

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Навчально-науковий центр заочної форми навчання
(повна назва)

Кафедра Інформаційних управляючих систем
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА Пояснювальна записка

рівень вищої освіти другий (магістерський)

Дослідження методів формування пояснень при зміні вимог користувача в
рекомендаційних системах.

(тема)

Виконав:

студент 2 курсу, групи ІУСТЗМ-22-1

Тимченко Лілія Олександрівна

(прізвище, ім'я, по батькові)

Спеціальність 122 Комп'ютерні науки

(код і повна назва спеціальності)

Тип програми освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Освітня програма Інформаційні управляючі
системи та технології

(повна назва освітньої програми)

Керівник Оксана ЧАЛА

(Власне ім'я ПРІЗВИЩЕ)

Допускається до захисту

Зав. кафедри



(підпис)

Костянтин ПЕТРОВ

(Власне ім'я ПРІЗВИЩЕ)

2024 р.

Харківський національний університет радіоелектроніки

Навчально-науковий центр заочної форми навчання

Кафедра Інформаційних управляючих систем

Рівень вищої освіти другий (магістерський)

Спеціальність 122 Комп'ютерні науки

(код і повна назва)

Тип програми освітньо-професійна

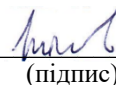
(освітньо-професійна або освітньо-наукова)

Освітня програма Інформаційні управляючі системи та технології

(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри



(підпис)

« 04 » грудня 20 23 р.

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові Тимченко Лілії Олександрівні

(прізвище, ім'я, по батькові)

1. Тема роботи Дослідження методів формування пояснень при зміні вимог користувача в рекомендаційних системах

затверджена наказом університету від 01 грудня 2023 р. № 259Стз

2. Термін подання студентом роботи до екзаменаційної комісії 19 січня 2024 р.

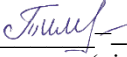
3. Вихідні дані до роботи наукова література, публікації, статті та інтернет джерела з досліджуваної проблеми, матеріали преддипломної практики

4. Перелік питань, що потрібно опрацювати в роботі: аналіз та дослідження методів в рекомендаційних системах, дослідження методів побудови рекомендацій в рекомендаційних системах, аналіз підходів до формування пояснень в рекомендаційних системах, удосконалення методу формування пояснень при зміні вимог користувача за допомогою використання методу з темпоральними правилами, програмна реалізація удосконаленого методу формування пояснень

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Отримання завдання	04.12.2023	виконано
2	Аналіз особливостей рекомендаційних систем	04.12 – 8.12	виконано
3	Дослідження методів формування рекомендацій	08.12 – 10.12	виконано
4	Постановка задачі дослідження	10.12 – 11.12	виконано
5	Дослідження методів формування пояснень в інтелектуальних системах	11.12 – 15.12	виконано
6	Аналіз методу побудови пояснень з урахуванням уподобань користувача	15.12 – 18.12	виконано
7	Удосконалення методу побудови пояснень з урахуванням змін уподобань користувача	18.12 – 25.12	виконано
8	Виконання інформаційної технології удосконаленого методу	25.12 – 29.12	виконано
9	Імплементация інформаційної технології	29.12 – 03.01	виконано
10	Програмна реалізація удосконаленого методу	03.01 – 06.01	виконано
11	Оформлення пояснювальної записки до кваліфікаційної роботи	06.01 – 10.01	виконано
12	Захист роботи в ЕК	22.01.2024	виконано

Дата видачі завдання 04 грудня 2023 р.

Студент _____

 (підпис)

_____ Лілія Тимченко
 (власне ім'я, прізвище)

Керівник роботи _____

 (підпис)

_____ проф. каф ІУС Оксана ЧАЛІА
 (посада, власне ім'я, прізвище)

РЕФЕРАТ

Пояснювальна записка до магістерської кваліфікаційної роботи містить: 91с., 4 розділи, 20 рис., 8 табл., 1 додаток, 30 джерел.

ЕЛЕКТРОННА КОМЕРЦІЯ, ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ, МЕТОД УРАХУВАННЯ ТЕМПОРАЛЬНОЇ ДИНАМІКИ, РЕКОМЕДАЦІЙНІ СИСТЕМИ, ФОРМУВАННЯ РЕКОМЕНДАЦІЙ.

У роботі виконано огляд методів формування пояснень у проектах рекомендаційних систем. Проаналізовано існуючі методи формування пояснень. На підставі проведеного аналізу запропоновано покращений метод формування пояснень у проектах рекомендаційних систем з урахуванням темпоральної динаміки змін уподобань користувача.

В процесі дослідження отримані такі результати: визначені існуючі методи надання рекомендацій; визначені існуючі властивості пояснень у рекомендаційних системах; визначені існуючі методи формування пояснень у інтелектуальних системах; визначені існуючі методи формування пояснень у рекомендаційних системах; виконана класифікація методів формування пояснень та опис удосконаленого методу на основі темпоріалу; проведено експериментальну перевірку удосконаленого методу.

ABSTRACT

The explanatory note to the master's qualification thesis contains: 91 p., 4 sections, 20 figures, 5 tables, 1 supplement, 30 sources.

E-COMMERCE, FORMATION OF RECOMMENDATIONS, INTELLIGENT SYSTEMS, METHOD OF TAKING INTO ACCOUNT TEMPORAL DYNAMICS, RECOMMENDATION SYSTEMS.

The paper reviews the methods of forming explanations in the projects of recommendation systems. The existing methods of forming explanations are analyzed. On the basis of the analysis, an improved method of forming explanations in the projects of recommendation systems has been proposed, taking into account the temporal dynamics of changes in user preferences. In the course of the study, the following results were obtained: the existing methods of providing recommendations were determined; the existing properties of explanations in recommender systems are determined; the existing methods of formation of explanations in intelligent systems are defined; the existing methods of forming explanations in recommendation systems are defined; a classification of methods for the formation of explanations and a description of the improved method based on the temporal are carried out; Experimental verification of the improved method was carried out.

ЗМІСТ

Скорочення та умовні позначки.....	7
Вступ.....	8
1 Аналіз пояснень в рекомендаційних системах.....	10
1.1 Аналіз особливостей рекомендаційних систем.....	10
1.2 Дослідження методів формування рекомендацій.....	17
1.3 Дослідження методів формування пояснень в інтелектуальних системах.....	27
1.4 Постановка задачі дослідження.....	32
2. Удосконалення методу побудови пояснень при зміні вимог користувача рекомендаційної системи.....	34
2.1 Методи побудови пояснень на основі правил.....	34
2.2 Опис динаміки вподобань користувачів з використанням темпоральних правил.....	40
3 Інформаційна технологія побудови пояснень.....	47
3.1 Опис інформаційної технології	47
3.2 Опис вхідних даних для інформаційної технології.....	50
4 Експериментальна перевірка методу побудови пояснень.....	52
4.1 Програмна реалізація методу побудови пояснень.....	52
4.2 Експериментальна перевірка удосконаленого методу.....	57
Висновки.....	61
Перелік джерел посилання.....	63
Додаток А Графічні матеріали.....	67

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ

РС – Рекомендаційна система

ШІ – Штучний інтелект

МН – Машинне навчання

ШІ – Штучний інтелект

БД – База даних

XAI - "Explainable Artificial Intelligence" Зрозуміла Штучна
Інтелектуальність

ВСТУП

Сучасний розвиток інформаційних технологій створює нові можливості для пропозицій продуктів і послуг кінцевим споживачам. В системах електронної комерції персоналізація ґрунтується на факторах, таких як регіон проживання, перегляд товарів та історія покупок. Ці дані служать основою для формування власниками компаній пропозицій, які максимально відповідають індивідуальним уподобанням кожного клієнта у відповідному регіоні.

Системи рекомендацій використовуються для створення процесу персоналізації продажів, надаючи персоналізовану інформаційну підтримку вибору користувача та пропонуючи споживачам індивідуальні рекомендації на основі їхнього профілю користувача. Ці системи є ефективними в електронній комерції, веб-пошукових системах, а також у формуванні рекомендацій для програм цифрового телебачення та інших галузях. Використання рекомендацій дозволяє підвищити обсяги продажів у сферах електронної комерції, покращити рівень обслуговування у банківських системах, особисто налаштувати та підвищити ефективність процесів навчання, а також швидко встановлювати зв'язки в соціальних мережах, враховуючи індивідуальні інтереси користувачів. Через це такі системи отримали широке розповсюдження протягом останнього десятиріччя.

Пояснення відіграють важливу роль у рекомендаційних системах, оскільки вони можуть впливати на вибір та задоволеність користувачів. Останнім часом виробники товарів онлайн збільшують інвестиції у дослідження та розробку сучасних методів та підходів з метою надання якісних пояснень користувачам.

Сучасні методи рекомендацій не вирішують низку взаємопов'язаних проблем, які негативно впливають на якість рекомендацій. Серед цих проблем варто відзначити холодний старт, спам-атаки, неузгодженість інтересів

користувачів і проблему масштабованості. Холодний старт виникає внаслідок ускладненості надання рекомендацій користувачам, які тільки що приєдналися до системи. У таких випадках користувачі ще не мали достатньо можливостей оцінити різноманіття товарів, що призводить до труднощів у передбаченні їхніх інтересів рекомендаційною системою.

Одним з основних та широко використовуваних методів є рекомендації на основі методу з часовими залежностями (темпоріального методу), де система пропонує товари, які отримали високі рейтинги від користувачів з аналогічними смаками та інтересами. У системах, які використовують цей метод, профіль користувача формується на основі оцінок та часових норм, які користувач присвоює різним елементам. Надання пояснень разом із рекомендаціями приводить до кращого розуміння рекомендаційної системи.

Таким чином введення темпоральних часових правил, які відображають зміну кількості продажів з часом, може сприяти вирішенню багатьох проблем та покращенню процесу надання рекомендацій.

