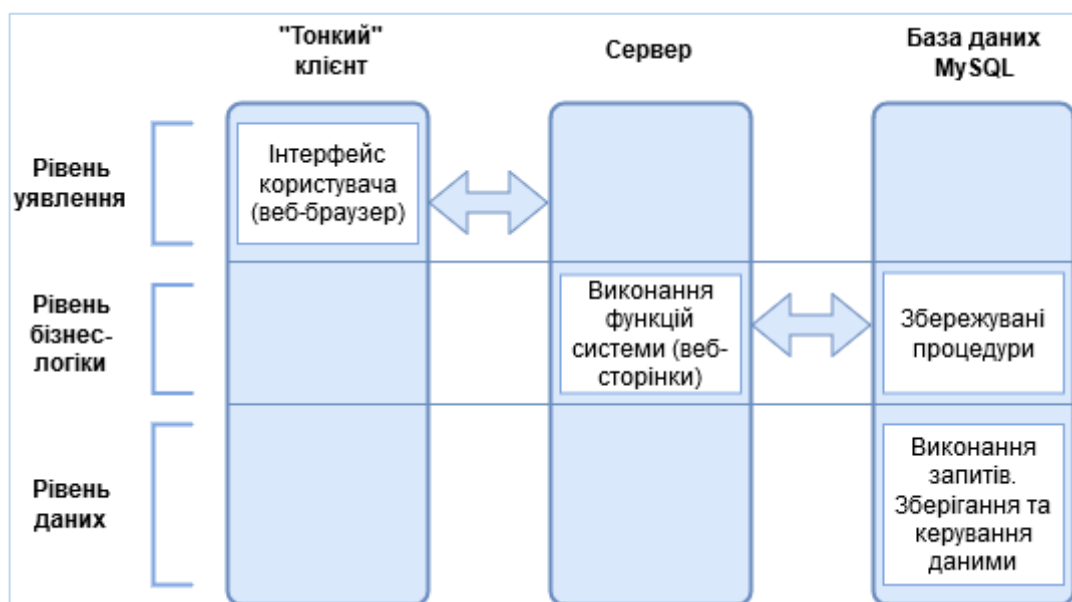


Рисунок 4.1 – Схеми триланкової архітектури клієнт-сервер

Враховуючи цілі створення, ІС електронної комерції з інтегрованим алгоритмом РС має відповідати вимогам:

- серверна частина має включати БД, що знаходиться у ЗНФ для забезпечення цілісності даних шляхом руйнування залежностей;
- БД має бути розроблена з використанням СУБД MySQL;
- клієнтська частина інформаційної системи повинна бути реалізована у вигляді веб-сайту, використовуючи мову програмування JavaScript;
- ІС має формувати переліки рекомендованих товарів шляхом урахування вподобань користувачів;
- ІС має дозволяти користувачам оцінювати товари від 1 (найгірша оцінка) до 5 (найкраща оцінка);
- інтегрований алгоритм має дозволяти користувачам отримувати список рекомендованих товарів в інтернет-магазині жіночого одягу;
- клієнтський інтерфейс доступу до розробленої ІС має бути зрозумілим та простим для користування.

4.2 Розробка діаграм варіантів використання

Для порівняння результатів інтеграції алгоритму РС до системи електронної комерції (інтернет-магазину жіночого одягу) розроблена діаграма варіантів використання (Use case diagram).

Use case допомагає проектувати систему з точки зору кінцевого користувача. Це ефективна техніка для передачі поведінки ІС в термінах користувача шляхом визначення всієї зовнішньої поведінки системи [26].

Системне оточення розробленої діаграми складають три основні актори системи електронної комерції: «Користувач», «Авторизований користувач» та «Адміністратор».

На рис. 4.2 наведена діаграма варіантів використання для системи електронної комерції до впровадження алгоритму РС.

