Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет комп'ютерних наук
(повна назва)

Кафедра програмної інженерії
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти другий (магістерський)

Дослідження технологій рекомендацій для систем електронної комерції

Виконав:
студент (ка) 2 курсу, групи ІПЗм-22-4

Демченко М. О.

Спеціальність 121 – Інженерія програмного забезпечення
(код і повна назва спеціальності)
Тип програми освітньо-наукова

Керівник доц. каф. ПІ Каук В. І.
(посада, прізвище, ініціали)

Допускається до захисту
Зав. кафедри _____ З.В.Дудар
(підпис) (прізвище, ініціали)

2024 р.

Харківський національний університет радіоелектроніки

Факультет комп'ютерних наук
Кафедра програмної інженерії
Рівень вищої освіти другий (магістерський)
Спеціальність 121 – Інженерія програмного забезпечення
Тип програми освітньо-наукова програма
Освітня програма Інженерія програмного забезпечення
(шифр і назва)

ЗАТВЕРДЖУЮ:
Зав. кафедри _____
(підпис)
«___» _____ 2024 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові Демченку Максиму Олександровичу

1. Тема роботи «Дослідження технологій рекомендацій для систем електронної комерції»

Затверджена наказом по університету від 20.03.2024 р. № 250Ст

2. Термін подання студентом роботи до екзаменаційної комісії 19.06.2024

3. Вихідні дані до роботи Дослідження технологій рекомендацій для е-комерції. Програмні засоби, що використовуються: Операційна система Microsoft Windows 10, середовище розробки VS Code 2024, мова програмування JavaScript, система управління базами даних MongoDB.

4. Перелік питань, що потрібно опрацювати в роботі мета роботи, аналіз предметної галузі і постановка задачі, огляд та аналіз літературних джерел з дослідження, дослідження теоретичне, дослідження практичне.

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Аналіз предметної галузі	01.05.2024 – 03.05.2024	*Виконано*
2	Розробка постановки задачі	04.05.2024 – 06.05.2024	*Виконано*
3	Розробка ПЗ	07.05.2024 – 10.05.2024	*Виконано*
4	Проведення експерименту	13.05.2024 – 14.05.2024	*Виконано*
5	Оформлення пояснювальної записки	15.05.2024 – 27.05.2024	*Виконано*
6	Підготовка презентації та доповіді	01.06.2024 – 02.06.2024	*Виконано*
7	Попередній захист	05.06.2024	*Виконано*
8	Нормоконтроль, рецензування	14.06.2024	*Виконано*
9	Здача роботи у електрониий архів	16.06.2024	*Виконано*
10	Допуск до захисту у зав. кафедри	18.06.2024	*Виконано*

Дата видачі завдання 27.04.2024 р.

Студент _____ Демченко М. О.
(підпис)

Керівник роботи _____ доц. каф. ПІ, Каук В. І.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ / ABSTRACT

Робота містить: 54 сторінки, 35 рисунків, 2 таблиці, 4 додатки, 20 джерел.

ВЛАСТИВОСТІ ОЦІНЮВАННЯ, КОЛАБОРАТИВНА ФІЛЬТРАЦІЯ, ПРОЦЕС ОЦІНЮВАННЯ, РЕКОМЕНДАЦІЙНІ СИСТЕМИ.

Робота присвячена дослідженню технологій рекомендацій для електронної комерції (е-комерції). Розвиток інтернет-торгівлі та зростання обсягів онлайн-продажів ставлять перед підприємствами завдання забезпечення ефективності та персоналізації користувацького досвіду. В цьому контексті рекомендаційні системи виявляються ключовим інструментом для досягнення цих цілей.

Об'єктом дослідження є різноманітні технології та методи, використовувані в рекомендаційних системах для покращення процесів в електронній комерції. Розглядаються як традиційні методи рекомендацій, так і новітні підходи, що базуються на штучному інтелекті, машинному навчанні та аналізі великих даних.

У роботі детально аналізуються принципи функціонування та переваги різних рекомендаційних систем, таких як колаборативна фільтрація, контент-фільтрація, гібридні методи та інші. Здійснюється порівняльний аналіз ефективності цих систем у контексті електронної комерції.

Отримані результати та висновки роботи визначають перспективні напрямки вдосконалення технологій рекомендацій для подальшого використання в сучасній електронній комерції. Робота має практичне значення для бізнесу та науково-дослідницької галузі, сприяючи розумінню та впровадженню передових рекомендаційних технологій в сфері е-комерції.

COLLABORATIVE FILTERING, EVALUATION PROCESS, PROPERTIES OF EVALUATION, RECOMMENDATION SYSTEMS

The work paper is dedicated to exploring recommendation technologies for electronic commerce (e-commerce). The development of internet trade and the increase

in online sales volumes pose challenges for businesses to ensure efficiency and personalize user experiences. In this context, recommendation systems emerge as a key tool to achieve these goals.

The object of the research is various technologies and methods used in recommendation systems to enhance processes in electronic commerce. Traditional recommendation methods as well as innovative approaches based on artificial intelligence, machine learning, and big data analysis are considered.

The paper provides a detailed analysis of the functioning principles and advantages of various recommendation systems, such as collaborative filtering, content filtering, hybrid methods, and others. A comparative analysis of the effectiveness of these systems in the context of electronic commerce is conducted.

The obtained results and conclusions of the research identify promising directions for improving recommendation technologies for further application in modern electronic commerce. The paper has practical significance for businesses and the research community, contributing to the understanding and implementation of advanced recommendation technologies in e-commerce.

Я, Демченко Максим Олександрович студент групи ІПЗм-22-4, здобувач освіти на другому (магістерському) рівні кафедра Програмної інженерії, заявляю: моя кваліфікаційна робота на тему «Дослідження технологій рекомендацій для систем електронної комерції», що буде представлена в екзаменаційну комісію для публічного захисту, виконана самостійно, в ній не містяться елементи плагіату і вона може бути опублікована в електронному архіві відкритого доступу ElArKhNURE. Всі запозичення з друкованих та електронних джерел мають відповідні посилання.

Я ознайомлений з діючим положенням «Про протидію академічному плагіату в ХНУРЕ», згідно з яким виявлення плагіату є підставою для відмови в допуску кваліфікаційної роботи до захисту та застосування дисциплінарних заходів.

ЗМІСТ

Вступ 7
1 Аналіз предметної області 9
1.1 Інформаційні ресурси та системи 9
1.2 Загальний аналіз області 9
1.2 Оцінка існуючих систем 11
1.4 Виявлення проблем, що можуть виникнути 20
1.5 Постановка завдання 21
2 Дослідження методів генерації рекомендацій 23
2.1 Колаборативна фільтрація 23
2.2 Контентна фільтрація 25
2.3 Метод, що розглядається у дослідженні 26
3 Підготування до проведення дослідження 27
3.1 Порівняння методів рекомендацій 27
3.2 Опис роботи методу, що досліджується 30
4 Проведення дослідження 32
Висновки 39
Перелік джерел посилання 40
Додаток А 42
Додаток Б 43
Додаток В 50
Додаток Г 53
Додаток Д 54

ВСТУП

В сучасному світі, де конкуренція в електронній комерції є надзвичайно жорсткою, розуміння та використання передових технологій рекомендацій стає стратегічно важливим завданням для підприємств. Рекомендаційні системи дозволяють створювати індивідуалізовані пропозиції для кожного користувача, що сприяє підвищенню його задоволення від покупок та збільшенню ймовірності повторних відвідувань.

Найбільшою перевагою рекомендаційних систем є здатність аналізувати великий обсяг даних, можливість зберігати інформацію про вподобання та дії користувачів, а в результаті обробки цих даних, робити прогнози майбутніх дій клієнтів. Цей процес надає можливість підприємствам надавати своїм клієнтам ті продукти та послуги, які дійсно необхідні, що в перед усім, є суттєвою причиною збільшення продажів та покращення відношення користувачів.

У цій роботі розглядаються різноманітні аспекти рекомендаційних систем, починаючи від традиційних методів, таких як колаборативна фільтрація, і закінчуючи новітніми тенденціями у використанні та аналізу великих даних [1]. Колаборативна фільтрація – це один з найбільш розповсюджених підходів, що бере за основу аналіз вподобань різних користувачів. Проте, на рівні з цим методом, великої популярності набувають й інші підходи, наприклад така як контентно-орієнтовані рекомендації.

Важливо визначити, як рекомендаційні системи впливають на електронну комерцію, враховуючи такі фактори, як забезпечення безпеки даних, оптимізація процесів масштабування та розв'язання проблем рідкісності даних. Проблема рідкісності даних, або ж проблема холодного старту, виникає тоді, коли інформації про нового користувача або продукт є новим, в результаті чого, недостатньо даних для надання точних рекомендацій [2]. Вирішення цієї проблеми є важливим аспектом ефективності рекомендаційних систем.

У роботі також розглядається потенційні обмеження та стратегії для їх вирішення, щоб висвітлити практичну важливість дослідження для підприємств та академічного середовища. Серед проблем, важливо значення набувають

важкість інтеграції рекомендаційних систем у існуючі платформи, великі затрати на їх підтримку та розробку, а також оновлення алгоритмів.

Отримані результати надають підприємствам підстави для більш ефективного впровадження рекомендаційних технологій у їхню стратегію електронної комерції. Використання даних систем може значно покращити користувацький досвід, роблячи його більш привабливим та простішим для клієнтів. Це, насамперед, проявляється підвищенням рівня задоволеності та лояльності клієнтури, в свою чергу, саме це є ключовими факторами для існування підприємства у сучасному світі, котрий наповнений конкуренцією.