1 АНАЛІЗ ПОЯСНЕНЬ В РЕКОМЕНДАЦІЙНИХ СИСТЕМАХ

1.1 Аналіз особливостей рекомендаційних систем

Швидкі темпи росту обсягу доступної цифрової інформації та збільшення числа користувачів Інтернету породили потенційну проблему інформаційного перенасичення, що ускладнює своєчасний доступ до цікавого контенту в Інтернеті. Інформаційно-пошукові системи, такі як Google та Bing, частково розв'язали цю проблему, але в них відсутня пріоритизація та персоналізація інформації, де система порівнює доступний вміст з інтересами та вподобаннями користувача. Це призвело до збільшення попиту на рекомендаційні системи більше, ніж будь-коли раніше. Рекомендаційні системи є інформаційними фільтрами, що вирішують проблему інформаційного перенасичення, відокремлюючи важливий фрагмент інформації з об'ємної та динамічно генерованої маси відповідно до вподобань, інтересів та спостереженої поведінки користувача. Система рекомендацій має можливість передбачати, які елементи чи продукти будуть віддані перевагою конкретному користувачу на основі його профілю.

В наш час рекомендаційні системи мають широке поширення та застосування у багатьох галузях. В першу чергу, їх використовують у сфері інтернет-комерції для допомоги користувачам у виборі відповідних товарів. Такі сервіси аналізують інформацію про вподобання користувачів та стараються пропонувати їм корисні товари. Наразі існує безліч методів формування рекомендацій, кожен з яких має свої переваги та недоліки. Тому дослідження в даній області є актуальним.

Якість рекомендаційної системи залежить від різноманітних властивостей її вироблених рекомендацій, які зараз оцінюються за допомогою показників продуктивності, що відображають обмежене розуміння поведінки системи. Деякі

розробники оцінюють точність рекомендацій за допомогою метрики, яка вимірює, наскільки ефективно рекомендаційна система відображає набір відомих користувачевих уподобань. Точність визначається шляхом виділення частини набору даних з оцінками у вигляді відомих уподобань та використання інших оцінок для навчання рекомендаційної системи та генерації рекомендацій. Відомі переваги можуть бути використані в розрахунках деяких показників, таких як точність, повнота, середньоквадратична помилка і інші. Для розрахунку оцінки того, наскільки точно алгоритм рекомендацій працює з цим конкретним набором даних.

Структура рекомендаційної системи представлена на рисунку 1.1.

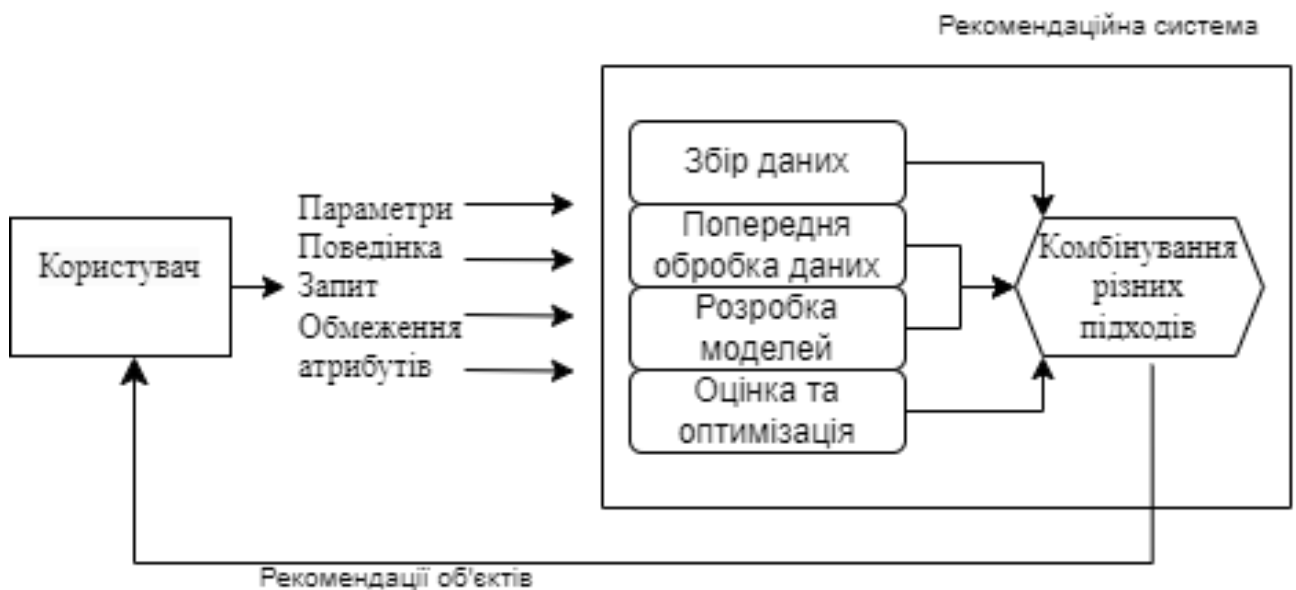


Рисунок 1.1 – Структура рекомендаційної системи

Останніми роками спостерігається зростаючий інтерес до метрик оцінки рекомендаційних систем, які є більш орієнтованими на користувача. Також визнано, що багато рекомендаційних систем працювали як чорні ящики, не забезпечуючи прозорості в процесі рекомендацій, і не надавали жодної

додаткової інформації, яка супроводжує рекомендації, крім самих рекомендацій. Пояснення може забезпечити прозорість, викриваючи міркування та дані, які стоять за рекомендацією. Це може бути, наприклад, пояснення, розміщені на Amazon, як от: "Споживачі, що придбали цей товар, також хотіли купити...". Пояснення може також служити іншим цілям, таким як допомога створення довіри та забезпечення лояльності користувачів, підвищення рівня задоволення, полегшення та прискорення пошуку користувачів, виявлення їх вподобань і заохочення їх спробувати чи придбати рекомендований продукт.

Отже, ми виділяємо різні види пояснень, одні, наприклад, надають пояснення того, як спрацьовує система надання рекомендацій (прозорість), та пояснення вибору товару користувачем (ефективність). Пояснення що може бути більш ефективним можна сформулювати так "Рекомендуємо розглянути вам варіант X, бо ...". Експертні системи можна вважати попередниками рекомендаційних систем. Тому ми коротко розглянемо дослідження методів оцінювання пояснень в експертних системах та інтелектуальних системах, а також оцінено розвиток рекомендаційних систем, які привертають дедалі більший дослідницький інтерес з моменту появи експертних систем. У вищенаведених прикладах ефективних та прозорих пояснень рекомендаційних систем ми виявили, що два критерії оцінки можуть взаємо виключатися. Процес роботи пояснень ілюстровано на рисунку 1.2, де видно, що пояснення можуть набирати різних форм, використовуючи текст та графіку.

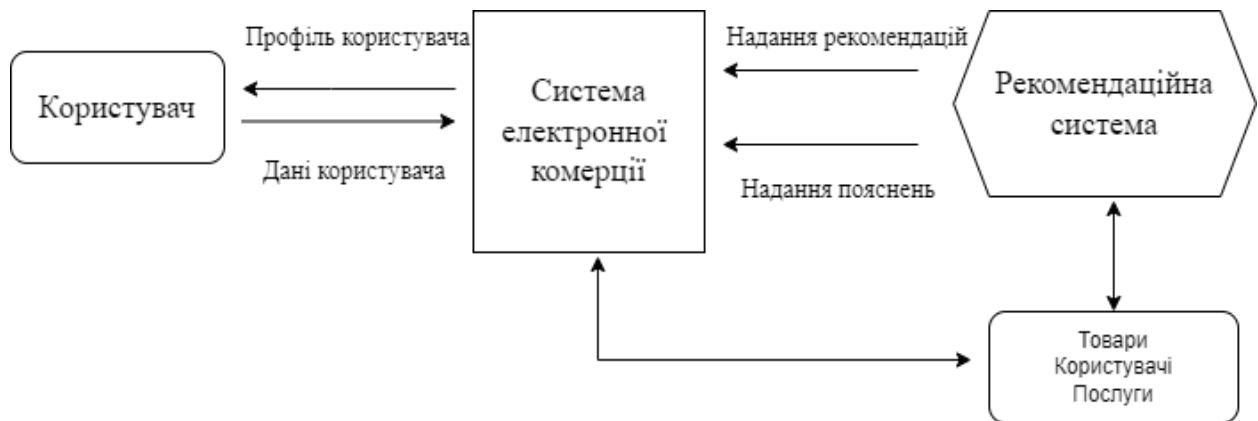


Рисунок 1.2 – Приклад процесу роботи пояснень рекомендаційної системи

На мою думку, база рекомендаційної системи має вплив на оцінювання вимог та пояснень, і я розглядаю це в контексті показників оцінки, що зазвичай застосовуються у рекомендаційних системах (наприклад, точність та охоплення). У продовж аналізу ми розглядаємо та надаємо приклади пояснень, а також оцінюємо основний алгоритм роботи пояснень та вимог у системах рекомендацій. При цьому, пояснення не від'єднуються від самих рекомендацій та методу комунікації користувачів із системою рекомендацій: ці фактори взаємодіють між собою, і можливість їх ефективного формулювання впливає на здобуття мети пояснення. Крім того, основний алгоритм системи рекомендацій може мати вплив на види пояснень РС, які будуть сформовані, при цьому існує можливість, що обрані розробником системи пояснень можуть не відобразити основний алгоритм. Це переважно стосується обчислювальних складних алгоритмів, де формулювання пояснень може бути ускладнене через параметри, такі як загальна фільтрація. У подібних ситуаціях розробник має зважувати компроміси різних цілей пояснення, таких як прозорість і задоволеність. Рекомендаційні системи використовуються для створення персоналізованого списку предметів, таких як товари або послуги, які співпадають з інтересами конкретного користувача.