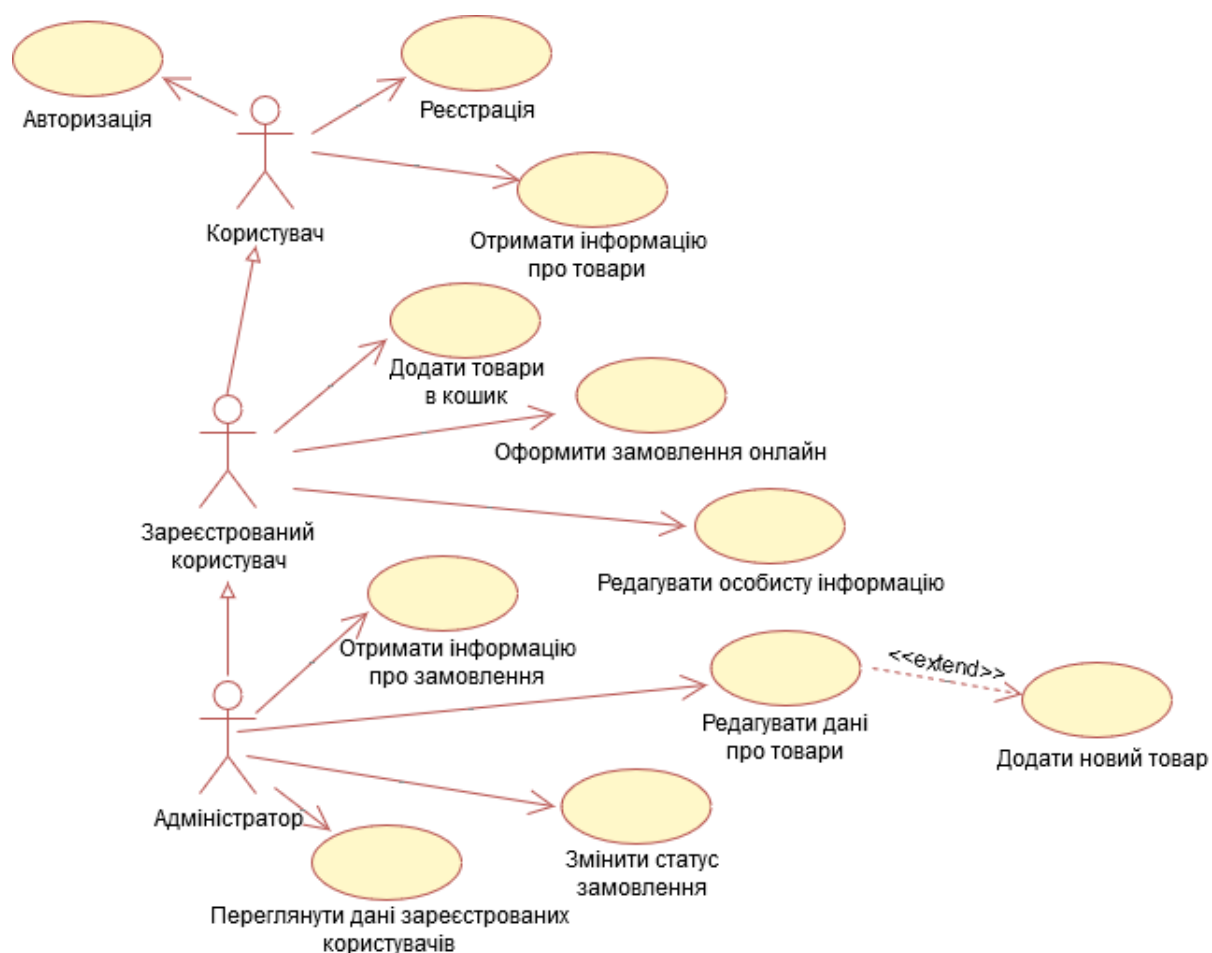


Рисунок 4.2 – Діаграма варіантів ІС електронної комерції до впровадження алгоритму РС

Адже одним з етапів кваліфікаційної роботи є інтеграція розробленого алгоритму рекомендаційної системи до системи електронної комерції необхідно порівняти діаграму варіантів використання до впровадження алгоритму та після. На рис. 4.3 наведена діаграма ІС електронної комерції після впровадження алгоритму РС.

Набір дій, які може виконувати актор «Зареєстрований користувач», розширений двома новими прецедентами: «Оцінити товари» та «Переглянути рекомендовані товари».

Так як метод колаборативної фільтрації здатен обробляти інформацію про товари не тільки після того, як їх оцінять користувачі, а й рекомендувати товари згідно вподобань схожих користувачів із групи користувачів, до якої додано нового

Зареєстрованого користувача, то прецедент «Переглянути рекомендовані товари» є незалежним від прецеденту «Оцінити товари».

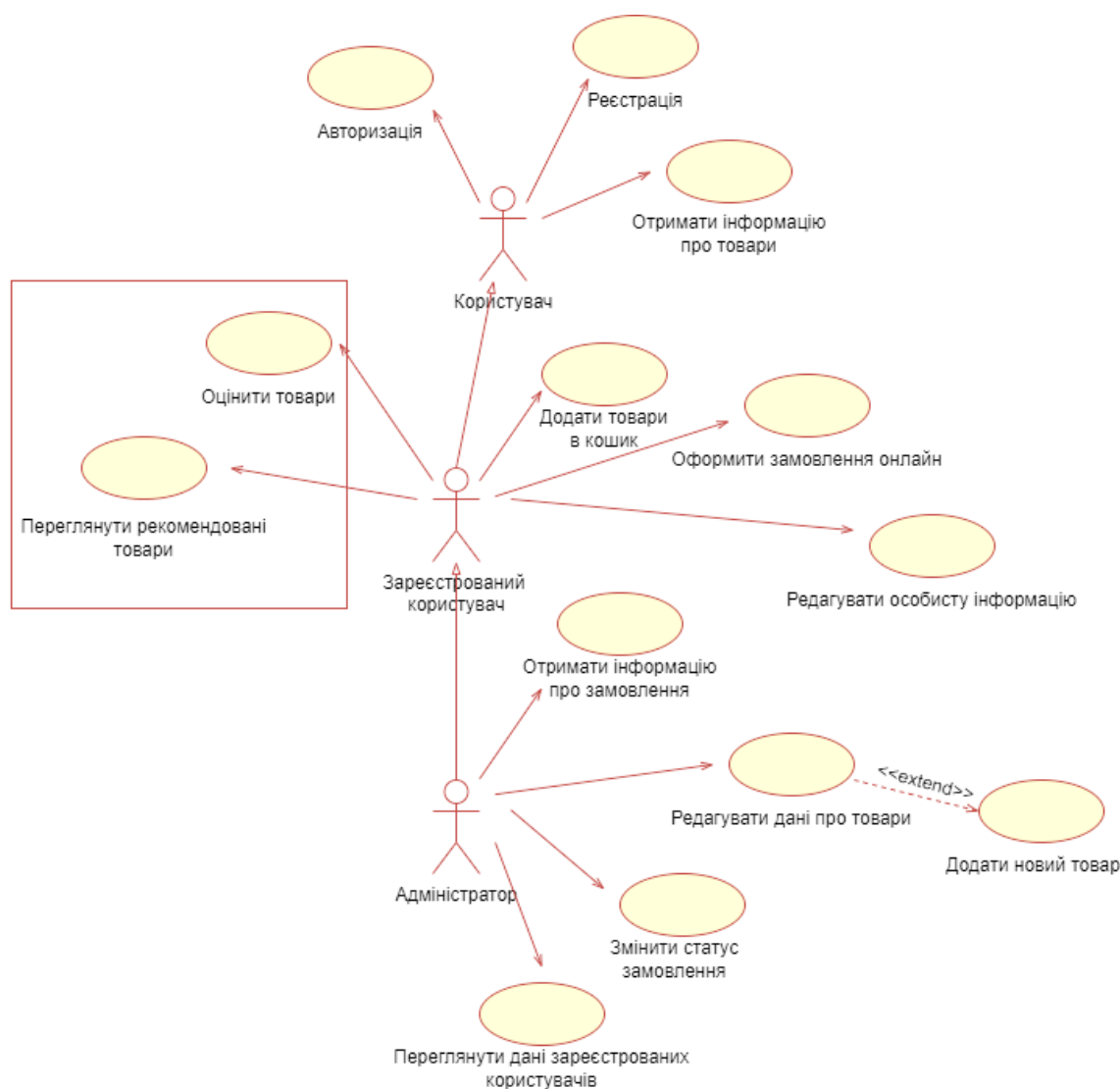


Рисунок 4.3 – Діаграма варіантів ІС електронної комерції з інтегрованим алгоритмом РС

Оновлена діаграма варіантів використання з інтегрованим алгоритмом РС дозволила зрозуміти семантику основних сценаріїв використання та визначити сутності, що взаємодіють з системою електронної комерції ззовні.