Ця робота прагне сприяти розумінню та використанню передових інструментів в сфері рекомендацій, спрямованих на оптимізацію користувацького досвіду та підвищення конкурентоспроможності у світі онлайн-торгівлі. Вона виділяє необхідність внесення інноваційних рішень у сфері електронної комерції, а саме рекомендаційних систем, для досягнення успішної торгівлі для підприємств різного масштабу та напрямків. Залучення передових технологій, надає доступ до нових горизонтів розвитку в електронній комерції.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Інформаційні ресурси та системи

У сучасний період часу важливість інформації для компаній стала невід'ємним активом, незалежно від їхнього сектору, профілю або розміру. Успіх підприємства в сучасній конкурентній грі значно залежить від його здатності збирати, зберігати та ефективно використовувати інформацію. Зростання важливості інформації для підприємств обумовлене такими факторами, як швидка глобалізація, розвиток інформаційних технологій та зміни у споживчому поведінці. Це також стосується компаній, які діють на ринку електронної комерції. Навіть незважаючи на розмір або сферу діяльності, вони опрацьовують величезний обсяг інформації, що перевищує той, що зазвичай мається в традиційній торгівлі, оскільки вони функціонують на віртуальних ринках. Це викликає зростання інтересу до інформації та її потенційного використання для створення конкурентних переваг в електронній комерції. Зі зручнішим доступом до джерел інформації збільшується різноманітність доступної інформації як за обсягом, так і за якістю. Ці зміни ускладнюють взаємини між різними компаніями електронної комерції, породжуючи проблеми, які можна охарактеризувати як патології інформації. Відтак, виникає потреба у розгляді інформації як активу, а інформаційна логістика починає грати все більш ключову роль у підприємствах електронної комерції.

Поняття "інформаційні системи" включає різноманітні мережі, обладнання та програмне забезпечення, які використовуються людьми та компаніями для аналізу та використання даних. Окрім інтерпретації даних, інформаційні системи також здійснюють збір, обробку та поширення цих даних [3].

1.2 Загальний аналіз області

Електронна комерція функціонує через інтернет. Це означає, що торгівля відбувається в цифровому середовищі, де покупець і продавець взаємодіють через веб-платформи, додатки або інші онлайн-канали. Електронна комерція включає в себе різні види діяльності, такі як онлайн-магазини, інтернет-аукціони, електронні

платіжні системи, цифрові служби доставки товарів і багато іншого. Загалом, e-commerce надає можливість проводити торгівлю без фізичної присутності покупця і продавця в одному і тому ж місці [4].

В цілому, система електронної комерції надає певні можливості, використовуючи інтернет-технології, які надають користувачам наступні можливості, які можна побачити на рисунку 1.1.

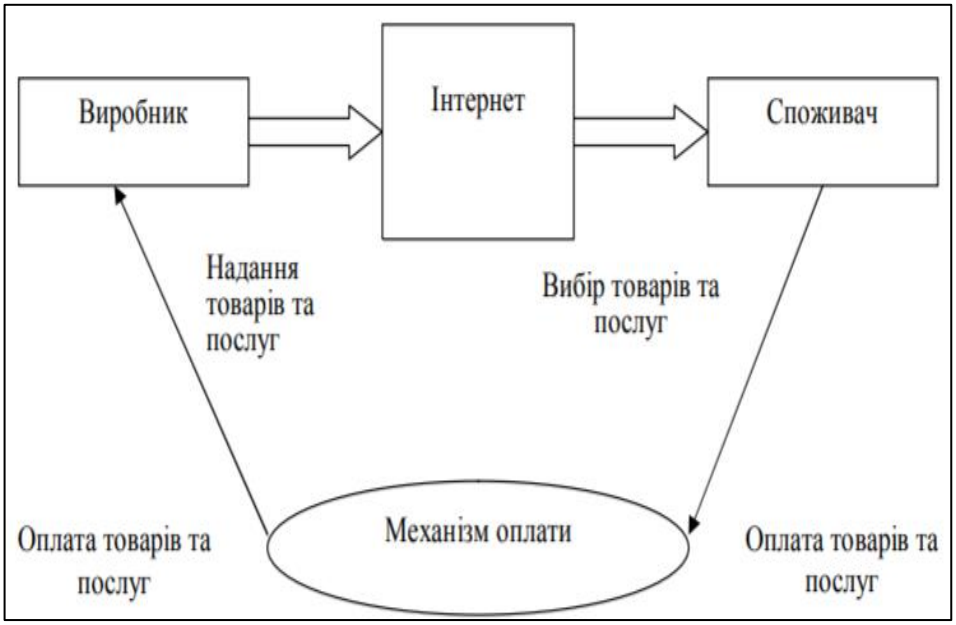

Рисунок 1.1 – Схема праці електронної комерції

У системах електронної комерції може використовується рекомендаційна підсистема, для полегшення етапу вибору товарів користувачами [5].

Рекомендаційні системи є дуже важливим інструментом в електронній комерції, який спрямований на полегшення пошуку необхідних товарів, що в свою чергу підвищує кількість продажів та покращує користувацький досвід, що є покращенням зворотного зв’язку з користувачами.

Для надання рекомендацій користувачам, використовується інформація, яка отримується з різних підсистем електронної комерції, як приклад можна навести підсистему, що відповідає за відстеження дій користувачів з товарами системи. На основі цих даних, можна надати коефіцієнти вподобання для окремих дій користувача та виконати обчислення для подальшого надання рекомендацій для користувачів.

Рекомендаційні системи включаються в себе наступні методи [6]:

- фільтрація на основі вмісту: рекомендації формуються на основі властивостей товарів, які сподобалися користувачу раніше;
- колаборативна фільтрація: рекомендації формуються з включенням вподобань інших користувачів зі схожими інтересами;
- гібридні методи: комбінація кількох підходів.

Рекомендаційні системи допомагають вирішувати наступні проблеми:

- збільшення обсягів продажу;
- підвищення рівня задоволеності користувачів;
- зменшення кількості відмов від товарів (видалення з кошику).

1.2 Оцінка існуючих систем

Існує безліч систем електронної комерції, що відзначаються різною функціональністю, гнучкістю та можливостями. Нижче наведено огляд деяких популярних систем електронної комерції разом з їхніми основними характеристиками:

nike.com – сайт Nike є офіційним інтернет-магазином компанії Nike, одного з провідних світових виробників та дистриб'юторів взуття, одягу та аксесуарів для спорту та повсякденного використання (див.рис.1.2). Розроблений на платформі Magento [7]. Аналіз сайту Nike.

Плюси сайту Nike:

- широкий вибір продукції: велика кількість товарів для різних видів спорту та повсякденного використання;
- інноваційні технології: Nike відома своїми технологічними рішеннями у виробництві взуття та одягу;
- ексклюзивні колекції: представлення унікальних та ексклюзивних моделей товарів.

Мінуси сайту Nike:

- цінова політика: деякі товари можуть бути занадто дорогими, що робить їх недоступними для всіх клієнтів;

– спеціалізація на спортивних товарах: якщо ви шукаєте товари для інших потреб, а не пов'язаних із спортом.

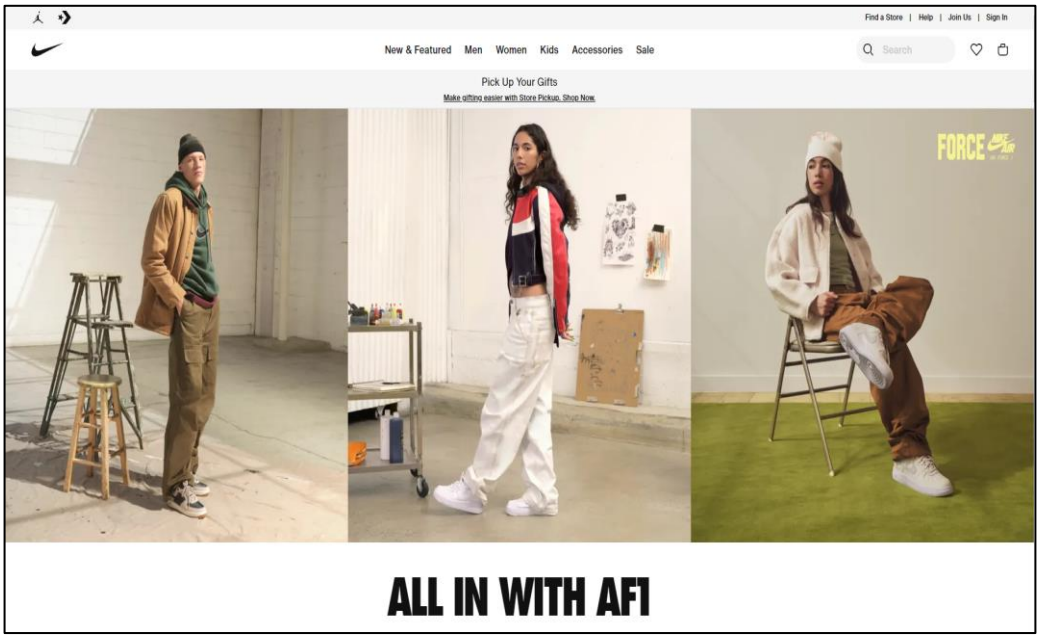

Рисунок 1.2 – Елементи управління сайту Nike

ugmonk.com – це інтернет-магазин (див.рис.1.3), який спеціалізується на продажі стильних і мінімалістичних товарів, таких як одяг, аксесуари та предмети декору. Розроблений на платформі WooCommerce [8]. Аналіз сайту UGMONK.