Ці системи генерують рекомендації, базуючись на схожості між предметами або відповідних критеріях користувачів. Для створення рекомендацій необхідно отримати вхідні дані, це інформація про покупки товарів та їх рейтинги, що надані іншими користувачами систем. Рекомендаційні системи допомагають користувачеві простіше то швидше зробити вибір, що призводить до збільшення обсягів продажу. Таким чином, ці системи отримали широке розповсюдження в сфері продажу товарів, у сервісах переглядів онлайн фільмів, у розповсюдженні туристичних турів, наданні рекомендацій для різноманітних онлайн-конференцій та інших сферах.

Під час здійснення покупок люди купують товари, які їм подобаються або які рекомендують інші люди, спираючись на довірі їхнього вибору. В наш час цифрових технологій, інтернет-магазини пропонують своїм клієнтам величезний асортимент товарів. Щоб сприяти їм у пошуку та придбанні найбільш підходящих товарів серед цього розмаїття, вони використовують системи рекомендацій. Крім того, постачальники контенту, такі як музичні та фільмові сервіси, а також соціальні мережі, використовують системи рекомендацій для управління вмістом та створення ефективних рекомендацій для користувачів.

Простіше кажучи, системи рекомендацій працюють, як автоматизована форма помічника клієнтів, яка не лише показує сам продукт, який ви запитували, але також показує ті, які якимось чином пов'язані або можуть вам також сподобатися під час перегляду. Вони є одними з найпопулярніших сервісів машинного навчання, які використовуються в бізнесі для персоналізації контенту для клієнтів та споживачів.

Таким чином, система рекомендацій(РС) — це порадишник покупок, який тихо спостерігає за кожним вашим кроком, щоб краще порозуміти ваші вподобання. Він використовує технологію, яка вивчає дані користувача та надає персоналізовані пропозиції.

Системи рекомендацій широко використовуються в різних галузях,

включаючи електронну комерцію (ЕК), розваги та контент-платформи. Головна задача рекомендаційної системи полягає у покращенні взаємодії з користувачем, пропонуючи привабливі варіанти, які відповідають вашим індивідуальним уподобанням.

Очевидною метою системи рекомендацій є надання користувачам рекомендацій відповідних продуктів. Як сказав Стів Джобс: «Часто люди не знають, чого вони хочуть, доки їм це не покажуть» [6]. По словам Йова, однією з другорядних цілей системи рекомендацій є показ користувачам продуктів, які вони можливо не бачили раніше, але які їм можуть сподобатися. Вдалий вибір рекомендацій сприяє підвищенню загальної задоволеності користувачів, збільшуючи ймовірність повторного використання веб-сайту або додатку. Основні задачі РС представлені на рисунку 1.3.



Рисунок 1.3 – Задачі рекомендаційних систем

Системи рекомендацій працюють на основі складних алгоритмів, які обробляють та аналізують великі обсяги даних, такі, як поведінка користувачів, історія перегляду та покупок. Вони використовують різноманітні методи, включаючи фільтрацію на основі вмісту та використання методів машинного навчання, щоб виявити закономірності в поведінці користувачів.

Популярність систем рекомендацій пояснюється їх здатністю розвиватися та адаптуватися. Вона аналізує ваші відгуки, чи то явні оцінки користувачів, чи неявні сигнали, щоб відповідати неперервним змінам у ваших уподобаннях. Крім того, системи рекомендацій усвідомлюють, що з часом ваші смаки можуть зазнати змін. Таким чином, система постійно аналізує вашу поведінку, аби гнучко адаптуватися до будь-яких змін.

Головною частиною у побудові системи рекомендацій є інформація, яку можна розділити на три типи: явні дані, неявні дані та опис продукту.

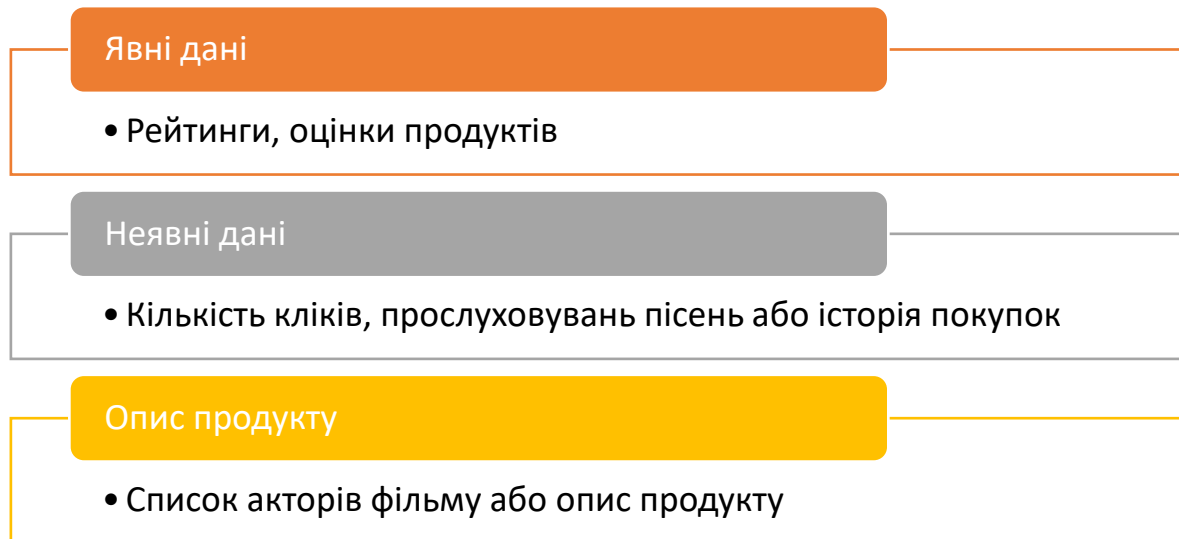


Рисунок 1.4 – Типи вхідних даних у системі рекомендацій

Зазвичай явні дані представлені у числовій формі, такі як, наприклад, рейтинг у формі 5 зірок, який користувач виставляє продукту. Прикладами явних

даних можуть бути оцінки продуктів, або популярність курсів, які користувачі вказують на навчальній платформі. Збір такого типу даних виявляється важким, оскільки він потребує додаткового внесення даних від користувачів, що призводить до необхідності витрат більше часу для отримання достатньої кількості оцінок для створення ефективної моделі машинного навчання.

Збір неявних даних виявляється простим. Це будь-які, що зосереджуються на тому, як користувач взаємодіє з доступними продуктами або контентом. Однак основна проблема цього типу даних полягає в трансформації поведінки користувача в його уподобання. Незважаючи на це, існують ефективні методи для вирішення цієї задачі. Наприклад, кількість відтворень пісні на Spotify, кількість кліків на посилання продукту або історія покупок на Amazon можуть служити прикладами неявних даних.

Останнім типом даних є опис продукту. Оскільки ці дані часто мають неструктурований вид (наприклад, представлені у вигляді вільного тексту), потрібна додаткова попередня обробка для отримання відповідної інформації та її подальшого структурування. Прикладами описів продукту можуть служити перелік акторів у фільмі на Netflix, інформація про виконавців на Spotify або опис товару на Amazon.

1.2 Дослідження методів формування рекомендацій

Нові тенденції у сфері інформаційних технологій з'являються завдяки величезній кількості інформації, яка доступна і постійно накопичується у сферах електронної комерції, реалізації та управління бізнес-процесами в торгівлі та формування зворотного зв'язку з користувачами щодо товарів і послуг. Метою

цього сектору є сприяння продажу товарів та надання персоналізованих пропозицій, пристосованих до індивідуальних уподобань споживачів.

Для обробки великих обсягів даних необхідне використання автоматизації і застосування рекомендаційних систем. Такі комплекси можуть знайти застосування не лише в електронній або оффлайн комерції, а й у сфері інформаційної безпеки, машинобудування та відбір продуктів та послуг. В даному контексті рекомендаційні системи виконують роль інструмента підтримки процесу прийняття рішень. Здатність системи ідентифікувати групи подібних користувачів чи об'єктів дозволяє використовувати її як альтернативний інструмент для пошуку інформації в Інтернеті та виявляти об'єкти, що не знаходять традиційні пошукові алгоритми. Крім того, РС застосовуються для фільтрації спаму або іншого шкідливого контенту. РС прогнозує, які об'єкти можуть зацікавити користувачів, на основі інформації про зібрані об'єкти (продукти, сайти, книги тощо) у програмне забезпечення. Існує три види систем рекомендацій:

- фільтрація на основі контенту (contentbased filtering);
- колаборативна фільтрація (collaborative filtering);
- гібридні методи (колаборативна +контентна фільтрація).