4.3 Обґрунтування вибору технологій створення системи електронної комерції

Клієнтський інтерфейс системи електронної комерції з продажу жіночого одягу розроблений з використанням мови гіпертекстової розмітки HTML та CSS (каскадних таблиць стилів).

Для повноцінної роботи клієнтської сторони системи була застосована мова програмування JavaScript.

Окрім цього, для розробки веб-додатку системи електронної комерції використаний патерн MVC. Схема роботи якого представлена на рис 4.4.

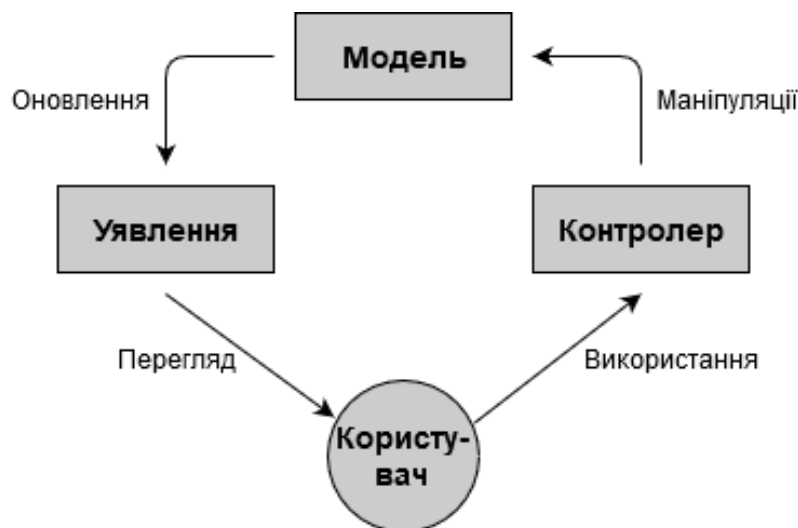


Рисунок 4.4 – Схема роботи патерну MVC

Перевагою що до використання model-view-controller вважають чітке розділення логіки інтерфейсу користувача та логіки додатку [27].

Для розробки серверної частини системи обрана мова програмування PHP. У зв'язці з PHP використовувався фреймворк Laravel, який допоміг розробити гнучку систему електронної комерції, а також алгоритм колаборативної фільтрації PC.

4.4 Обґрунтування вибору СУБД

Для розробки бази даних системи електронної комерції обрана СУБД MySQL версії 5.7 з відкритим початковим кодом. MySQL зазвичай використовується разом з PHP та веб-сервером Apache.

До переваг можна віднести низку вартість, високу продуктивність та гнучкість [28]. Більш того, MySQL підходить для створення та підтримки невеликих систем, як розроблена в рамках кваліфікаційної роботи система електронної комерції з урахуванням уподобань користувачів.

Тому СУБД MySQL відповідає цілям та завданням розробки.

4.5 Огляд розробленої бази даних з використанням СУБД MySQL

На рис. 4.5 зображена схема розробленої бази даних для компонентів веб-системи.