Плюси сайту UGMONK:

– мінімалістичний стиль: привабливий та модний дизайн, відповідний сучасним трендам мінімалізму;
– унікальний асортимент: ексклюзивні та унікальні товари, які важко знайти в інших магазинах;
– зосередженість на якості: висока якість виробів та увага до деталей.

Мінуси сайту UGMONK:

– цінова політика: товари можуть бути занадто дорогими, що робить їх недоступними для всіх клієнтів.
– обмежений асортимент: у порівнянні з великими магазинами асортимент товарів може бути обмеженим.

UGMONK вражає своїм стильним виглядом, унікальними товарами та фокусом на мінімалізмі, звертаючи увагу на простоту та якість.

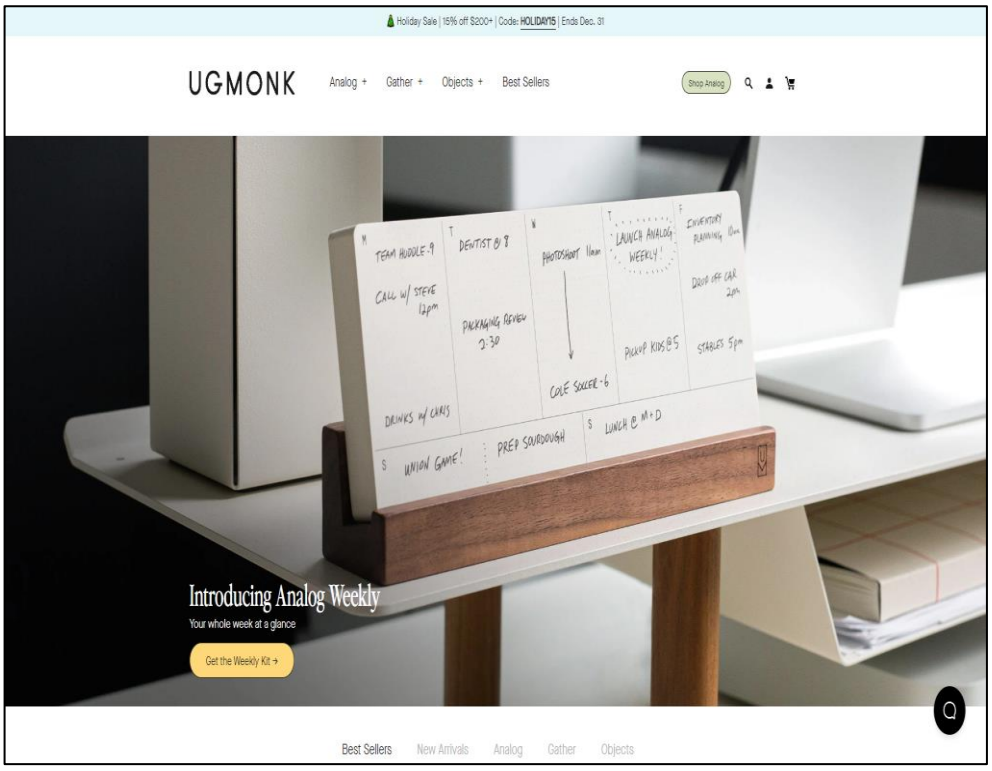

Рисунок 1.3 – Елементи управління сайту UGMONK

allbirds.com – це бренд взуття, який визначається своєю спрощеною філософією, зосередженою на сталевому дизайні та екологічності (див.рис.1.4). Сайт Allbirds відзначається не тільки торгівлею взуттям, але й акцентом на сталевий підхід до дизайну та виробництва. Розроблений на платформі Shopify [9]. Аналіз сайту Allbirds.

Плюси сайту Allbirds:

– екологічна установка: зосередженість на використанні екологічних матеріалів та відповідальному виробництві;
– зручність та мінімалізм: легкий для навігації та мінімалістичний дизайн сайту;
– інновації в дизайні: використання новаторських технологій у виробництві взуття.

Мінуси сайту Allbirds:

– цінова категорія: продукція Allbirds може бути набагато дорожчою порівняно з іншими марками взуття;

– обмежений асортимент: можливо, не надає такого широкого вибору моделей та стилів порівняно з іншими брендами.

Allbirds вирізняється своєю відданістю сталості, використанням екологічних матеріалів та інноваційними підходами до дизайну взуття.

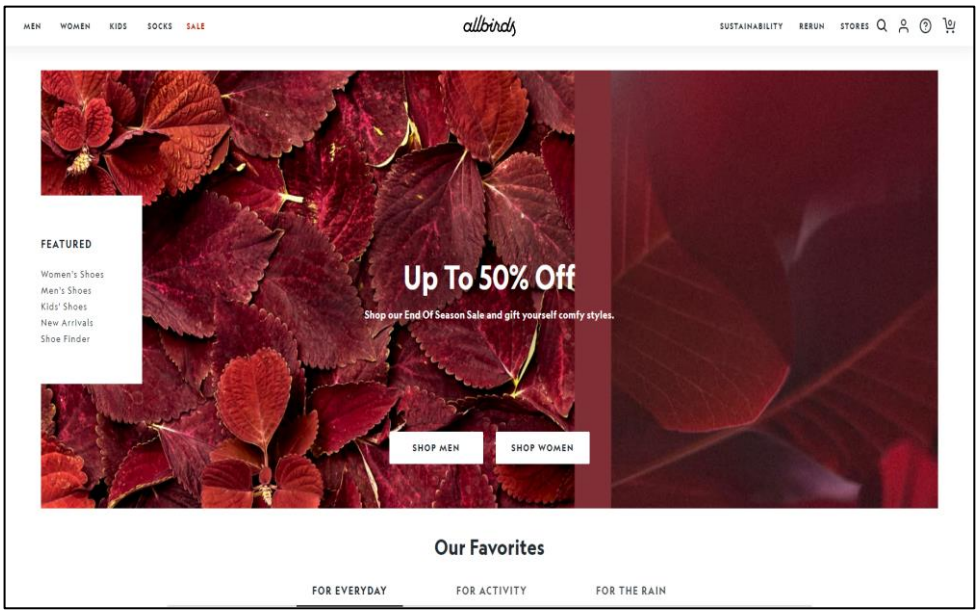

Рисунок 1.4 – Елементи управління сайту Allbirds

valentino.com – офіційний веб-сайт італійського бренду Valentino, відомого своєю високою модою та елегантними виробами в галузі одягу, взуття, сумок та аксесуарів (див.рис.1.5). Розроблений на платформі PrestaShop [10]. Аналіз сайту Valentino.

Плюси сайту Valentino:

– ексклюзивність та Елегантність: вишуканий дизайн та висока якість виробів, що відображають елегантність та стиль бренду;
– велика кількість інформації: ретельна інформація про кожен товар, включаючи фотографії, розміри та матеріали;
– інноваційні підходи: використання віртуального приміщення магазину для унікального онлайн-досвіду.

Мінуси сайту Valentino:

– цінова політика: продукція Valentino зазвичай належить до високої цінової категорії, що робить її недоступною для багатьох клієнтів;

– спеціалізація в високій моді: орієнтований на високу моду, що може не задовольняти всіх клієнтів.

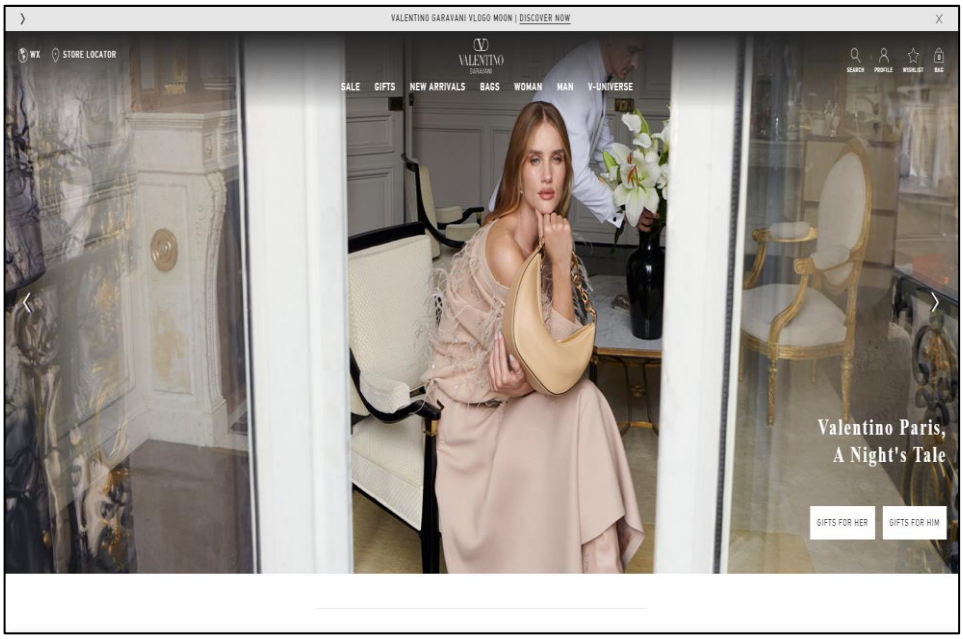

Рисунок 1.5 - Елементи управління сайту Valentino

benjerry.com – це офіційний веб-сайт компанії Ben & Jerry`s, відомої своїм морозивом та замороженими продуктами (див.рис.1.6). Сайт не лише пропонує інформацію про асортимент продукції, але також розповідає про цінності компанії, участь у соціальних питаннях та ініціативах зі сталості. Аналіз сайту Ben & Jerry`s.