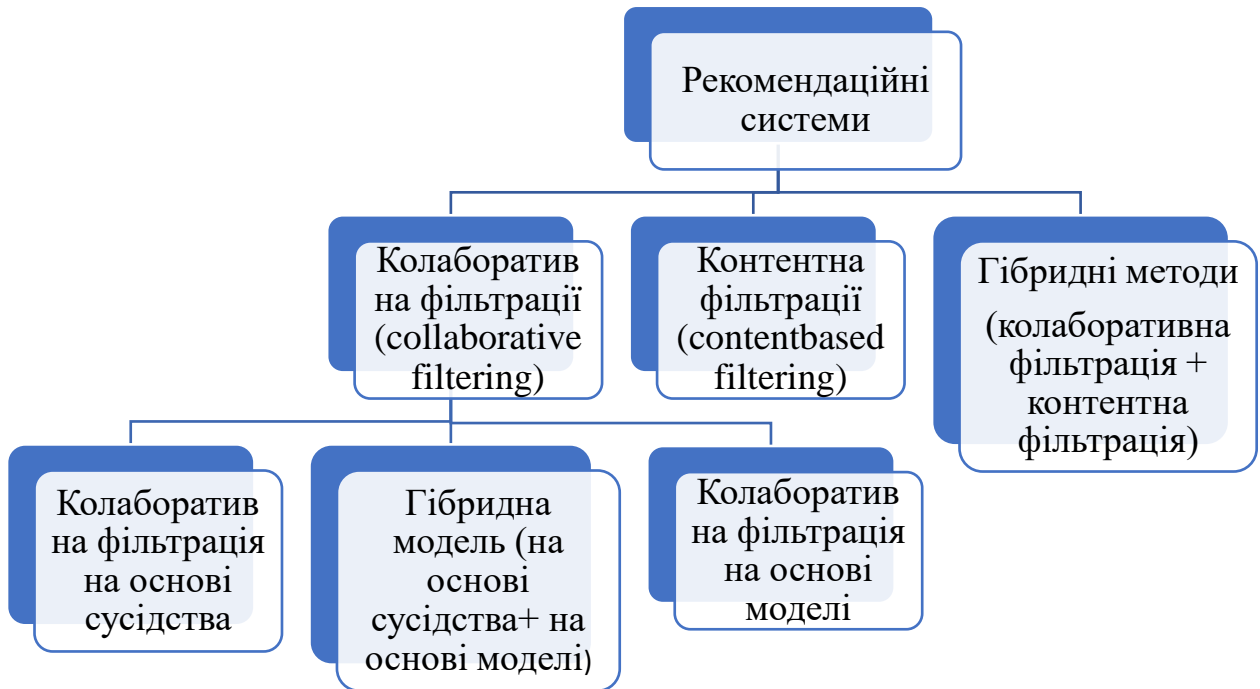


Рисунок 1.5 – Класифікація рекомендаційних систем

Один із способів створення рекомендаційних систем (РС) є метод колаборативної фільтрації. Цей метод створений на аналізі вподобань споживачів, де враховується використання історичних даних, такі як попередні покупки користувачів, їх оцінки та відгуки. Він також надає можливість створювати рекомендації на базі визначених груп клієнтів зі схожими вподобаннями. На сьогоднішній день різноманітні алгоритми включають в себе методи колаборативної фільтрації, які можна умовно виділити в одному з двох класів. Загальну класифікацію методів колаборативної фільтрації можна знайти у таблиці 1.1.

Таблиця 1.1 – Методи для побудови РС

Назва	Опис
1.Колаборативна фільтрація (КФ)	заснований на використанні інформації про оцінку, яку користувач дає об'єкту, для визначення схожості користувача та об'єкта.
1.1 Колаборативна фільтрація на базі сусідства	використовує коефіцієнти схожості між компонентами системи (споживачами, об'єктами) для створення рекомендацій на базі їхньої схожості.
1.2 КФ на основі моделі	розробляються з використанням алгоритмів машинного навчання, щоб спрогнозувати оцінку користувача на продукт що немає ще оцінки.
2. Фільтрація на основі вмісту	використовує атрибути товару для того, щоб запропонувати інші товари, схожі на вподобаний користувачем, на основі його минулої поведінки та відкритих відгуків.
3.Гібридні методи	поєднують у собі кілька методів та/або моделей роботи рекомендаційних систем (РС).

Колаборативна фільтрація (КФ) - це методологія, в якій основна концепція полягає в використанні інформації щодо попередньої поведінки, оцінок або інших додаткових даних вже існуючої групи користувачів для передбачення, які об'єкти поточного користувача системи ймовірно будуть йому до вподоби або зацікавлять. Ці типи систем широко використовуються в різних інформаційних системах, зокрема, як інструмент на інтернет-магазинах для персоналізації контенту відповідно до потреб конкретного клієнта та стимулювання додаткових покупок. Для ефективної роботи цього підходу необхідна значна кількість оцінок від користувачів чи оцінок елементів рекомендації. Цей метод не ефективний для

нових користувачів, нових елементів або в обох випадках через відсутність оцінок у системі. [2, 5].

Формальну роботу РС з використання колаборативної фільтрації можна представити наступним чином на рисунку 1.6.

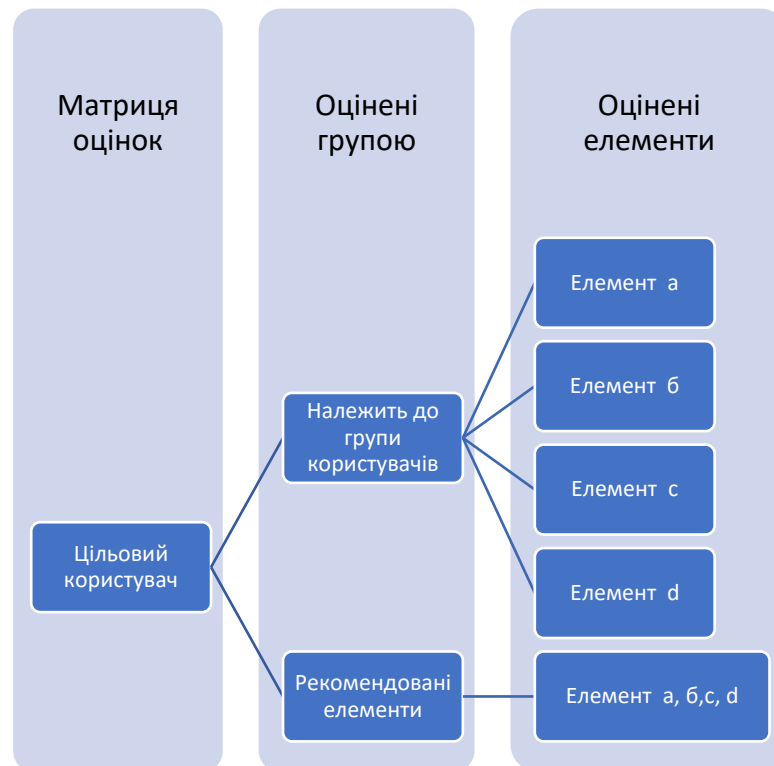


Рисунок 1.6 – Принцип роботи рекомендаційної системи на основі КФ

Контентна фільтрація надає рекомендації, опираючись на дії (поведінки) користувача, такі як відвідані раніше веб-сторінки чи поставлені оцінки товарів.

Гібридні методи об'єднують обидва підходи. Злиття контентної та колаборативної фільтрації дозволяє отримати більш точні рекомендації. В гібридних методах спочатку використовується контентна фільтрація, а потім її результати уточнюються за допомогою колаборативної фільтрації [3]. Існують дві обширні категорії колаборативної фільтрації і одна гібридна:

- КФ на базі сусідства;
- КФ на базі моделі;
- гібридна фільтрація КФ на базі пам'яті та КФ на базі моделі.

КФ на базі сусідства знайома нам, як методика на базі пам'яті, будує взаємозв'язки між елементами або між користувачами. Даний підхід формує уподобання користувача до елемента, враховуючи оцінки тих самих елементів, які були виставлені цим користувачем. Він рахує подібності стосовно двох користувачів та між об'єктами. Далі створює прогноз для споживача, застосовуючи середнє зважене усіх оцінок. Методи на базі сусідства відзначаються своєю простотою, ефективністю і здатністю надавати більш точні та конкретні рекомендації. Даний метод найчастіше використовуються в невеличких інформаційних системах(ІС) через свою доступність і простоту в реалізації [5].

Колаборативна фільтрація(КФ) на базі моделі базується на зібранні та споживанні інформації стосовно користувачів та їх взаємодії з певними елементами. Даний підхід являється вмістовним та надає більш точніші рекомендації, так як спрямований на розкритті прихованих частин поведінки, які можуть зрозуміти взаємозв'язок між споживачами та елементами.

Фільтрація на базі вмісту (контентна фільтрація) формує рекомендації товарів на основі уподобань, враховуючи минулі взаємодії. У цьому методі робиться профіль споживача та профіль елемента системи, і відповідні параметри цих елементів мають відповідати вподобанням споживача. Даний підхід бере дані тільки для конкретного споживача та надає рекомендації тільки для нього, та не бере до уваги інших користувачів [5]. На рисунку 1.6 показаний приклад фільтрації вмісту.

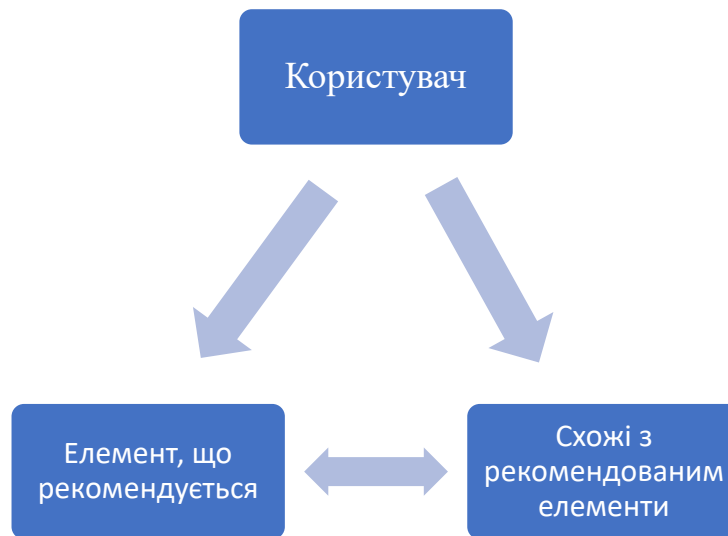


Рисунок 1.6 – Фільтрація на базі вмісту

У гібридну фільтрацію входить обидва методи - той, що базується на моделі, і той, що базується на сусідстві. Цей підхід відрізняється серед інших і найчастіше використовується при розробці РС для комерційних баз даних. Гібридний метод допомагає вирішити численні недоліки, такі як покращення точності прогнозів вибору користувачів та проблему втечі інформації. Але враховуючи всі переваги порівняно з іншими методами, гібридний підхід є більш складним і вимагає великих витрат на реалізацію та підтримку.