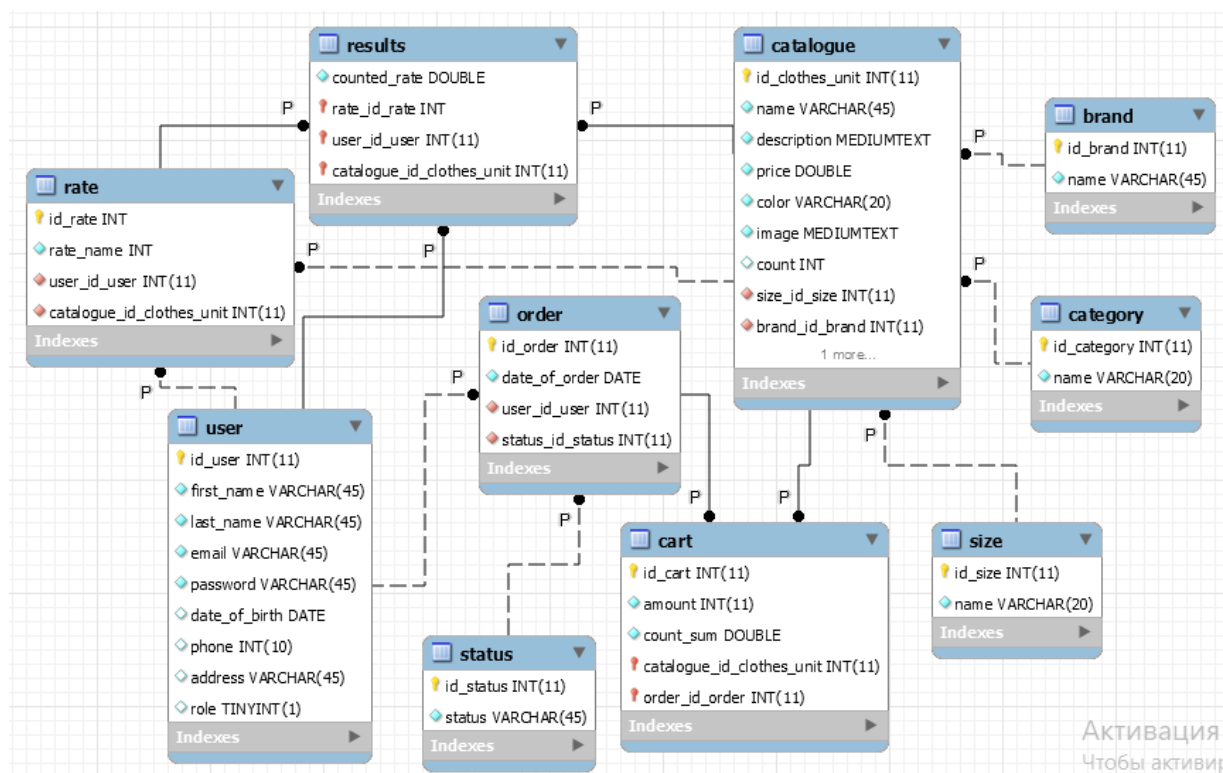


Рисунок 4.5 – Схема розробленої бази даних

Створена база даних MySQL в програмному засобі MySQL Workbench повністю відповідає встановленим вимогам в процесі моделювання.

4.6 Розробка інтерфейсу клієнтської частини системи електронної комерції

Створена головна сторінка для обраної галузі електронної комерції – інтернет-магазину жіночого одягу. На рис. 4.6 наведена частина сторінки.

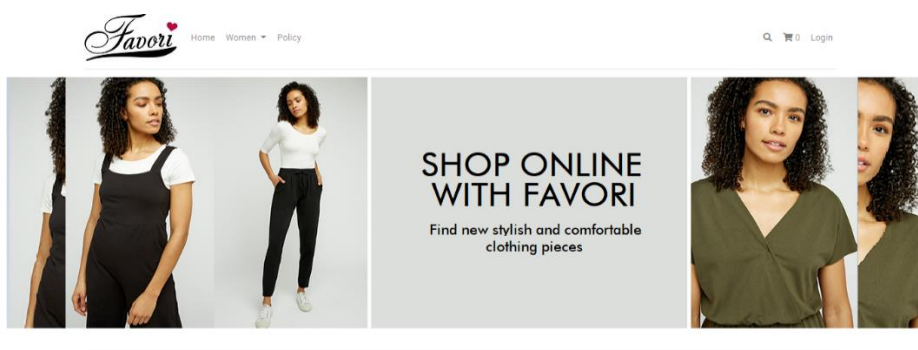


Рисунок 4.6 – Інтерфейс головної сторінки ІС

Форма реєстрації та авторизації наведена на рис. 4.7. Для реєстрації користувач має вказати ім'я, прізвище, електронну пошту та пароль. Для авторизації електронну пошту та пароль (рис. 4.8).

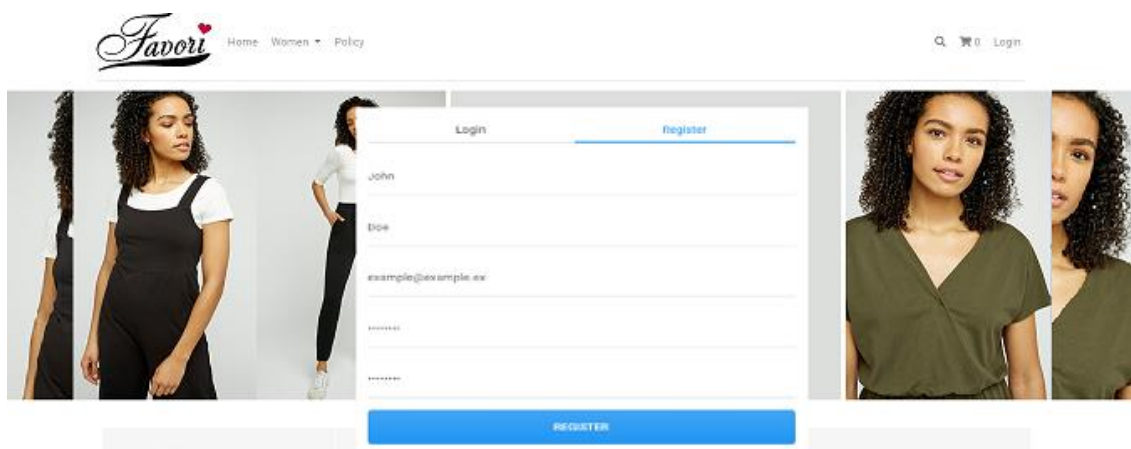


Рисунок 4.7 – Форма реєстрації ІС електронної комерції

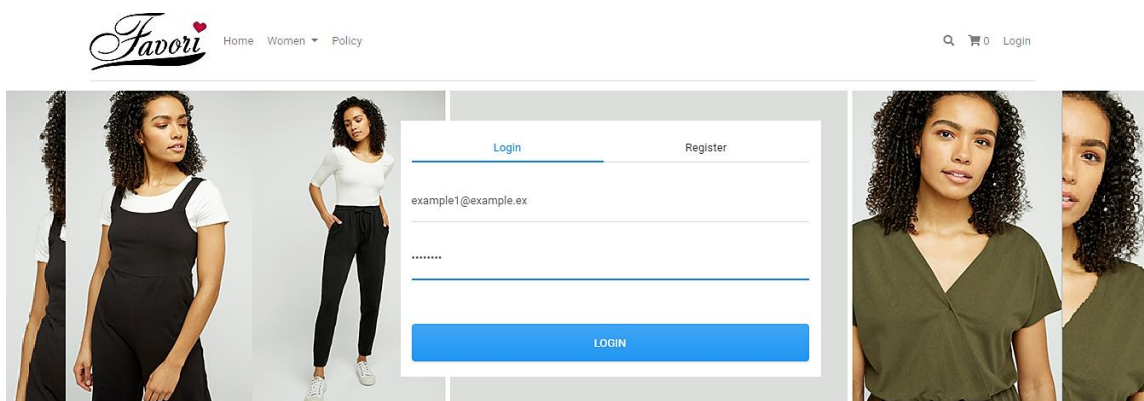


Рисунок 4.8 – Форма авторизації в системі

На рис. 4.9 наведений інтерфейс каталогу товарів системи електронної комерції. На наведеній сторінці з товарами категорії «Sweaters» користувачам надається можливість оглянути товари, оцінити їх, а також побачити рейтинг того чи іншого товару.