Плюси сайту Ben & Jerry`s:

– широкий асортимент смаків: різноманіття унікальних та інноваційних смаків морозива;
– спільна відповідальність: зосередженість на соціальній відповідальності та участі у глобальних ініціативах;
– залучення спільноти: створення спільноти фанатів та активна взаємодія з клієнтами через відгуки та соціальні мережі.

Мінуси сайту Ben & Jerry`s:

– обмеженість продуктів у деяких регіонах: деякі унікальні смаки можуть бути доступні лише в обмежених регіонах;

– можливість несприятливих відгуків: відгуки споживачів можуть бути несприятливими в деяких випадках.

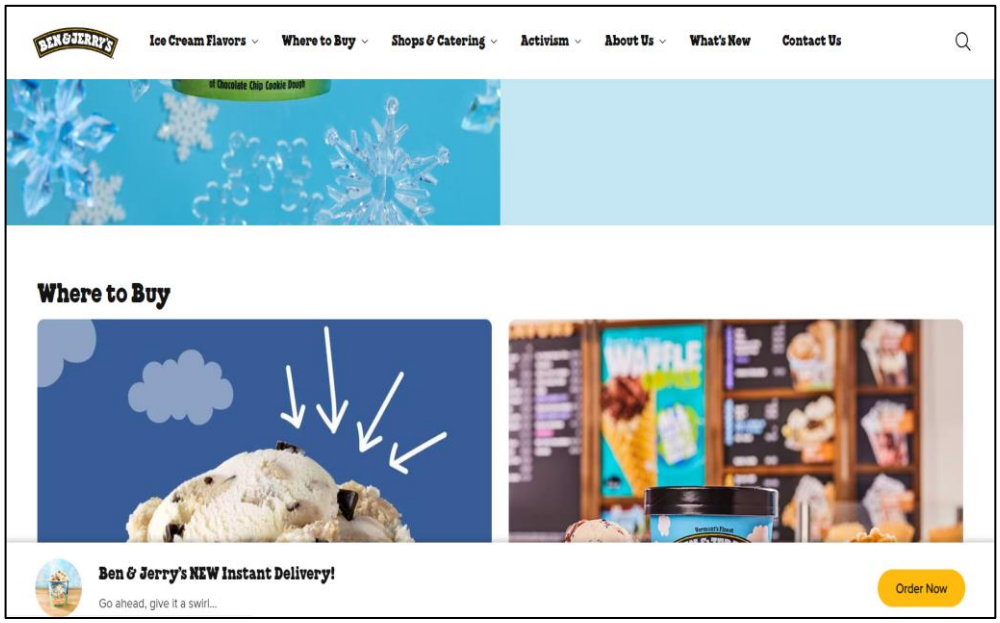

Рисунок 1.6 – Елементи управління сайту Ben & Jerry`s

У розглянутих системах відзначається відсутність можливості отримання рекомендацій для користувачів. У цих платформах відсутні функції, які надають персоналізовані поради або рекомендації в залежності від історії використання або інших параметрів користувача. Така відсутність може вплинути на індивідуальний досвід користувача, оскільки вони не мають можливості отримати персоналізовану інформацію чи рекомендації, які допомагають спростити процес взаємодії з платформою та зробити його більш вдосконаленим.

1.3 Системи автоматичного підбору товарів

Алгоритми рекомендацій представляють собою особливі механізми фільтрації необхідних даних, спрямовані на вибірку та надання користувачам елементів, що можуть викликати найбільшу зацікавленість. Зазвичай подібні системи порівнюють профілі користувачів з певним набором характеристик та прогнозують оцінку, яку особа може призначити конкретному елементу, який ще не враховувався. Основною метою таких систем є відтворення особистих

вподобань користувачів для надання порад щодо покупки товарів чи інших об'єктів.

Ці системи мають велике значення в електронній комерції і доступі до даних, допомагаючи ефективно обробляти обсяги великих інформаційних просторів і спрямовувати користувачів на найбільш підходящі елементи згідно їхніх потреб і вподобань.

За використання методів виявлення знань системи рекомендацій успішно застосовуються у сфері електронної комерції, де з'явлення інтернету зробило їх особливо актуальними. Зростання кількості клієнтів і продуктів ставить перед такими системами три ключові завдання:

- розробка ефективних рекомендацій за вподобаннями користувачів;
- оптимізація алгоритмів для великого набору даних;
- максимально широке охоплення в умовах обмеженості даних.

Найпоширенішим методом створення рекомендацій – це фільтрація, яка враховуючи уподобання користувачів, порівнює їх з іншими користувачами при генерації рекомендацій [11].

При формуванні профілю користувача використовуються різноманітні методи збору даних, які можна класифікувати на явні та неявні. Явний метод передбачає активну участь користувача у процесі, включаючи оцінку елементів за різними критеріями. Наприклад, користувач може визначати свої уподобання та надавати оцінки продуктам чи послугам.

З іншого боку, неявний метод базується на спостереженні за діяльністю користувача без його прямої участі. Це може включати аналіз переглядів товарів в інтернет-магазині, вивчення часу, який користувач витрачає на конкретний товар, а також моніторинг покупок та взаємодії з іншими елементами платформи.

У неявному методі дані збираються під час природної активності користувача, не вимагаючи його активної участі. Цей підхід дозволяє отримати об'єктивну картину вподобань та поведінки користувача, що допомагає у покращенні якості рекомендацій та персоналізації досвіду використання платформи.

Механізм дії рекомендаційної системи (див. Рис 1.7):

– збір даних;
– обробка та аналіз даних;
– створення моделей машинного навчання;
– генерація рекомендацій;
– надання рекомендацій користувачеві.

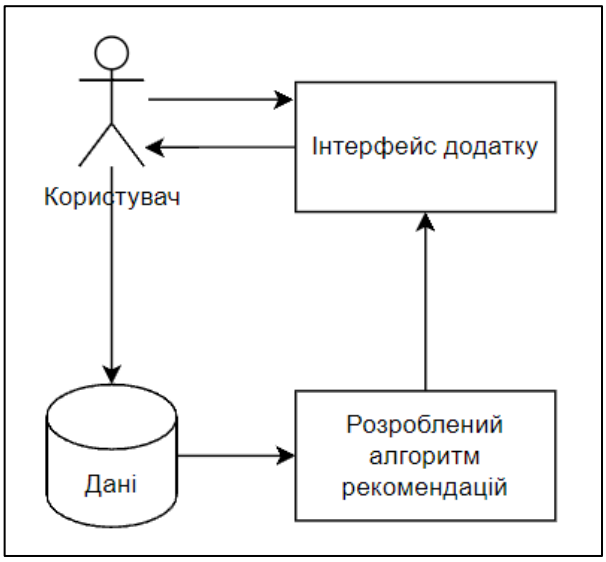

Рисунок 1.7 – Механізм дії рекомендаційної системи

Існує багато платформ та сервісів, які використовують рекомендаційні системи для поліпшення користувацького досвіду та збільшення конверсії. Ось деякі приклади [12]:

– Apple Music: надає персоналізовані рекомендації для музичних треків на основі історії прослуховування та уподобань користувачів(див.рис.1.8);

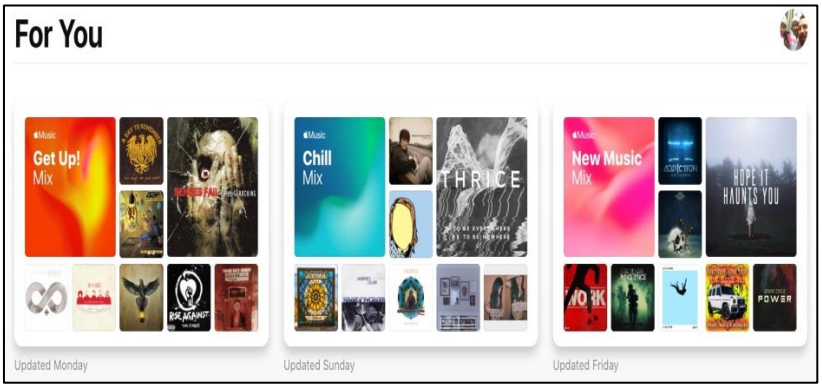

Рисунок 1.8 – Рекомендаційний модуль Apple Music

– Goodreads: рекомендує книги та авторів на основі читацьких вподобань, рецензій та списків прочитаних книг (див.рис.1.9)

Рисунок 1.9 – Рекомендаційний модуль Goodreads

– IMDb: застосовує рекомендації для фільмів та телепередач на основі історії перегляду та оцінок користувачів (див.рис.1.10);