Існують кілька основних способів комбінування при створенні гібридних систем:

- реалізація колаборативних і контентних алгоритмів окремо та поєднання їх спрощень;
- додання деяких точних правил у колаборативну методику;
- додавання незначних колаборативних правил у методику контенту;
- створення загальної моделі з використанням правил двох методик.

Гібридні РС відрізняються від інших своєю спроможність складання оцінок з різних складових з точки зору представлення, а не методом

комбінування передбачуваних оцінок. Найчастіше рекомендовані елементи базуються поруч один з одним. Таким чином, основним описом даних систем являється комбінація презентації, а не навпаки комбінація передбачуваних оцінок.

Однак багато РС є фактично гібридними, мало теоретичних досліджень спрямовано на те, як перетворити алгоритми та коли можна очікувати вигоди від роботи гібридного методу.

Застосування рекомендаційних систем має численні переваги, такі як залучення нових користувачів, надання вигідних пропозицій, підвищення конкурентоспроможності та підняття прибутку. Але, як у будь-якій системі, рекомендаційні системи не обійшлися без недоліків. Розглянемо детальніше їх переваги та недоліки найпоширеніших підходів при створенні рекомендаційних систем.

Таблиця 1.2 – Переваги та недоліки рекомендаційних систем (РС)

РС	Переваги	Недоліки
1	2	3
РС на базі контента	<ul style="list-style-type: none"> – не враховуючи дані інших користувачів при створенні рекомендації; – не має проблеми холодного старту для оновлених товарів, так як маючи ознаки товарів, є можливість легко знайти схожі товари; – висновки та результати створення рекомендацій – інтерпретовані дані. 	<ol style="list-style-type: none"> 1. Проблема взаємодії нових користувачів із системою полягає у важкості старту через недостатню інформацію та відсутність попередніх взаємодій. 2. Створення груп схожих товарів може призвести до обмеження рекомендацій для користувачів, які отримують однакові пропозиції з обмеженої кількості товарів. 3. Недостатня кількість даних про товари ускладнює розрізнення та групування товарів, що негативно впливає на якість рекомендацій для користувачів.
РС на базі колабораційних систем, які засновані на сусідстві	<ul style="list-style-type: none"> – просте впровадження та зрозумілість в реалізації системи; – очікуваний результат для користувача полягає в отриманні схожих рекомендацій з урахуванням його минулих вподобань, 	<ol style="list-style-type: none"> 1. Недостатня кількість оцінок від користувачів для багатьох товарів у системі виникає через їх велику кількість, що особливо характерне для систем з колабораційним підходом на основі сусідства. 2. Шахрайство є загальною проблемою, оскільки користувачі можуть намірено надавати завищені оцінки деяким товарам та

Кінець таблиці 1.2

1	2	3
	<ul style="list-style-type: none"> – легке оновлення нових даних – можливість масштабування. 	<p>штучно занижувати їх для конкурентів.</p> <p>3. Відсутність різноманітності виявляється в тому, що користувачам постійно рекомендуються ті ж самі товари, не пропонуючи нових цікавих пропозицій.</p> <p>4. Проблема холодного старту виникає при з'яві нових користувачів або товарів у системі, які зустрічають труднощі отримання рекомендацій.</p>
<p>РС на базі колаборац. систем заснов. на моделі</p>	<ul style="list-style-type: none"> – гарна обробка розріджених матриць; – найкраща інтерпретація масштабності даних 	

Гібридні методи допомагають вирішити більшість труднощів, поєднуючи найкращі аспекти обох методів колаборативної фільтрації. Це дозволяє подолати обмеження, яке може виникнути від використання лише одного підходу, наприклад, того, що часто спостерігається у методі, який заснований на сходстві. Крім того, гібридні системи сприяють поліпшенню якості результатів рекомендацій. Незважаючи на всі ці переваги, важкість впровадження та високі витрати залишаються проблемами, з якими зіткнеться гібридний підхід.

Незважаючи на усі переваги, рекомендаційні системи стикаються з важливою проблемою, відомою як проблема холодного старту. Ця проблема виникає, коли до системи додається новий користувач, і в ній відсутні достатні дані для формування ефективних рекомендацій для нього. Це особливо актуально при розробці рекомендаційних систем для онлайн-кінотеатрів, де надходження нових фільмів і користувачів є постійним явищем. Реєстрація нових користувачів щоденна, і вони стають холодними користувачами, зовсім або мало взаємодіючи з каталогом впродовж першого дня після реєстрації. Пряме включення їх до робочої системи рекомендацій стає неможливим без належних даних. Чекати, коли вони стануть «теплыми користувачами» не є виходом, оскільки якісні рекомендації для них мають важливе значення, і непослідовні рекомендації можуть вплинути на їх подальше користування сервісом та призвести до втрати клієнтів. Реакція на цю проблему включає впровадження алгоритмів, які вирішують проблему холодного старту та запобігають втраті клієнтів.

На рис. 1.7 представлено 3 категорії проблем холодного старту:

– зміна вимог користувача: коли запускається в роботу новий користувач, то система не розуміє, які його вподобання, що ускладнює надання надійних рекомендацій;

– нові товари та послуги: нові об'єкти системи(товари чи послуги), може бути відповідна інформація про вміст, але з користувачем не повинна бути взаємодія;

– новий користувач: нові користувачі мають право входити в систему без будь-якої взаємодії з системами та не мати ніяких персональних рекомендацій гостям.

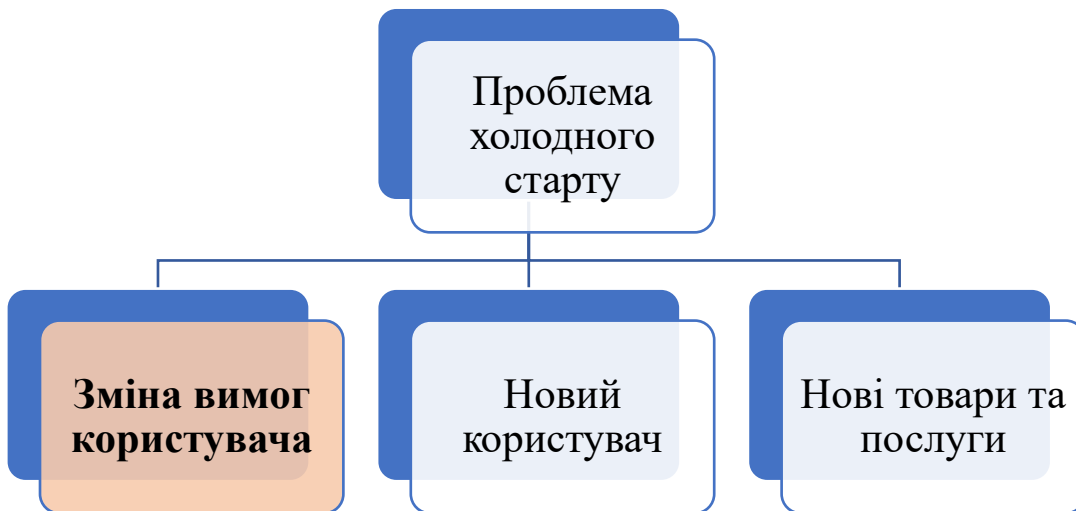


Рисунок 1.7 – Проблеми холодного старту

1.3 Дослідження методів формування пояснень в інтелектуальних системах

Оцінка рекомендаційних систем продовжує залишатися об'єктом інтенсивних досліджень в цьому напрямку. Від початку існування перших РС їх ефективність традиційно пов'язується із точністю прогнозування рейтингів. Щоб виразити це інакше, прогнозовані рейтинги порівнюються з реальними, та все, що відрізняється між ними оцінюють за допомогою критеріїв, таких як середня абсолютна похибка та середньоквадратична похибка [8]. Велика кількість досліджень спрямована на оцінці переліку рекомендацій з використанням метрик.

Оцінка РС стала ключовим об'єктом вивчення у цій галузі ще на початку їх існування та і зараз є предметом постійних досліджень. Зазвичай розрізняють

два основних підходи щодо оцінювання : онлайн та офлайн. В роботі розглядається онлайн оцінювання.

Застосовуючи заходи машинного навчання та підходи пошуку інформації, оцінка РС спирається на вилученні з системи елементів наявних знань щодо вподобань користувачів (дані з тестів). На практиці це порівняння проводиться за рахунок вимірювання точності прогнозування рейтингів, застосовуючи показники, які базуються на помилках.

З точки зору ефективності рекомендацій стосовно користувачів можемо сказати, що якість ранжування рекомендованих об'єктів є важливішою, ніж точність у прогнозах конкретних показників рейтингу [10]. Стосовно проблеми рекомендацій можна сказати, що вподобання користувачів представлені у формі оцінок, та мета рекомендаційного алгоритму полягає у прогнозі скритих оцінок на основі відкритих оцінок та, додаткової інформації стосовно споживачів, товарів та вибору. В такому випадку точність рекомендацій оцінюється способом вимірювання похибки між прогнозованими і явними рейтингами. Використовуються в такому випадку такі метрики, як середня абсолютна похибка та середньоквадратична.

Метрики, спрямовані на точність, визначають кількість релевантних та нерелевантних рекомендованих елементів, що повторно перевіряються. Раніше було розроблено у напрямку інформаційного пошуку дуже велику кількість метрик, різних методологій та різних наборів даних. Рекомендації можна прирівняти до задачі інформаційного пошуку, де користувачі мають неявні потреби у просторі об'єктів, що можуть задовольнити їхні цілі.