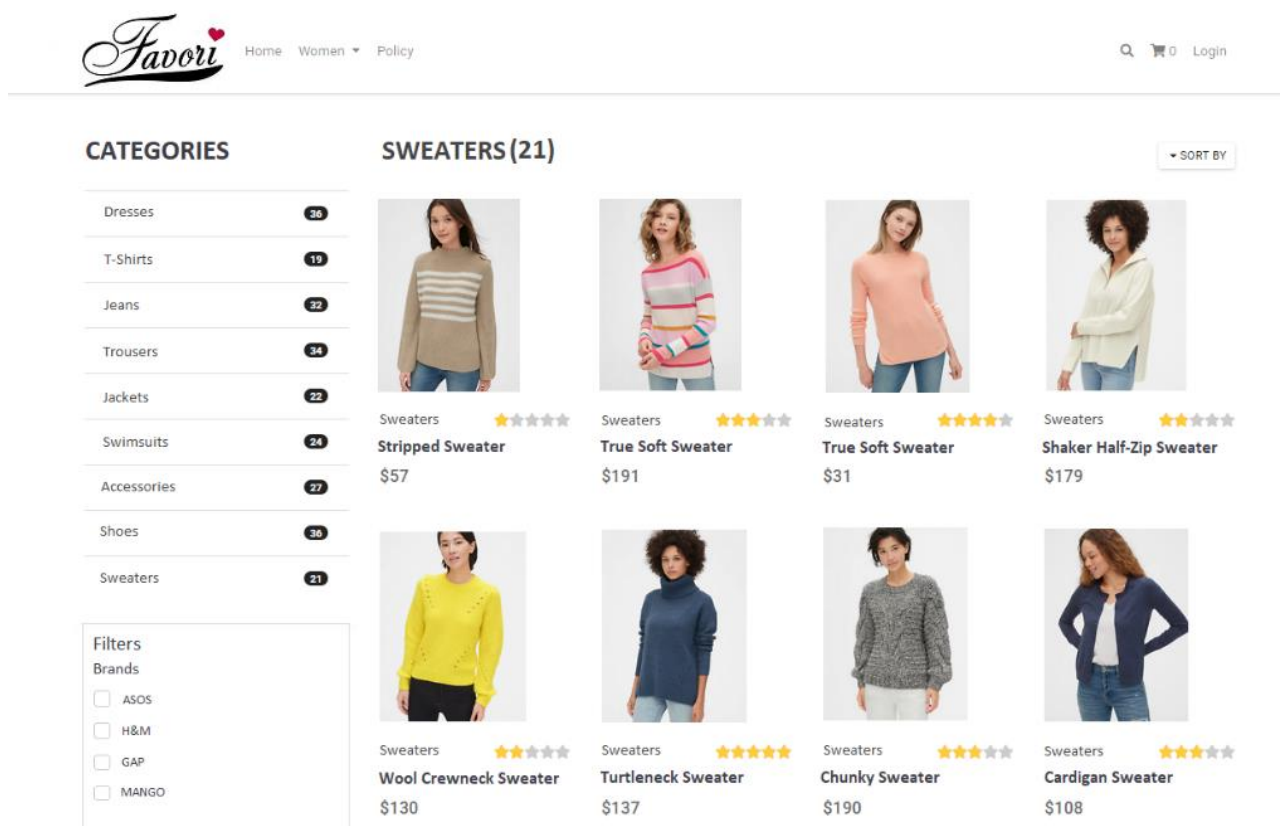


Рисунок 4.9 – Каталог одягу

На рис. 4.10 наведений інтерфейс, що відображає список рекомендованих користувачеві товарів на основі його уподобань та покупок, а також користувачів, що входять разом з ним до спільної групи користувачів, виділеної системою, згідно методу колаборативної фільтрації.

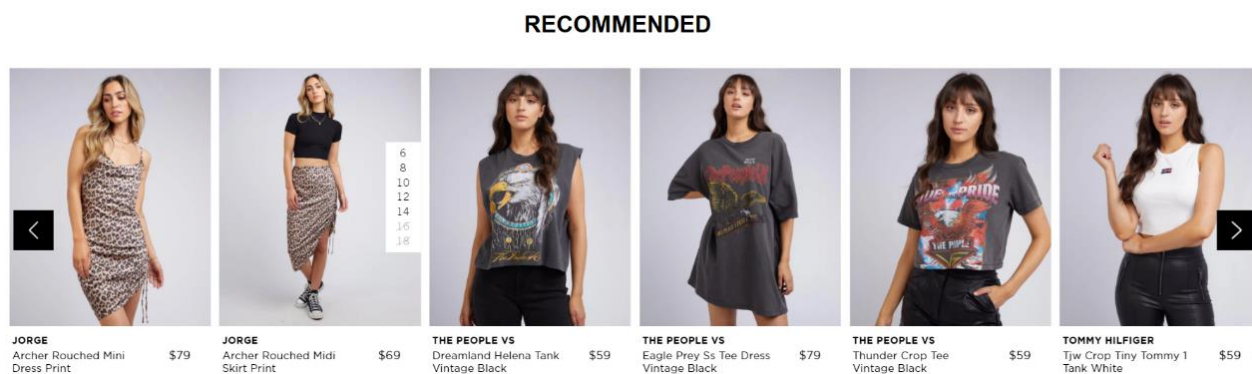


Рисунок 4.10 – Підбірка рекомендованих товарів

Отже, клієнтський інтерфейс розроблений в мінімалістичному стилі згідно встановлених вимог.

4.7 Тестування розроблених компонентів системи

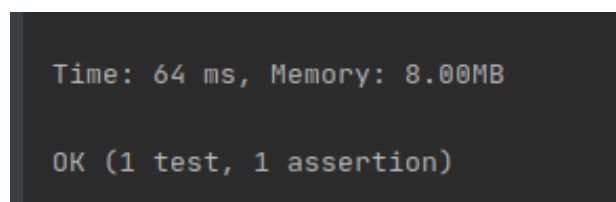
На етапі тестування проведено дослідження та виявлення дефектів та помилок, що могли бути допущені на етапі системи електронної комерції.