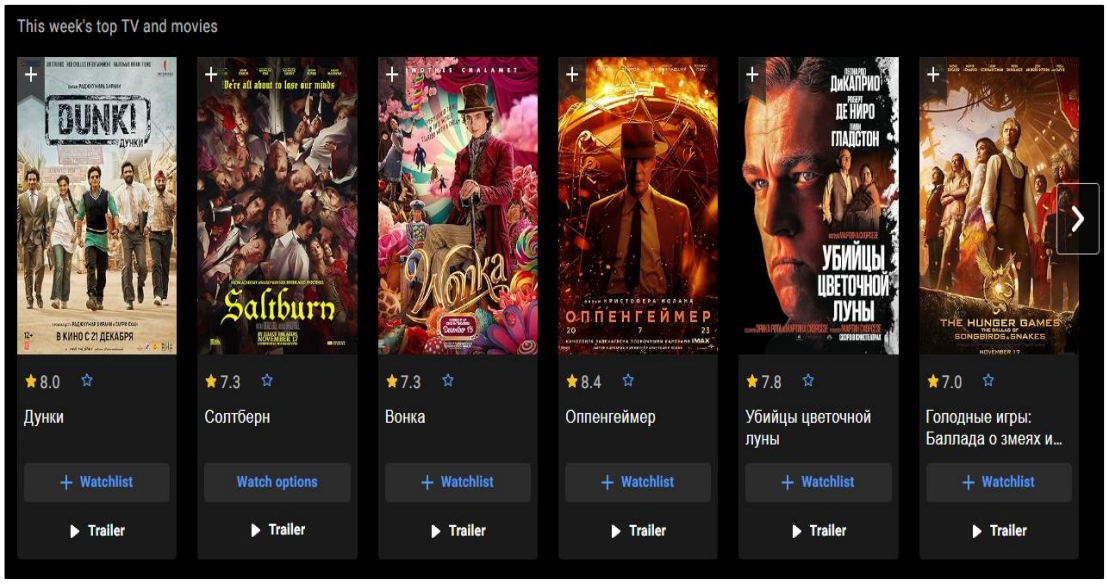

Рисунок 1.10 – Рекомендаційний модуль IMDb

– Reddit: використовує рекомендаційні системи для підбору тематичних обговорень та контенту на основі інтересів користувачів.Last.fm: Рекомендує музичні виконавці та треки на основі слухань користувачів та їхніх музичних уподобань (див.рис.1.11);

Рисунок 1.11 – Рекомендаційний модуль Reddit

– Steam: надає рекомендації для ігор на основі історії геймплею та уподобань гравців (див.рис.1.12);

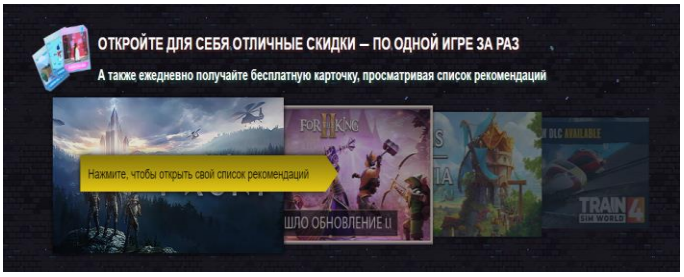

Рисунок 1.12 – Рекомендаційний модуль Steam

Ці приклади вказують на різноманітність застосувань рекомендаційних систем у різних галузях та середовищах.

1.4 Виявлення проблем, що можуть виникнути

Незважаючи на те, що рекомендаційні системи мають великий потенціал для покращення досвіду користувачів та збільшенню обсягів продажів у сфері електронної комерції, але ж таки є деякі проблеми, які мають великий вплив на ефективність та результативність електронної системи [13].

Неповні та шумні дані: однією з найбільших проблем, пов’язаних з рекомендаційною системою, є якість даних, що обробляється. Недостатня кількість або неповнота даних може призвести до формування неточних та неефективних рекомендацій. Це призводить до ускладнення процесу аналізу та формування корисних рекомендацій для користувачів.

Проблема холодного старту: ще одна велика проблема полягає в тому, що недостатньо інформації про нових користувачів та нові продукти. Ця ситуація відома як проблема холодного запуску. Без попередніх даних про поведінку користувачів і характеристики продукту система рекомендацій не може точно передбачити інтереси користувачів і актуальність продукту, що робить процес надання точних рекомендацій надзвичайно складним [14].

Масштабованість: сучасні рекомендаційні системи часто обробляють величезні обсяги даних і потребують великих обчислювальних ресурсів для обробки та аналізу. Забезпечення масштабованості таких систем є складним завданням, оскільки алгоритми повинні швидко і ефективно обробляти великі обсяги даних без шкоди для продуктивності або точності рекомендацій. Це вимагає розробки оптимізованих алгоритмів та використання потужної обчислювальної інфраструктури.

Конфіденційність даних: збір та аналіз даних користувачів є невід'ємною частиною роботи системи рекомендацій, але це ставить питання про захист конфіденційності. Забезпечення конфіденційності та безпеки персональних даних користувачів - дуже важливий аспект, що вимагає пильної уваги. Використання анонімізації, шифрування та інших технологій захисту даних необхідно для підтримки довіри користувачів і дотримання нормативних вимог до конфіденційності.

1.5 Постановка завдання

Об'єктом дослідження в цій роботі є процес створення рекомендацій для користувачів систем електронної комерції.

Предметом дослідження є методи генерування онлайн-рекомендацій у таких системах.

Метою цієї роботи є дослідження методів формування рекомендацій на основі контентної фільтрації та колаборативної фільтрації, які використовуються в процесі створення рекомендацій в електронних системах.

В рамках дослідження було поставлено наступні задачі:

- аналіз структури та функціональних компонентів, які забезпечують роботу рекомендаційної системи, а також вивчення основних факторів, що впливають на її ефективність;
- вивчення існуючих підходів та алгоритмів, які використовуються для створення рекомендацій, зокрема переваг та недоліків кожного з методів;
- детальний аналіз алгоритмів колаборативної фільтрації, які базуються на аналізі поведінки користувачів та їх взаємодії з товарами, а також визначення їхнього впливу на точність рекомендацій;
- дослідження алгоритмів контентної фільтрації, що ґрунтуються на властивостях та характеристиках товарів, а також їх застосування для формування персоналізованих рекомендацій;
- удосконалити метод колаборативної фільтрації з використанням моделей прихованих факторів: розробка покращених алгоритмів, що враховують приховані фактори для підвищення точності рекомендацій та зменшення впливу проблеми холодного старту;
- здійснення тестування та аналіз результатів експериментів для оцінки ефективності запропонованих удосконалень та їх впливу на загальну продуктивність системи рекомендацій.

Виконання цих завдань дозволить не лише теоретично обґрунтувати доцільність застосування різних методів формування рекомендацій, але й на практиці покращити якість та ефективність рекомендаційних систем у сфері електронної комерції.

2 ДОСЛІДЖЕННЯ МЕТОДІВ ГЕНЕРАЦІЇ РЕКОМЕНДАЦІЙ

2.1 Колаборативна фільтрація

Колаборативна фільтрація – основний метод рекомендаційних систем, який часто використовується в електронній комерції. Цей метод базується на аналізі вподобань та відстеження дій користувачів. Основна ідея полягає в тому, що користувачі з однаковими вподобаннями у минулому часі, будуть мати однакові вподобання і в майбутньому. Існує декілька основних підходів до створення методу колаборативної фільтрації: метод на основі пам’яті та метод на основі моделі [15].

Методи, засновані на пам’яті:

Методи, засновані на пам’яті використовують схожість між користувачами або товарами для генерації рекомендацій. Виділяються декілька основних підходів: користувацько-орієнтований та товаро-орієнтована фільтрація.

Користувацько-орієнтована фільтрація: в цьому методі увага приділяється схожості користувачів на основі оцінок, що вони залишають для однакових товарів. Використовується наступна формула для обчислення схожості між користувачами U та V (див. рис. 2.1):

$$\mathrm{sim}(u,v) = \frac{\sum_{i \in I_{uv}} (r_{ui} - \bar{r_u})(r_{vi} - \bar{r_v})}{\sqrt{\sum_{i \in I_{uv}} (r_{ui} - \bar{r_u})^2 \sum_{i \in I_{uv}} (r_{vi} - \bar{r_v})^2}}$$

Рисунок 2.1 – Формула коефіцієнта кореляції Пірсона

В цій формулі:

- sim(u,v) – схожість між користувачами u та v;
- I_{uv} – набір товарів, оцінених обома користувачами;
- r_{ui} – оцінка користувачем u товару i;
- r_u – середня оцінка користувача u.

Для формування рекомендацій необхідно обчислити суму оцінок схожих користувачів (див. рис 2.2):

$$\hat{r_{ui}} = \bar{r_u} + \frac{\sum_{v \in N(u)} \mathrm{sim}(u,v) \cdot (r_{vi} - \bar{r_v})}{\sum_{v \in N(u)} |\mathrm{sim}(u,v)|}$$

Рисунок 2.2 – Формула обчислення суми оцінок схожих користувачів

В цій формулі:

– r_{ui} – передбачувана оцінка користувачем u товару i;

– N(u) – набір k найбільш схожих користувачів до користувача u.

Методи на основі моделі.

Метод, заснований на моделях, використовує машинне навчання для побудови моделі, яка передбачає оцінку користувачем. Метод використовує розкладання матриці оцінок на дві матриці меншого розміру (див.рис.2.3):

$$R \approx P \cdot Q^T$$

Рисунок 2.3 – Формула створення матриці оцінок користувачів

В цій формулі:

– R – матриця оцінок користувачів;

– P – матриця користувачів;

– Q – матриця товарів.

Основною метою є мінімізація функції втрат (див. рис. 2.4):

$$\min_{P,Q} \sum_{(u,i) \in K} (r_{ui} - p_u^T q_i)^2 + \lambda(||p_u||^2 + ||q_i||^2)$$

Рисунок 2.4 – Мінімізація функції втрат

В цій формулі:

– K – набір відомих оцінок;

– λ – параметр регуляризації для запобігання перенавчанню.