Завдання РС лежить у обранні, ранжуванні та представленні користувачеві набору об'єктів, які найкращим чином задовольняють його потреби. Природа потреб користувача та причини того, чому той чи інший об'єкт задовольняє їх, часто не можуть бути точно та повністю виявлені. Адаптація методології оцінювання інформаційного пошуку, що ґрунтується на релевантності, стає

природною, і вона включає в себе отримання ручних міток ревалентності для рекомендованих об'єктів щодо потреб користувача [15]. На рисунку 1.7 представлені методи для оцінки ефективності РС, які найбільше досліджувалися.

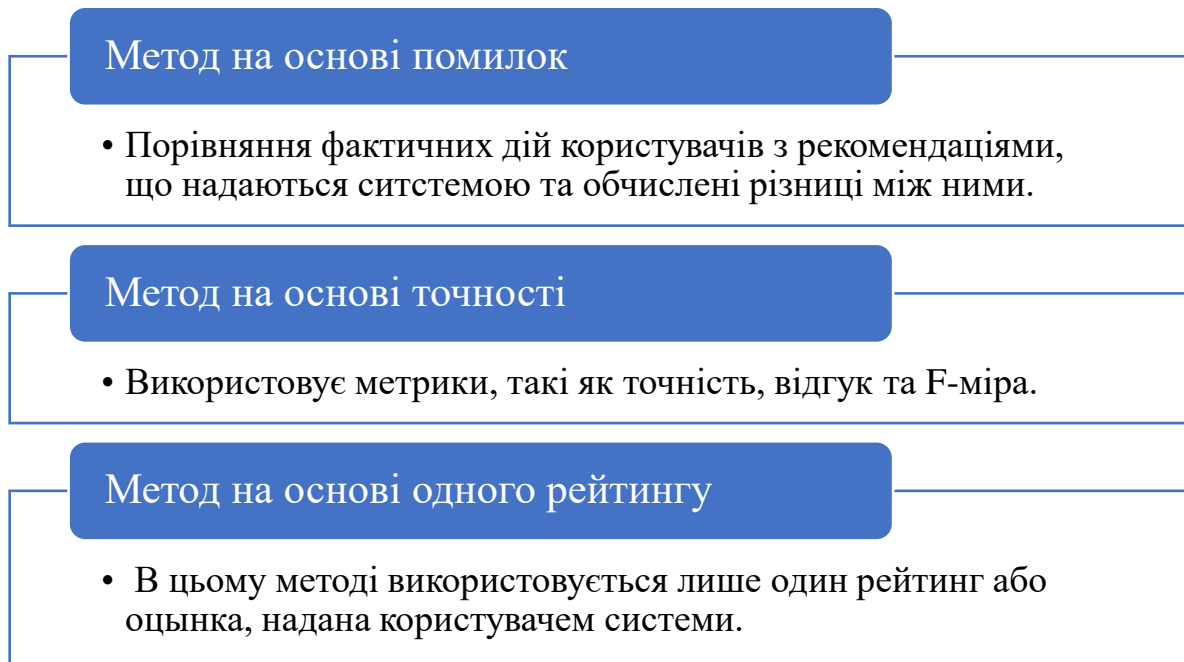


Рисунок 1.8 – Існуючі методи оцінки ефективності РС

Один із класичних припущень у літературі, що досліджує рекомендаційні системи, полягає в тому, що користувачі віддають перевагу системам, які забезпечують значно точніші прогнозування. Методи, які направлені на досягнення точності, підрозділяються на три складові: метрики, що застосовують лише 1 рейтинг; метрики, які оцінюють за допомогою порівняння два рейтинги (один із них основний або ідеальний); та метрики з галузі машинного навчання. До прикладів методів, що ґрунтуються на одному рейтингу, відносяться точність, згадування, середня точність та ранг. Кожен підхід представляє якість ранжування з невеличкою варіацією в оцінці. Наприклад, точність відображає відсоток відповідних рекомендованих позицій,

а взаємний ранг показує відсоток відповідних позицій, які запропоновані були. Ці два показники являються обернено пропорційними, так як збільшення рівня відновлення, та призводить до зниження їх точності.

Значний прогрес в галузі машинного навчання викликав появу нового етапу у розвитку програм штучного інтелекту, охоплюючи такі галузі, як транспорт, безпека, медицина, фінанси та оборона. Ці програми пропонують значні переваги, але одночасно стикаються з проблемою пояснення своїх прийнятих рішень і дій людям. Програма пояснюваного штучного інтелекту (ХАІ) DARPA націлена на створення систем штучного інтелекту, які володіють навченими моделями та прийняттям рішень, зрозумілими для кінцевих користувачів, і які можуть викликати довіру.

Щоб досягнути таку мету, необхідні методи, які роблять моделі більш зрозумілими, розробка ефективних інтерфейсів для пояснень та врахування психологічних аспектів для ефективного розуміння пояснень. Команди розробників ХАІ активно працюють над першими двома аспектами, створюючи методи машинного навчання та розробляючи принципи та стратегії взаємодії людини з комп'ютером для створення ефективних пояснень. Ще одна команда в галузі ХАІ розв'язує третю задачу, узагальнюючи та застосовуючи психологічні теорії пояснень. Це допомагає оцінювачам ХАІ визначити відповідну структуру оцінювання, яку групи розробників будуть використовувати для тестування своїх систем [18]. Розглянемо методи формування пояснень в інтелектуальних системах(ІС).

Таблиця 1.3 - Методи формування пояснень в ІС

Методи	Характеристика
1	2
Метод оцінювання пояснень з використання теорії можливостей	Метою цього методу є розробка оцінки пояснень з урахуванням відмінностей у вхідних даних та відповідному рішенні системи штучного інтелекту.
Ансамблевий метод оцінки ефективності РС	Ці підходи включають розрахунок помилок для кожної рекомендації, яку система надає користувачеві, шляхом визначення різниці між прогнозованим та фактичним значенням. Після цього сумуються отримані помилки, і середнє значення використовується як показник ефективності рекомендаційної системи.
Метод фільтрації на основі вмісту	Вони мають можливість рекомендувати нові елементи, навіть якщо немає оцінок, наданих користувачами. Таким чином, навіть якщо база даних не містить налаштувань користувача, точність рекомендацій не впливає.
Контентна фільтрація	Концепція фільтрації на основі вмісту полягає в визначенні категорій продуктів за допомогою конкретних ключових слів, розумінні вподобань користувача, пошуку цих ключових слів у базі даних і рекомендації різних продуктів з схожими характеристиками.

Кінець таблиці 1.3

1	2
Спільна фільтрація - методи засновані на пам'яті - модельні методики	Спільна фільтрація - це метод прогнозування, який використовується для контенту, що не може бути легко та адекватно описано метаданими (наприклад фільми або музика). У цьому методі спільної фільтрації створюється база даних, яка відображає уподобання користувачів до різних елементів у вигляді матриці користувача. Потім система порівнює користувачів із схожими інтересами, обчислюючи схожість їхніх профілів, для надання рекомендацій.

1.4 Постановка задачі дослідження

Існуючі підходи до створення рекомендацій ґрунтуються на застосуванні колаборативного фільтрування, яке встановлює латентні фактори схожості користувачів та товарів. Однак цей механізм є непрозорим для користувача, що знижує довіру до отриманих рекомендацій і як наслідок може зменшити кількість покупок в електронній комерції. Для підвищення довіри до рекомендацій доцільно використовувати пояснення. Такі пояснення мають бути придатними для нових споживачів РС та враховувати можливі зміни у потребах користувача. Тому для побудови пояснень при зміні вимог користувача доцільно враховувати темпоральний аспект та використовувати темпоральні правила.

Об'єктом дослідження в цій магістерській атестаційній роботі є процес побудови пояснень в рекомендаційних системах.

Предметом дослідження являються методи побудови пояснень в рекомендаційних системах (РС).

Метою даної кваліфікаційної роботи є дослідження методів формування пояснень при зміні вимог користувача в рекомендаційних системах з використанням темпоральних правил.

Для досягнення мети роботи треба визначити такі задачі:

- проведення аналізу особливостей рекомендаційних систем;
- дослідження методів формування рекомендацій;
- дослідження методів формування пояснень в інтелектуальних системах;
- аналіз методів побудови пояснень на основі правил;
- удосконалення методу надання пояснень з урахуванням темпоральної динаміки вподобань користувача;
- проведення експериментальної перевірки удосконаленого методу надання пояснень у проектах рекомендаційних систем.

2 УДОСКОНАЛЕННЯ МЕТОДУ ПОБУДОВИ ПОЯСНЕНЬ ПРИ ЗМІНІ ВИМОГ КОРИСТУВАЧА РЕКОМЕНДАЦІЙНОЇ СИСТЕМИ

2.1 Методи побудови пояснень на основі правил

Рекомендаційні системи, навіть при наданні користувачам нових рекомендацій, іноді розглядаються як непрозорі, тобто як чорні ящики, які не мають пояснювальної інформації стосовно них. Таким чином додавання пояснень до РС є дуже важливим елементом. Вони роз'яснюють причини рекомендацій для конкретних елементів, полегшують сам процес прийняття рішень, таким чином сприяють задоволенню та довіри до самої системи, яка в свою чергу підвищує коефіцієнти конверсії.

Самі пояснення представляють собою опис товару чи послуги, завдяки якому користувач краще розуміє підходить йому рекомендаційний товар чи ні [23]. Додавання пояснень до рекомендацій значною мірою може покращити розуміння РС.

Протягом останніх років встало питання зробити процес надання рекомендацій більш прозорим для споживачів. Об'єктом дослідження стала велика проблема автоматизованого генерування в цій галузі та представлення пояснень. Вже сьогодні запропоновані незначні засоби пояснень, які вже використовуються на сторінках сайтів електронної комерції. Ці пояснення представлені у різному вигляді інтерфейсів, наприклад описами, гістограмами, графіками, діаграмами тощо. В галузі РС визначено такий факт, що показники точності можуть частково оцінити систему (точність, середня помилка і запам'ятовування).