Для перевірки невеликої, ізольованої частини коду розробленого ПЗ використаний такий вид тестування, як модульне або юніт-тестування. З використанням мови програмування PHP, а також фреймворку PHPUnit розроблені декілька тестів, один з яких наведений на рис. 4.11. Наведений тест перевіряє, як інформація про товар або товари підготовляється до зберігання, наприклад, у випадку, коли інформація про товари була змінена адміністратором.

```
public function testPrepareData()  
{  
    $model = $this->getModel();  
    $data = [  
        ShopImport::SKU_COL => 'test'  
    ];  
    $class = new ReflectionClass($model);  
    $method = $class->getMethod('prepareData');  
    $method->setAccessible('true');  
    $this->assertEquals($data, $method->invokeArgs($model, [$data]));  
}
```

Рисунок 4.11 – Код тесту для перевірки підготовки інформації до зберігання

Результат проведення перевірки модуля наведений на рис. 4.12.



```
Time: 64 ms, Memory: 8.00MB  
  
OK (1 test, 1 assertion)
```

Рисунок 4.12 – Тест першого модуля

Другий юніт-тест, наведений на рис. 4.13 перевіряє, чи правильно отримується список атрибутів категорії товару (її назва, опис тощо), який необхідний для урахування уподобань.

```

public function testGetCategoryAttribute()
{
    $this->helperConfig->expects($this->any())
    ->method('getShopCategoryAttribute')
    ->willReturn('test');

    /** @var \ReflectionClass $class */
    $class = new ReflectionClass($this->model);

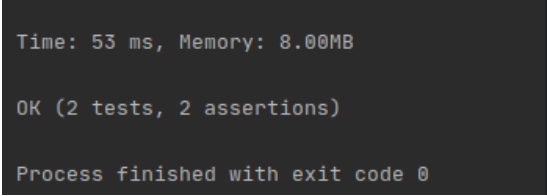
    $method = $class->getMethod('getCategoryAttribute');
    $method->setAccessible('true');

    $this->assertEquals('test', $method->invoke($this->model));
}

```

Рисунок 4.13 – Код тесту для перевірки правильності отримання атрибутів

Результат проведення перевірки модуля наведений на рис. 4.14.



```

Time: 53 ms, Memory: 8.00MB

OK (2 tests, 2 assertions)

Process finished with exit code 0

```

Рисунок 4.14 – Результат проведених тестів

Отже, проведені модульні тести успішно пройшли перевірку, тому, можна зробити висновок, що функції системи електронної комерції працюють належним чином.

4.8 Аналіз ефективності функціонування системи електронної комерції з інтегрованим алгоритмом

Для проведення аналізу ефективності функціонування інтегрованого модулю алгоритма колаборативної фільтрації заснованої на сусідстві (user-based) потребується виділення критеріїв ефективності.

Для їхньої оцінки зазвичай застосовують явний або неявний види зворотнього зв'язку. Під явним розуміють виставлення оцінок системі користувачами в будь-якому вигляді: заповнення форм зворотнього зв'язку, отримання відповідей на поставлені запитання в телефонному режимі (опитування), заповнення форми оцінки роботи систем електронної комерції на сторінках веб-сайту користувачами тощо.

Неявним називають вид, який дозволяє аналізувати дії та поведінку користувача, юзаючи на даних, які збираються та зберігаються в базі даних системи. Окрім цього, підхід заснований на відборі користувачів за схожістю [29].

Тож в якості критеріїв оцінки ефективності алгоритму обрано:

- зростання кількості продажів;
- точність роботи алгоритму;
- явний збір оцінок користувачів.

Рекомендації схожих товарів користувачам можна вважати одним із шляхів стимуляції покупок. Тож аналіз зміни кількості продажів дозволить оцінити ефективність інтеграції модулю алгоритма колаборативної фільтрації.

Шляхом залучення користувачів, протягом 6 місяців проводилося дослідження. В результаті проведення порівняння кількості продажів в системі електронної комерції, здійснених залученими користувачами, що відбулися в період з лютого 2021 року по листопад 2021 року, отримано графік, який ілюструє зростання сумарної кількості продажів щомісяця (рис. 4.15).

Інтеграція модуля проведена в жовтні 2021 року. З отриманих результатів спостереження зрозуміло, що впровадження алгоритму колаборативної фільтрації

дозволило перевищити максимальну кількість продажів, яка склала 76 у червні до його інтеграції, а в листопаді, після інтеграції 116.

Тож можна припустити, що подальше використання алгоритму, а також його оптимізація, дозволить підвищувати кількість продажів постійно.



Рисунок 4.15 Графік зміни кількості продажів товарів

Окрім кількості продажів, зростає кількість зареєстрованих користувачів (табл. 4.1). Адже лише зареєстровані користувачі можуть оцінювати товари та отримувати більш точні рекомендації.

Таблиця 4.1 Порівняння кількості зареєстрованих користувачів в системі електронної комерції щомісяця

Дата	Кількість зареєстрованих користувачів
Травень 2021	12
Червень 2021	16
Липень 2021	16
Серпень 2021	23
Вересень 2021	24
Жовтень 2021	35
Листопад 2021	49

Порівняно з травнем, кількість зареєстрованих користувачів зростає в 4.1 рази в листопаді.