Використання матричної метрики дозволяє виділяти приховані фактори, що впливають на вподобання користувачів, та зробити більш точні рекомендації.

Незважаючи на те, що колаборативна фільтрація є дуже гарним інструментом для надання рекомендацій для користувачів, вона має великий недолік: проблема холодного старту та масштабованість.

2.2 Контентна фільтрація

Контентна фільтрація використовує аналіз властивостей продукту та дозволяє згенерувати рекомендації на основі подібності характеристик продуктів та інтересів користувачів [16].

Контентна фільтрація використовує ключові слова, опис, властивості та інші характеристики товарів. Для визначення подібності товарів, використовується косинусна схожість(див.рис.2.5):

$$\mathrm{sim}(i,j) = \frac{\sum_k w_{ik} \cdot w_{jk}}{\sqrt{\sum_k w_{ik}^2} \cdot \sqrt{\sum_k w_{jk}^2}}$$

Рисунок 2.5 – Косинусна схожість

В цій формулі:

- w_{ik} – вага атрибуту k для товару i;
- w_{jk} – вага атрибуту k для товару j.

Профілі користувачів створюються на основі користувацьких уподобань та історії взаємодії користувачів с продуктами. Для кожного користувача розраховується вектори переваг, що відображають важливість різних атрибутів продукту. Він розраховується за наступною формулою (див. рис. 2.6):

$$\vec{p_u} = \frac{1}{|I_u|} \sum_{i \in I_u} \vec{w_i}$$

Рисунок 2.6 – Формула обчислення вектору вподобань

В цій формулі:

– I_u – набір товарів, оцінених користувачем u;

– w_i – вектор атрибутів товару i.

Рекомендації формуються на основі подібності профілю користувача і атрибутів продукту (див.рис.2.7):

$$\boxed{\hat{r_{ui}} = \mathrm{sim}(\vec{p_u}, \vec{w_i})}$$

Рисунок 2.7 – Формула розрахунку передбачуваної оцінки

В цій формулі:

– $\hat{r}_{ui}$ – передбачувана оцінка користувачем u товару i;

– $\mathrm{sim}(p_u, w_i)$ – схожість між профілем користувача та атрибутами товару.

Переваги контентної фільтрації:

– відстуність проблеми холодного старту;

– можливість сформувати рекомендації для нових товарів.

Недоліки контентної фільтрації:

– залежить від якості наповненості характеристик товарів;

– немає можливості надати рекомендації для товарів, які несхожі на старі інтереси користувачів.

2.3 Метод, що розглядається у дослідженні

Результатом роботи повинен стати метод надання рекомендацій, на основі колаборативного методу з використанням контентної фільтрації для вирішення деяких проблем колаборативного методу.

Якщо даних для використання колаборативного методу буде недостатньо, буде використовуватись метод контентної фільтрації, в іншому випадку рекомендації згенеруються з трошки видозміненим колаборативним методом. Основна різниця методу – використання негативних факторів

3 ПІДГОТУВАННЯ ДО ПРОВЕДЕННЯ ДОСЛІДЖЕННЯ

3.1 Порівняння методів рекомендацій

Для порівняння було обрано два методи створення рекомендацій для електронної комерції [17]:

– контент-фільтр;
– колаборативна фільтрація.

Обидва методи мають свій набір недоліків та переваг, зважаючи на напрямок використання їх системою. Завдання полягає у порівнянні ефективності використання цих методів з удосконаленим методом, що розглядається в умовах дослідницької роботи. Необхідно розглянути важливі нюанси методів та занести загальний опис до таблиці порівняння.

Контент-фільтрація – це метод рекомендаційних систем, що базується на аналізі характеристик контенту об’єкту та зіставленні їх з вподобаннями користувачів.

Наприклад, онлайн магазини, такі як eBay або Amazon, використовують контент-фільтрацію для рекомендацій товарів користувачам [18]. Система аналізує характеристики товарів, описи, категорії, бренди тощо, а потім пропонує схожі товари або додаткові товари, які можуть зацікавити користувача.

Перевагою такого підходу можна визначити високу ефективність виконання, що значно впливає на якість системи рекомендацій, через значне зменшення часу відгуку сторінки сайту, що відповідає нормам оптимізації.

Очевидно, як і будь-якого іншого методу, контентна фільтрація має свої недоліки. Серед значних, можна відзначити, що точність рекомендацій сильно залежить від кількості заповненої інформації про об’єкти. Припустимо, існує інтернет магазин, де система рекомендацій використовує контент-фільтрацію для пропозицій покупок користувачам. Однак, якщо для всіх товарів у цьому магазині надано лише інформацію про виробника, а не про характеристики або категорії, система може давати рекомендації, які не враховують справжній інтерес покупця. Наприклад, якщо покупець переглядає категорію "меблі для ванної кімнати" і всі

товари в цій категорії відносяться до одного виробника, то система може неправильно припустити, що покупець зацікавлений тільки в цьому бренді. Це може призвести до невдалих рекомендацій, наприклад, пропозицій придбати додаткові товари того ж виробника, які покупець насправді не бажає. Таким чином, точність системи рекомендацій значно залежить від якості та повноти даних про товари в магазині.

Колаборативна фільтрація – це метод рекомендаційних систем, який базується на аналізі зв'язків між користувачами та об'єктами для надання рекомендацій.

Онлайн магазини, такі як Amazon або eBay, використовують колаборативну фільтрацію для рекомендацій товарів на основі покупок та переглядів інших користувачів з подібними інтересами.

Найкраще колаборативна фільтрація працює в умовах, коли профіль користувача вже сформовано, і з використанням історії взаємодії користувача з товарами сайту, передбачити необхідні для нього пропозиції. Ще однією функцією методу, є знаходження нових вподобань, надаючи набір рекомендацій на основі подібних на користувача зареєстрованих акаунтів [19].

Метод не обробляє інформацію про об’єкти, а лише використовує дані взаємодії користувачів з об’єктами.

Найбільший недолік колаборативної фільтрації – це так званий «холодний старт» [20]. Продуктивність системи падає, через те, що нові предмети або користувачі представляють велику проблему для рекомендаційних систем. Для нових предметів можна використовувати контентні характеристики, які дозволяють зробити припущення щодо їхніх властивостей та взаємозв'язків з іншими предметами. Це дозволить почати рекомендації для цих предметів, навіть якщо недостатньо інформації про їх взаємодію з користувачами. Таким чином, ми можемо налаштувати, щоб система більше рекомендувала холодні пропозиції і відповідно менше гарячі.

Метод, що розглядається в умовах дослідницької роботи, поєднує в собі обидва методи фільтрації. Але замість стандартного методу колаборативної

фільтрації, що використовує позитивні дані про об’єкти. Метод збирає інформацію про негативні дані. Враховуючи розглянуту систему, а саме електрону комерцію. Можна збирати інформацію стосовно можливих повернень товарів, відмов від заказів, тощо. Враховуючи цю інформацію, можна зробити більш обширну картину вподобань користувачів, чим підвищиться точність стандартного колаборативного методу фільтрації.

Враховуючи це, при «холодному старті», буде вирішено проблему рекомендації товарів користувачам. Оскільки новим користувачам, що не встигли покористуватись системою, буде надано до рекомендації об’єкти, що схожі між собою, без урахування інших даних.

Виходячи с цієї можливості, отримуємо також і мінус системи. А саме – точність надання рекомендацій новим користувачам. Вона буде сильно залежати від контенту, що описує об’єкти.

Також великим недоліком методу є те, що він використовує багато обчислень інформації, що викликає великі технічні затрати.

Таблиця 3.1 – Порівняння методів рекомендацій для електронної комерції

Фільтрація	Сильні сторони	Слабкі сторони
Контентна	1. Не потребує роботи з профілями користувачів 2. Не виникає проблема «Холодного старту», бо не потребує заповнення профілю користувачів.	1. Надає рекомендації лише на основі інтересів користувача.
Колаборативна	1. Може надавати рекомендації товарів, що не розглядаються користувачем, але можуть його заінтересувати.	1. Пропорційно до кількості користувачів, які обирають певний товар, зменшується

Кінець таблиці 3.1

		кількість альтернатив, які рекомендує система. 2. Проблема «Холодного старту».
Метод, що досліджується	1. Може надавати рекомендації товарів, що не розглядаються користувачем, але можуть його заінтересувати. 2. Не потребує роботи з профілями користувачів 3. Не виникає проблема «Холодного старту», бо не потребує заповнення профілю користувачів.	1. Залежність від сильного технічного обладнання. 2. Залежність від правильного заповнення даних про об’єкти.

З таблиці можна побачити: метод, що досліджується, поєднує переваги обох підходів: він може рекомендувати товари на основі інтересів користувача та уникнути проблеми "холодного старту", при цьому не потребує роботи з профілями користувачів. Однак він також залежить від технічного обладнання та правильного заповнення даних про об’єкти."

Цей опис допомагає зрозуміти ключові відмінності між різними методами фільтрації та їхні взаємні переваги та обмеження.

3.2 Опис роботи методу, що досліджується

Метод, що досліджується у роботі, комбінує в собі колаборативну на контекстну фільтрації (див. рис. 3.1).