Пояснення високої якості може сприяти формуванню довіри користувачів та їхньої лояльності під час вибору товарів, підвищити їхнє задоволення та

полегшити пошук необхідного, а також переконати користувачів придбати або випробувати новий рекомендований товар [25].

В таблиці 2.1 можемо побачити низку цілей, які можуть виконувати різноманітні пояснення в РС.

Таблиця 2.1 – Цілі пояснень у РС

Властивість пояснень РС	Надання товару
Прозорість	Вказує на ясність та зрозумілість інформації, наданої користувачам щодо рекомендацій
Перевірка	Користувачам надається можливість повідомити системі про те, що вона функціонує некоректно.
Довіра	Змінює довіру користувачів у сторону зростання.
Переконливість	Стимулюю користувачів випробувати чи придбати товар.
Ефективність	Сприяє у користувачів у прийнятті вірного рішення
Оперативність	Сприяє у швидшому прийнятті рішення користувачами
Задоволення	Підвищую зручність використання або задоволення

Після визначення цілей ми розглядаємо пояснення, які були впроваджені на основі конкретного базового алгоритму та різних "стилів пояснення". Важливо відзначити, що стиль пояснення може відповідати "стилю" конкретного алгоритму, незалежно від того, як саме були отримані чи вирішені відповідні рекомендації. Часто спостерігається розбіжність між методами одержання рекомендацій та відповідним стилем, використовуваним для пояснень. Таким чином, різноманітні методи пояснень можуть не завжди відповідати меті прозорості, але вони можуть служити іншим цілям пояснення. У таблиці 2.2 наведені приклади пояснень.

Таблиця 2.2 – Приклади пояснень у комерційних та академічних системах, визначені за стилем пояснення

Стиль пояснення	Приклад пояснення	Система
Випадковий	Якщо ви зосередили свій пошук на цьому фільмі, то ми ще можемо запропонувати вам схожі фільми, які отримали позитивні оцінки раніше	LovelyFilms.com
Колаборативна	«Споживачі, які купили даний товар також купили ще й і цей товар...»	Amazon.com
Контентна	Виступають упорядковані теги по перевагам або актуальності	MovieFilm.com
Діалогова	Діалог: «Що ми можемо запропонувати вам на сніданок? Може омлет чи каву?»	Advisor-one.com
Демографічна	«Для дітей це дуже привертає увагу і потребує зосередженості та серйозності, а для дорослих має високу історичну цінність»	Intriguee.com
На основі знань	«Нижча ціна і відповідно більше поїздок у метро», «Бокс 45 відрізняється від вашого прохання тільки кольором і буде найкращим боком при виборі будь-якого варіанту»	Trains.com

Прозорість — це не єдина ціль визначення пояснень, яку варто враховувати при відборі стилю пояснення. На сьогодні порівняння стилів пояснення, їхньої ефективності та пояснювальних цілей залишається не вивченими до кінця. Однак інші наукові дослідження досліджували, як різноманітність інтерфейсів пояснень впливає на різні мети надання пояснень. Незважаючи на це, типи пояснень, які можна створити, в певній мірі визначаються основним алгоритмом механізму рекомендацій. Для комерційних систем, де конфіденційність інформації має значення, можна використовувати методи на основі припущень. Навіть коли діалогові системи розглядаються як підхід, ми розглядаємо їх більше як стиль взаємодії, ніж конкретний алгоритм. У випадку стилів пояснень, що ґрунтуються на спільній роботі, передбачуваними

вхідними даними для механізму рекомендацій є рейтинги користувачів для елементів. Ці рейтинги використовуються для визначення подібних користувачів. Таких користувачів часто називають "сусідами". Далі прогноз для рекомендованого показника екстраполюється відповідно по оцінкам сусідів. З точки зору комерції одним з найголовніших критеріїв пояснень спільної праці є те, що використовується на Amazon.com: "Споживачі, що придбали даний товар замовили ще щось...". Таке пояснення показує, що користувач дивиться на товар, який його зацікавив. Система пропонує схожих користувачів (купивших цей товар) та пропонує товари, які вибрали інші подібні користувачі. Такі пояснення базуються на основі взаємодії, де рейтинги неявно показуються через споживача.

Отже, при зміні вимог користувача традиційні підходи колаборативної фільтрації можуть призводити до неправильних фільтрацій. Тому рекомендації слід надавати на основі характеристик об'єктів, таких як товари та послуги.

Користувач може змінювати свої вимоги відповідно до зміни демографічних характеристик, таких як завершення освіти, зміна сімейного стану, переїзд на нове місце проживання, зміна роботи і т.д. Однак частіше зміна вимог відбувається внаслідок подій, що відбуваються в регіоні проживання споживача. Таким чином, в даній ситуації є доцільним використання методу динаміки вподобань користувача з використанням темпоральних правил.

Ці методи представлені у вигляді підкатегорії, як самостійні методи моделювання [26].

В сучасному житті важко недооцінити значення машинного навчання, яке охоплює широкий спектр від розваг до різних галузей. Багато компаній намагаються використовувати моделі машинного навчання, оскільки вони демонструють більшу швидкість та точність порівняно із застарілими системами. Однак не треба ігнорувати той факт, що моделі машинного навчання можуть взяти на себе функції прийняття рішень, які раніше здійснювалися

людьми у різних галузях. Незважаючи на це, відсутність прозорості стосовно їхніх процедур прийняття рішень, внутрішніх механізмів чи неочікуваної поведінки може викликати розчарування та непорозуміння серед користувачів. Це також може мати негативний наслідок на життя людей, особливо в критичних областях, таких як фінансові операції, страхові послуги та виявлення раку в системі охорони здоров'я.

Таким чином проведемо систематичний огляд самостійних методів самопояснювального моделювання ХАІ та розглянемо їх більш детально.

Самопояснювальне моделювання представляє собою два підходи внутрішньої пояснюваності та спеціальної пояснюваності. Графічна модель всіх категорій та підкатегорій представлена на рисунку 2.1.

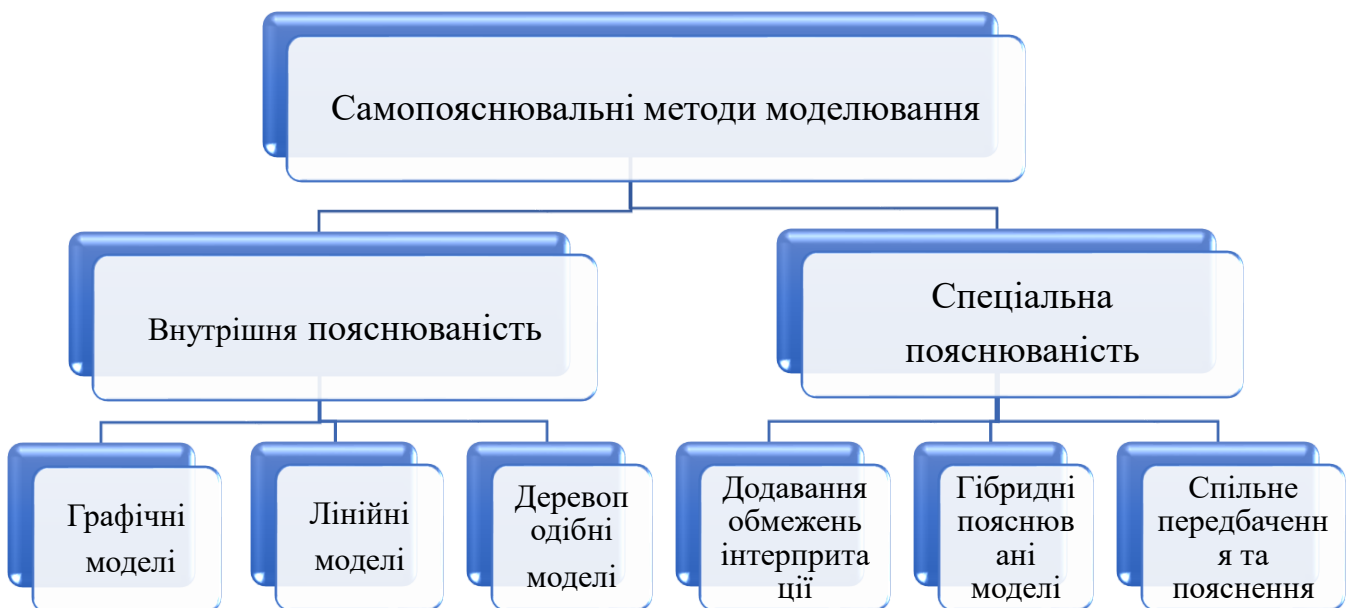


Рисунок 2.1 – Самопояснювальні методи моделювання на основі правил

Методи внутрішньої пояснюваності акцентуються на пояснюваності моделі, а не на її високій ефективності навчання. У цьому варіанті внутрішня та проста модель вважаються більш прагматичними, ніж складні моделі "чорного

ящика". Наприклад, у галузях соціальних наук та психології дослідники віддають перевагу моделюванню структурними рівняннями, оскільки воно може краще відобразити взаємозв'язок між прихованими факторами, ніж складна модель машинного навчання з недостатньою пояснюваністю. Підхід внутрішньої пояснюваності розділяється на три основні категорії: лінійні моделі (зокрема, узагальнені лінійні моделі, логістична регресія і т.д.), моделі на основі дерев (включаючи дерево рішень, правилові відповідності і т.д.) і, нарешті, графічні моделі, такі як байєсівські моделі, циклічні спрямовані моделі тощо.