Також, в результаті проведеного дослідження отримана розріджена матриця з наступною структурою:

Таблиця 4.2 Розріджена матриця

	Плаття «Альфа»	Плаття «Літо»	Плаття «Колібрі»	Плаття «Мрія»	Блуза «Клео»	Костюм бавовняний	Спідниця «Єва»
Юлія	4	5	3.1184	1	3.2852	2.8	3.2277
Ганна		2		4		5	4
Олена	5		3	2		1	
Марта		3			5	4	

Для оцінки ефективності роботи алгоритму першим кроком обчислена подібність користувачів за формулою 4.1 для алгоритма колаборативної фільтрації заснованого на сусідстві (user-based).

$$\widehat{r}_{uo} = \bar{r}_u + \frac{1}{\sum_{u' \in R(u)} |sim(u, u')|} \sum_{u' \in R(u)} sim(u, u') (r_{u',i} - \bar{r}_{u'}), \quad (4.1)$$

де \bar{r}_u середня оцінка користувача Юлія, $\sum_{u' \in R(u)} sim(u, u')$ сума схожості за користувачами, $\bar{r}_u = \frac{1}{|R(u)|} \sum_{o \in R(u)} r_{ui}$ значення вподобань за користувачами, $sim(u, u')$ міра схожості.

$$\sum_{u' \in R(u)} sim(u, u') = 1.1475$$

Результати обчислень суми схожості користувача Юлія та користувачів Ганна, Олена та Марта представлені в таблиці 4.3.

Таблиця 4.3 Результати обчислень суми схожості за користувачами

	Ганна	Олена	Марта	Сума
Юлія	0.2766	0.5436	0.3273	1.1475

Перевірка оцінки за товаром «Костюм двійка», обчислення коефіцієнта схожості з кожним користувачем (Ганна, Олена, Марта).

$$\sum_{w \in R(u)} \text{sim}(u, u')(r_{u',i} - \bar{r}_{u'}) = 0.27 * (5 - 3.75) = 0.34,$$

$$\sum_{w \in R(u)} \text{sim}(u, u')(r_{u',i} - \bar{r}_{u'}) = 0.54 * (1 - 2.75) = -0.945,$$

$$\sum_{w \in R(u)} \text{sim}(u, u')(r_{u',i} - \bar{r}_{u'}) = 0.32 * (4 - 4) = 0,$$

Попередньо обчислена середня оцінка товарів користувачами, яка дорівнює $\bar{r}_u = 3.33$.

$$\widehat{r}_{uo} = 3.33 + \frac{0.34 - 0.945 + 0}{1.1475} = 2.8$$

Тож для товару «Костюм двійка» отримана передбачувана оцінка користувача, яка дорівнює 2.8.

За аналогією обчислені передбачувані оцінки для інших товарів.

Також проведено оцінку ефективності роботи алгоритму, використовуючи метрику root mean square error (RMSE) за формулою (3.6).

$$RMSE = \sqrt{\frac{((3.1 - 3)^2 + (3.3 - 3)^2 + (2.4 - 2)^2 + (3.3 - 3)^2)}{4}} = \sqrt{\frac{0.03}{4}} = 0.087$$

Отже, середньоквадратична помилка дорівнює 8.7%, яка означає, що алгоритм є точним та працює ефективно на 91.3%. Цей показник є досить великим, враховуючи наявність деяких недоліків методу, які розглянуті в 2 розділі.

Для оцінки якості рекомендацій проведено опитування серед користувачів, шляхом розсилки електронного листа з переліком питань та варіантами відповідей.

Таблиця 4.4 Результати явного опитування серед користувачів

Ім'я користувача	Якість рекомендацій	Зручність сервісу	Середня оцінка
Юлія	5	5	5
Ганна	4	4	4
Олена	4	5	4.5
Марта	4	4	4
Карина	5	3	4
Олександра	5	5	5
Ольга	5	5	5
Валерія	5	4	4.5
Богдана	5	5	5
Діана	4	5	4.5
Загальний результат:			4.55

Отже, якість рекомендацій в досить невеликій та новій системі електронної комерції має високий рівень. Крім того, веб-система є досить зручною, проте може бути покращена в майбутньому.

Як покращення для системи можна розглядати дослідження та інтеграцію гібридного алгоритму фільтрації, а також розширення функціонала, шляхом збору та аналізу поведінки користувачів в системі (наприклад, аналіз товарів, які додані до кошика, але не куплені).

Інтегрований алгоритм є досить ефективним, конкурентноспроможним, досить дешевим в розробці та може бути застосований в досить нових системах, що не мають великого обсягу даних, який потрібно обробляти.

ВИСНОВКИ

В результаті виконання кваліфікаційної роботи, згідно поставленої задачі, проведено аналіз обраної предметної області, розглянуті та досліджені функції, призначення та типи систем електронної комерції. Проаналізовані способи врахування уподобань користувачів в системах електронної комерції. Проведений порівняльний аналіз алгоритмів рекомендаційних систем.

Визначено, що застосування рекомендаційних систем є популярною технікою та стратегією, яка допомагає користувачам обирати бажані товари. За останні кілька років рекомендаційні системи змінилися від новинок, які використовуються кількома великими сайтами електронної комерції, до серйозних бізнес-інструментів, які змінюють світ електронної комерції.

Під час теоретичних досліджень розглянуті різні алгоритми, які можуть бути використані для розробки систем рекомендацій для бізнесу, основним профілем якого є продаж товарів, зокрема для підбору найбільш релевантних варіантів товарів персонально для кожного клієнта. Крім того, наведено приклад інтеграції та використання одного з алгоритмів, алгоритма колаборативної фільтрації, для інтернет-магазину одягу, що надає можливість оцінки товарів та перегляду рекомендацій в подальшому, орієнтуючись на одиниці товарів, які були оцінені раніше користувачами, та, використовуючи аналіз оцінок інших користувачів, що, зокрема, є близькими за інтересами до клієнта, для якого відбувається підбір рекомендацій.