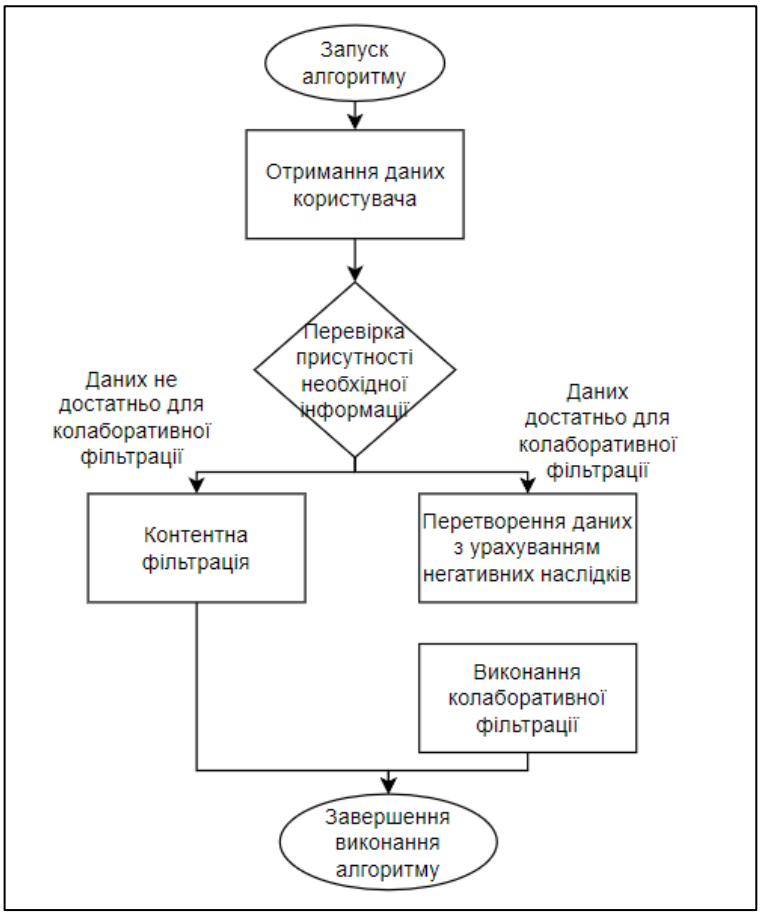

Рисунок 3.1 – Робота методу, що досліджується

Цей метод використовує колаборатвну фільтрацію, але для аналізу даних, використовується не тільки позитивні дії користувачів з товарами, а ще й негативні наслідки, а саме видалення товару з кошику покупок.

У разі, якщо після перетворення інформації про користувача, її недостатньо для виконання процесу колаборативної фільтрації, застосовується контентна фільтрація.

4 ПРОВЕДЕННЯ ДОСЛІДЖЕННЯ

Для проведення дослідження з відкритих джерел завантажено збірку файлів, а саме:

- events.csv – файл з даними поведінки;
- category_tree.csv – файл, що описує дерево категорій;
- item_properties.csv – файл із властивостями предметів.

Файл з даними поведінки, зберігає події, що були зроблені користувачами, а саме:

- view;
- transaction;
- remove;
- addtocart.

У завантаженому наборі даних, є 2 178 329 подій «View», 67 371 «addtocart» та 22 457 «other». Нижче наведено приклад рядків подій (див.рис.4.1):

1433221332117	257597	view	355908

Рисунок 4.1 – Приклад рядку для поведінки «view»

Цей рядок має формат даних, який, використовується для запису певних подій або транзакцій. Ось розшифровка полів цього рядка:

- 1433221332117 – це мітка часу або дата в форматі Unix Timestamp (кількість мілісекунд, що минула з 1 січня 1970 року);
- 257597 – це ідентифікатор користувача або об'єкта, який здійснив транзакцію або взяв участь у події;
- view – це тип події або транзакції. У цьому випадку це перегляд, але в інших випадках може бути іншим типом події, наприклад, "транзакція", "видалення" і т.д.;
- 355908 – це ідентифікатор, що пов'язаний з транзакцією або об'єктом, що стосується цієї транзакції.

Файл опису дерева категорій складається з 1698 рядків. Рядки вказують на зв’язок дочірньої категорії з батьківською. Нижче наведено приклад рядків (див.рис.4.2):

1016	213

Рисунок 4.2 – Приклад рядку для зв’язку дочірньої категорії з батьківською

Цей рядок має формат даних, який, використовується для запису зв’язку дочірньої категорії з батьківською. Ось розшифровка полів цього рядка:

– 1016 – це ідентифікатор категорії;
– 213 – це ідентифікатор батьківської категорії (якщо його немає, поле порожнє).

Файл опису властивостей об’єктів складається з 269 604 рядків опису, для 43 200 об’єктів. Рядки мають поле «timestamp», оскільки з плином часу, властивість може змінюватись.

Для вирішення задачі надання рекомендацій, використаємо коефіцієнти для подій, а саме:

– додання до кошику: 0.5;
– перегляд товару: 0.2;
– транзакція: 1;
– видалення товару з кошику: -1.

Використаємо метод, що перетворить позиції подій у базі даних, на наші коефіцієнти. Код представлено на рисунку 4.3.

```
eventsCollection.find({}).toArray(function(err, allEvents) {
        if (err) throw err;

        const mapping = {
            'view': 0.2,
            'addtocart': 0.5,
            'transaction': 1,
            'remove': -1
        };

        // Обробляємо кожну подію
        allEvents.forEach(function(event) {
            if (!event.timestamp) {
                return;
            }
            var rand = Math.random();
            if (rand < 0.15) {
                event.event = 'remove';
            }

            // Оновлюємо документ у колекції events
            eventsCollection.updateOne(
                { _id: event._id },
                {
                    $set: {
                        ProductCode: event.itemid,
                        Rating: mapping[event.event],
                        CustomerID: event.visitorid
                    }
                }
            );
    });
```

Рисунок 4.3 – Код перетворення подій на коефіцієнти

Перетворюємо файл дерева категорій, для подальшого використання у контекстній рекомендації. А саме перетворюємо об’єкт продукту в JSON рядок. Після чого, перенесемо дані до файлу, що зберігає ідентифікатор об’єкту та його опис. Код представлено на рисунку 4.4.

```
productsCollection.find({}).toArray(function(err, products) {
    products.forEach(function(el) {
        productsMappedCollection.insertOne({
            itemid: el.itemId,
            description: JSON.stringify(el, null, 2)
        }, function(err, result) {
            if (err) throw err;
        });
    });
```

Рисунок 4.4 – Перетворення файлу дерева категорій

Для того, щоб отримати дані про товари, користувачів, що взаємодіяли с цими товарами, було використано метод, що наведено на рисунку 4.5.

```
await client.connect();

while (await eventsCursor.hasNext()) {
    const eventDocument = await eventsCursor.next();
    const eventObject = JSON.parse(JSON.stringify(eventDocument));
    if (eventObject) {
        customers.push(eventObject.CustomerID);
        products.push(eventObject.ProductCode);
        marks.push(eventObject.Rating)
        events.push(eventObject);
    }
}
```

Рисунок 4.5 – Метод отримання даних

Після чого, вони використовуються для колаборативної фільтрації. Звичайна фільтрація, використовує лише ті дані, коефіцієнт котрих більший за 0. Для методу, що розглядається у дослідженні, будуть використовуватись також дані, що менше за 0.

На наступному кроці, необхідно створити матрицю предметів та коефіцієнтів взаємодії з цими предметами. Код стандартної фільтрацій зображено на рисунку 4.6, а код методу, що розглядається на рисунках 4.6 та 4.7.

```
events.filter(event => event.Rating > 0)
    .forEach(event =>{
        eventsData.push({
            customerId: event.CustomerID,
            itemId: event.ProductCode,
            quantity: event.Rating
        })
    })
const dataArray = eventsData.map(({ customerId, itemId, quantity }) => [customerId, itemId, quantity]);
const purchasesSparse = math.matrix(dataArray);
```

Рисунок 4.6 – Код стандартної фільтрацій

```
events.filter(event => event.Rating != 0)
    .forEach(event =>{
        eventsData.push({
            customerId: event.CustomerID,
            itemId: event.ProductCode,
            quantity: event.Rating
        })
    })
const dataArray = eventsData.map(({ customerId, itemId, quantity }) => [customerId, itemId, quantity]);
const purchasesSparse = math.matrix(dataArray);
```

Рисунок 4.7 – Код методу, що розглядається

Як можна побачити, різниця в тому, які дані підпадають виборці. В стандартному методі, лише більші за 0, а в тому, що розглядається у роботі, ще й менші за 0. Результати вибірки можна побачити на рисунках 4.8 та 4.9 відповідно.

```
    ... 2343962 more items
  ],
  _size: [ 2344062, 3 ],
```

Рисунок 4.8 – Результат для стандартного методу

```
    ... 2756001 more items
  ],
  _size: [ 2756101, 3 ],
```

Рисунок 4.9 – Результат для методу, що розглядається

Далі скористаємось функцією створення тестового набору, для тестування точності моделі рекомендацій. Він містить всі оригінальні рейтинги, які згенеровані користувачами для предметів. Частину цих рейтингів "замаскуємо" на нуль, щоб перевірити, наскільки добре модель може передбачити ці "пропущені" рейтинги. Код для перетворення, представлено на рисунку 4.10

```
const eventsData = events.filter(event => event.quantity !== 0)
    .map(event => ({ customerId: event.CustomerID, itemId: event.ProductCode, quantity: event.Rating }));
const jsonDataArray = eventsData.map(event => JSON.stringify(event));

const trainData = jsonDataArray.map(json => JSON.parse(json));
```

Рисунок 4.10 – Код для створення тестового набору

Після цих дій, необхідно перевірити, наскільки добре модель рекомендацій відповідає реальним діям користувачів. Одним зі способів цього аналізу є використання метрики, відомої як "область під кривою ROC" (Receiver Operating Characteristic). Ця метрика вимірює, наскільки добре модель може передбачити позитивні та негативні реакції.