Проведено тестування розроблених модулів системи електронної комерції та алгоритму. Проведений аналіз ефективності роботи розробленого алгоритму в системі електронної комерції, за результатами якого визначено, що алгоритм працює правильно та ефективно.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ

1. Формування сучасної науки: методика та практика: матеріали I Міжнародної студентської наукової конференції (Т.2), м.Кам'янець-Подільський, 29 жовтня, 2021 рік / ГО«Молодіжна наукова ліга».—Вінниця: ГО«Європейська наукова платформа», 2021. 128с. с. 21.
2. Ralph M. Stair, George W. Reynolds. Principles of information systems: a managerial approach, Sixth Edition. // Transaction Processing and Management Reporting Systems. Chapter 9. – 2001. – [Електронний ресурс] – Режим доступу до ресурсу: <https://www.umsl.edu/~joshik/msis480/chapt09.htm>.
3. Балабанов И.П. Электронная Коммерция / И.П. Балабанов. СПб:Питер, 2001. 336 с.
4. Закон України № 675. Про електронну комерцію [Електронний ресурс] // Верховна Рада України. – 2011. – Режим доступу до ресурсу: <https://zakon.rada.gov.ua/laws/show/675-19>.
5. Джонс, М. Рекомендаційні системи: Частина 1. Введення в підходи і алгоритми.: <http://www.ibm.com/developerworks/ru/library/os-recommender1.html> (дата звернення: 01.04.2021)
6. The Amazon Recommendations Secret to Selling More Online [Електронний ресурс] // Rejoiner. – Режим доступу до ресурсу: <https://rejoiner.com/resources/amazon-recommendations-secret-selling-online/>
7. Vivek Dubey. How Amazon, eBay, Netflix, and Trivago Recommendations System Works? [Електронний ресурс] // LinkedIn. – Режим доступу до ресурсу: <https://www.linkedin.com/pulse/how-amazon-ebay-netflix-trivago-recommendations-system-vivek-dubey>
8. Ким Фальк. Рекомендательные системы на практике / пер. с англ. Д. М. Павлова. – М.: ДМК Пресс, 2020. – 448 с.: ил.
9. S. A. Amelkin. Evaluation of Recommender Systems Efficiency. EUR-WS v. 934. XIV All-Russian Scientific Conference "Digital libraries: Advanced Methods and

Technologies, Digital Collections" Pereslavl-Zalessky, Russia, October 15-18, 2012. pp, 288-291.

10. К. Косолапов. Введение в рекомендательные системы. [Электронный ресурс] // habr.ru. – 2019. – Режим доступа до ресурсу: <https://habr.com/ru/post/476222/>

11. R.Manjula. Content Based Filtering Techniques in Recommendation System using user preferences. International Journal of Innovations in Engineering and Technology (IJJET), 2016, pp. 149-154.

12. Sitnikov, D., Titova, O., Minukhin, S., Kovalenko, A., Titov, S. Informativity of Association Rules from the Viewpoint of Information Theory. Conference: 2018 IEEE International Scientific-Practical Conference Problems of Infocommunications. Science and TechnologyAt: Kharkiv, Ukraine.

13. Burke R. Hybrid Web Recommender Systems // The Adaptive Web. Lecture Notes in Computer Science, Vol. 4321. Springer, Berlin, Heidelberg. – 2007. – С. 377-408.

14. Balabanovic, M. and Shoham, Y., Fab: Content-based, Collaborative Recommendation. Communications of the ACM, 40(3), 1997, pp. 66–72.

15. Linden G., Smith B., and York J. Item-to-Item Collaborative Filtering (англ.) // IEEE Internet Computing, Los Alamitos, CA USA: журнал. — 2003. — P. 76-80.

16. Stekh Y., Logvinenko A., Lobur M., Artsibasov V. Model and methods for building Web recommendation systems // Proc. of the VIth International Conference on Computer Science and Information Technologies (CSIT 2011). – Lviv, 2011. – P. 314-316.

17. Мелешко Є. В. Дослідження методів побудови рекомендаційних систем в мережі Інтернет / Є. В. Мелешко, С. Г. Семенов, В. Д. Хох // Системи управління, навігації та зв'язку. - 2018. - Вип. 1. С. 131-136.

18. R. N. Behera and S. Dash. A particle swarm optimization based hybrid recommendation system. International Journal of Knowledge Discovery in Bioinformatics, vol. 6, no. 2, pp. 1–10, 2016.

19. Bernardi L., Kamps J., Kiseleva J., Mueller M.J.I. The Continuous Cold Start Problem in e-Commerce Recommender Systems. – 2015. – 6 p. – [Электронный ресурс] – Access mode: <https://arxiv.org/abs/1508.01177>

20. X. Su and T. M. Khoshgoftaar. A survey of collaborative filtering techniques. Adv. In Artif. Intell., 2009.

21. Fleder D., Hosanagar K. Blockbuster Culture's Next Rise or Fall: The Impact of Recommender Systems on Sales Diversity // Management Science, Vol. 55, No. 5, May 2009, pp. 697-712: журнал. — 2009. — P. 1 - 49.

22. Isinkaye F. O. Recommendation systems: Principles, methods and evaluation / F. O. Isinkaye, Y. O. Folajimi, B. A. Ojokoh. // Egyptian Informatics Journal. – 2015. – №16. – С. 261–273.

23. Shuyu Luo. Introduction to Recommender System. [Электронный ресурс] // Режим доступа до ресурсу: <https://towardsdatascience.com/intro-to-recommender-system-collaborative-filtering-64a238194a26>

24. Stylianos (Stelios) Kampakis. Performance measures: RMSE and MAE. [Электронный ресурс] // Режим доступа до ресурсу: <https://thedata scientist.com/performance-measures-rmse-mae/>

25. Ольга Назина. Клиент-серверная архитектура в картинках. [Электронный ресурс] // Режим доступа до ресурсу: <https://software-testing.ru/library/testing/testing-for-beginners/3340-client-server-architecture>

26. Буч Г., Якобсон А., Рамбо Дж. UML. Классика CS. 2-е изд. / Пер. с англ.; Под общей редакцией проф. С. Орлова. – СПб.: Питер, 2006. – 736 с.

27. John Deacon. Model-View-Controller (MVC) architecture [Электронный ресурс] // Режим доступа до ресурсу: <https://www.rareparts.com/pdf/MVC.pdf>

28. MySQL. / Дюбуа П. : Пер. с англ. – М.: ООО «И. Д. Вильямс», 2007. – 1168 с.

29. Francesco Ricci, Lior Rokach, Bracha Shapira, Paul B. Kantor. Recommender Systems Handbook. Springer Science+Business Media, LLC, 2011, 845 с.