Чим більшою буде площа, що під кривою Receiver Operating Characteristic, тим більше рекомендовані товари, які потім придбаються користувачами, знаходяться вгорі списку рекомендованих товарів. Це дозволяє оцінити, наскільки ефективно модель рекомендацій розміщує товари.

Для розрахунку цієї площі, необхідно створити функцію. Код реалізації функції наведено на рисунку 4.11.

```
const auc = positivePredictions.join(negativePredictions)
    .values()
    .map(t => {
        let correct = 0;
        let total = 0;
        t._1().forEach(positive => {
            t._2().forEach(negative => {
                if (positive.rating() > negative.rating()) {
                    correct++;
                }
                total++;
            });
        });
        if (total === 0) {
            return 0.0;
        }
        return correct / total;
    })
    .mean();
```

Рисунок 4.11 – Код реалізації розрахунку площі

Нижче наведено результати виконання розрахунків для стандартного та методу, що розглядається відповідно. Дивись рисунок 4.12

```
AUC для стандартної фільтрації: 0.791
AUC для методу, що розглядається: 0.839
```

Рисунок 4.12 – Код реалізації розрахунку площі

Як можна побачити з результатів, метод, що розглядається надає більш точні рекомендації, ніж стандартний.

Тепер необхідно реалізувати метод контентної фільтрації. Необхідно розрахувати Term Frequency та Inverse Document Frequency для розпізнання важливості описів об’єктів. Таким чином, чим частіше вживається слово в описах об’єктів, тим воно має більшу вагу. На рисунку 4.13 наведено код реалізації алгоритму.

```
function getTF(docName, term) {
term = term.toLowerCase();
const doc = this.getDoc(docName);
let total = 0.0;

if (doc !== null) {
    if (isPhrase(term)) {
        for (const key in doc) {
            if (key === term) {
                total += doc[key];
            } else if (key.includes(term)) {
                total += doc[key] / 2.0;
            }
        }
    } else if (doc.hasOwnProperty(term)) {
        total = doc[term];
    }
}

return total / Object.keys(doc).length;


function getIDF(term) {
term = term.toLowerCase();
const numberOfDocsWithTerm = this.numberOfDocsWithTerm(term);

if (numberOfDocsWithTerm !== 0) {
    return Math.log(this.corpusBow.length / numberOfDocsWithTerm);
} else {
    return 0;
}
```

Рисунок 4.13 – Код реалізації алгоритму важливості

Порівняємо час формування рекомендацій. Дивись рисунок 4.14.

```
Час, за який створюються рекомендації, стандартний колаборитивний метод (с): 10.951
Час, за який створюються рекомендації, метод, що розглядається (с): 11.25
Час, за який створюються рекомендації, контекстна фільтрація (с): 0.812
```

Рисунок 4.14 – Порівняння часу формування рекомендацій

На основі отриманих даних, створимо таблицю 4.1 порівняння методів:

Таблиця 4.1 Аналіз різних підходів до формування рекомендацій

Фільтрація	AUC	Час
Контентна	-	0.812
Метод, що розглядається	0.839	11.25
Колаборативна	0.791	10.951

Таким чином, можна зробити висновок, що використання методу, що розглядається, є правильним рішенням, через більшу точність наданих товарів користувачам. З іншого боку, контентна фільтрація витрачає значно менше часу на надання рекомендацій. Тому в залежності від ситуації, необхідно обирати один з цих методів.

ВИСНОВКИ

У сучасному світі електронна комерція стає все більш конкурентною, а отже, розуміння та використання передових технологій рекомендацій стає стратегічно важливим завданням для підприємств. Рекомендаційні системи є ключовим інструментом для створення індивідуалізованих пропозицій для кожного користувача, що сприяє підвищенню його задоволення від покупок та збільшенню ймовірності повторних відвідувань.

Однак, впровадження та використання рекомендаційних систем супроводжується рядом викликів, таких як питання безпеки даних, рідкість даних та складність масштабування. Результати цього дослідження вказують на стратегії та рішення для подолання цих проблем, що визначають шляхи для подальшого вдосконалення технологій рекомендацій.

Ця робота досліджує різноманітні аспекти рекомендаційних систем, починаючи від традиційних методів, таких як колаборативна фільтрація, і закінчуючи новітніми тенденціями у використанні та аналізу великих даних. Важливо враховувати вплив рекомендаційних систем на електронну комерцію, зокрема, їхню роль у розв'язанні проблем рідкісності даних.

Отримані результати надають підприємствам підстави для більш ефективного впровадження рекомендаційних технологій у стратегію електронної комерції. Це дозволить покращити користувацький досвід та забезпечити клієнтам персоналізовані та високоякісні пропозиції, що відповідають їхнім унікальним потребам.

Ця робота прагне сприяти розумінню та використанню передових інструментів у сфері рекомендацій, спрямованих на оптимізацію користувацького досвіду та підвищення конкурентоспроможності у світі онлайн-торгівлі. Вона відображає необхідність внесення інноваційних рішень у сфері електронної комерції, а саме рекомендаційних систем, для досягнення успішної торгівлі для підприємств різного масштабу та напрямків.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Кім Фальк. Рекомендаційні системи на практиці. Прес, 2020. 448 с.

2. Короткий опис проблеми холодного запуску. STORYSHOTS. URL: https://www.getstoryshots.com/uk/books/the-cold-start-problem-summary/ (дата звернення 06.06.2024)

3. Інформаційні системи. Step by step. URL: https://step.org.ua/konspekt/inform/tema1 (дата звернення 06.06.2024)

4. E-Commerce: як працює електронна комерція. Voll. URL: https://voll.com.ua/uk/blog/e-commerce-yak-pracyuye-elektronna-komerciya (дата звернення 06.06.2024)

5. Від концепції до реалізації. Dou. URL: https://dou.ua/forums/topic/47504/ (дата звернення 06.06.2024)

6. Як працює система рекомендацій контенту? Mgid. URL: https://www.mgid.com/uk/blog/yak-praczyuye-sistema-rekomendaczij-kontentu (дата звернення 07.06.2024)

7. Що таке Magento? Smile. URL: https://smile-ukraine.com/ua/magento/introduction (дата звернення 07.06.2024)

8. Що таке WooCommerce. URL: https://www.actualidadecommerce.com/uk/WooCommerce/ (дата звернення 07.06.2024)

9. Платформа Shopify:що це таке та як працює. Walnut. URL: https://wallnut.digital/shops/shopify/review/ (дата звернення 07.06.2024)

10.PrestaShop: огляд та можливості платформи. Хабр. URL: https://habr.com/ru/companies/serverspace/articles/745556 (дата звернення 08.06.2024)

11.Як підвищити лояльність клієнтів та покращити комунікацію. Drukarnia. URL: https://drukarnia.com.ua/articles/yak-pidvishiti-loyalnist-kliyentiv-ta-pokrashiti-komunikaciyu-zavdyaki-crm-i-marketingu-trOnY (дата звернення 08.06.2024)

12.Оптимізація користувацького досвіду та конверсії. Ranktracker. URL: https://www.ranktracker.com/uk/blog/user-experience-and-conversion-rate-optimization-key-principles-for-success/ (дата звернення 08.06.2024)

13.Pros and Cons of Collaborative Filtering. Medium. URL: https://medium.com/@ashmi_banerjee/pros-and-cons-of-collaborative-filtering-9c3aa4ce44f6#:~:text=Cold%20start%20and%20Latency%20are,items%20with%20limited%20interaction%20history. (дата звернення 09.06.2024)

14. Лещинський, В.О., Лещинська, І.О. Моделювання вибору користувача в умовах обмежень холодного старту рекомендаційної системи//Математичне моделювання. Распізнавання образів. 2019. с. 14-19

15. User-Based Collaborative Filtering. Geeksforgeeks. URL: https://www.geeksforgeeks.org/user-based-collaborative-filtering/ (дата звернення 09.06.2024)

16. Step-by-Step Guide to Building Content-Based Filtering. Stratascrath. URL: https://www.stratascratch.com/blog/step-by-step-guide-to-building-content-based-filtering/ (дата звернення 09.06.2024)

17. A Guide to Content-Based Filtering In Recommender Systems. Turing. URL: https://www.turing.com/kb/content-based-filtering-in-recommender-systems (дата звернення 09.06.2024)

18. Поради з оптимізації оголошень. Ebay. URL: https://export.ebay.com/ua/listings/how-optimize-your-listings/tips-listing-optimization/ (дата звернення 09.06.2024)

19. Herlocker, J.L., Konstan, J.A., Terveen, L.G., & Riedl, J.T. Evaluating Collaborative Filtering Recommender Systems//Oregon State University, University of Minnesota. 2004. с. 7-53

20. Strategies to Tackle the Cold Start Problem in Recommender Systems. Medium. URL: https://baotramduong.medium.com/recommender-system-the-cold-start-problem-strategies-to-address-it-bddb177e723c (дата звернення 09.06.2024)