Графічні моделі використовуються в галузі охорони здоров'я. Дослідники представили нейронну мережу, яка забезпечує візуальну інтерпретацію для різних наборів даних, пов'язаних з біологією. Крім того, в компаніях, що використовують прогнозні моделі, почали впроваджувати алгоритми глибокого навчання. Розглядається використання байєсівських мереж у прогнозуванні діяльності, а їхні результати пояснюються за допомогою методів внутрішньої пояснюваності.

Лінійні моделі піддатливі для інтерпретації. Причина полягає в їхній здатності відтворювати монотонні лінійні залежності між функціями та цілями. Коефіцієнти лінійної моделі можна впорядкувати за важливістю ознак для загального узагальнення всієї моделі, за винятком випадків, коли існує значна взаємодія між ознаками.

Деревоподібні мають високу ступінь інтерпретації, оскільки вони апроксимуються до людської логіки та можуть враховувати нелінійні зв'язки між функціями та цілями, навіть при взаємодії функцій між собою. Дерева рішень визначають важливість ознак на основі критеріїв домішок, таких як ентропія або індекс Джині. Аналогічно до дерев рішень, інші ансамблеві моделі на основі дерев, такі як Random Forest, можуть також надавати рейтинг важливості ознак. Відносний ранг (глибина) функції може служити оцінкою її важливості як вузла прийняття рішень у дереві.

Спеціальні методи пояснюваності є внутрішньо інтерпретованими; проте, зазвичай вони не можуть забезпечити високу точність або мінімальну похибку для більшості практичних проблем реального світу. Таким чином, підхід спеціальної пояснюваності нещодавно привернув велику увагу. У методах спеціальної пояснюваності вони застосовуються або доповнюють модель чорного ящика під час навчання. Дослідники пропонували різні спеціальні методи пояснюваності для підвищення прозорості чорного ящика, які можна розділити на додавання обмежень інтерпретації, пояснювані гібридні моделі, спільне передбачення та пояснення, архітектурні коригування та потік інформації.

Додавання обмежень інтерпретації використовується для ідентифікації раку шкіри та виявлення захворювань рослин. У цьому випадку процедури вибору функцій і очищення застосовуються в шарах для забезпечення значущого пояснення прийняття рішення моделі. Зрозумілість моделі може бути покращена за допомогою різних методів регуляризації. Цей метод забезпечує більшу зрозумілість без втрати прогнозної ефективності.

2.2 Опис динаміки вподобань користувачів з використанням темпоральних правил

Вдосконалений метод пояснень ґрунтується на використанні темпоральної динаміки вибору споживача та рейтингу при побудові пояснень у рекомендаційних системах.

Системи рекомендацій створені для надання персоналізованого списку товарів чи послуг, які відповідають інтересам конкретного споживача. Дані системи роблять рекомендації на підставі подібності тем або аналогічних

уподобань інших користувачів. Вхідними даними для роботи при розробці рекомендацій є інформація про продажі товарів, їхні рейтинги, які надали користувачі. РС спрощують вибір користувача, що в свою чергу сприяє збільшенню обсягів продажів товарів. Таким чином, ці системи рекомендацій широко використовуються у сфері продажу товарів, рекомендацій для конференцій, перегляду фільмів, надання туристичних послуг та інших галузях.

В процесі роботи РС виникають проблеми, що інформації недостатньо на вході та вона не інформативна, спотворена проблема неповної інформації або проблема «холодного старту» виникає при появі нових, нерегулярних користувачів РС. Нові «холодні» користувачі ще не виконали свій вибір, а система не знає точно їхніх вподобань, тому не може дати чіткі рекомендації. У таких випадках система пропонує найчастіше просто популярні товари.

Також існує проблема надання хибних вхідних даних або ще її називають проблема шилінг атак, коли зловмисник навмисно штучно знижує або підвищує рейтинг деяких предметів. Якщо рейтинги піднімаються, то продажі збільшуються, і при зниженні падають продажі у конкурентів. Такі показники зміни рейтингів будуть показувати неактуальну інформацію у РС, що призведе до збою і таким чином у користувача зникне довіра до системи.

Щоб вирішити проблему холодного пуску застосовуються гібридні методи, які комбінують спільну фільтрацію та контекстний аналіз даних. Гібридні підходи також враховують зміни вимог користувача з часом, тобто темпоральну динаміку його вподобань. У цьому випадку застосовуються тимчасові обмеження [15], багаточаровий граф [8], часовий граф, модель нейронної мережі.

Що стосується методів по виявленню шилінг атак, то вони спрямовані на виявленні шаблону зловмисника, який атакує, і працюють на основі методів машинного навчання та статистичних та використовують його часову динаміку переваг.

Загальною особливістю вищесказаних підходів є те, що вони сортують товари для "холодного" споживача, виявляють спотворення рейтингів та все це виконують без активної участі користувача. Такі підходи не є корисними, вони можуть знизити довіру користувача до РС та призвести до відмови у використанні системи рекомендацій.

Отже, щоб підняти довіру користувачів використовується концепція екранів пояснення рекомендацій. Система пропонує ряд критеріїв для оцінки пояснень, які викликають довіру користувача до системи рекомендацій. Таким чином завдяки поясненням змінюються показники продажу товарів, та змінюється сама зручність у використанні самих рекомендацій для споживача. По результатам опитування споживачів та повторного їх відбору можна сказати, що після впровадження пояснень знизилася відмова користувачів при виборі низки товарів.

Узагальнюючи вищесказане, існуючі методи не враховують зміни уподобань з часом до побудови пояснень, і тому важливо розробити необхідні рекомендації які будуть враховувати тимчасову динаміку у вимогах споживачів, особливо, коли відсутня інформація стосовно нових користувачів та можливих спотвореннях рейтингів у результаті шилінг-атак.

Головною метою цього дослідження є розробка методу побудови пояснень для систем рекомендацій, який буде враховувати тимчасові залежності та показувати зміну в інтересах споживача щодо процесів покупки товарів впродовж обраного періоду часу. Це доповнення до рекомендацій з поясненнями спрямоване на підвищенні ефективності самої РС, яке можна оцінити за допомогою достовірності даних та їх ефективності.

Тобто ефективність залежить від кількості проданих товарів з поясненнями та без них, а достовірність відображає вплив пояснень на вибір споживача при купівлі товару (пояснення повинні направити споживача купити більш дорогій товар). Таким чином зріст ефективності РС з поясненнями проходить за рахунок

підвищення продажів загалом та з акцентом на продаж більш дорогих та модних товарів, через довіру до системи рекомендацій. Споживач повинен зробити усвідомлений вибір товарів спираючись на часову динаміку попиту.

Метод використовує адаптовані часові правила, кожне з яких встановлює часовий порядок. Наприклад візьмемо 2 продукти, вага правил для них буде залежати від ймовірності їх використання, наскільки вони популярні у відповідності з описом правила. Набір правил дає можливість описати порядок дій в часі при виборі товарів, які описані цим правилом. Таку послідовність ми можемо побачити при виборі товару споживачем, його причинно-наслідкові зв'язки.

Для розв'язання задачі побудови рекомендацій правила адаптуються шляхом призначення їм ваг, враховуючи не лише популярність конкретного товару, а й зміни у кількості проданих одиниць у системі електронної комерції.

Даний методу застосовує часові правила F , так як часовий оператор типу F може підібрати пару будь яких моментів в один часовий потік, то комбінація цих правил дає змогу описати зміну продажів усіх моментів, які відносяться до даного інтервалу часу K .

Загальний вид правила представлений формулою 2.1.

$$n = n_m F^{(w)} n_s , \quad (2.1)$$

де n – само правило;

F – тип правила;

W – вага;

n_m та n_s кількість покупок в момент часу K_m та K_s .

При використанні правил у цьому методі можна побачити динаміку попиту товарів впродовж заданого інтервалу часу K . Тривалість K залежить від групи товарів та предметної галузі застосування де видно цикли попиту споживача

(зріст або спад). Наприклад цикл для товарів святкових складає раз на рік протягом 1-2 місяців, а більш технологічних – раз на рік (моб телефон).

Період K ділиться на послідовність інтервалів $K_1, K_2, K_3, \dots, K_m$. Інтервали K_m показують рівень деталізації часу (години, доби, неділі). На кожному інтервалі K_m розраховується своя кількість покупок товару. Тоді формується набір F часових правил n_{jm} , які пов'язують кожен інтервал K_m з кінцевим інтервалом K_M .

Вага всіх правил залежить від різниці покупок між інтервалами часу: чим більше змінилася кількість проданих товарів, тим більша вага правила.

На рисунку 2.2 представлений приклад часових правил для одного пункту.

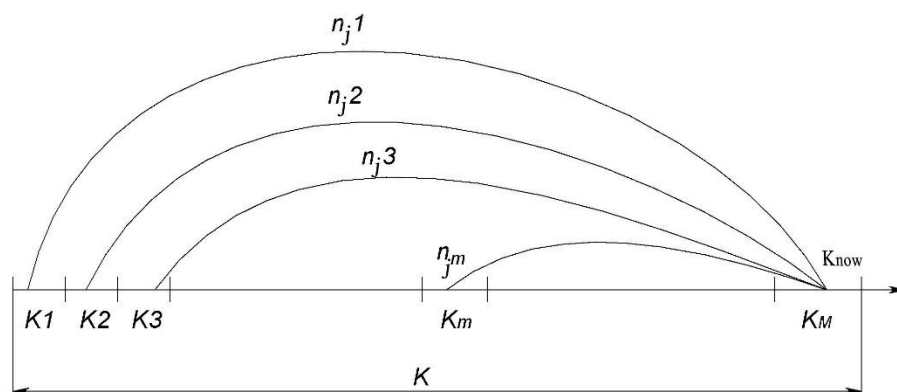


Рисунок 2.2 – Приклад розглянутих часових правил

Кожне правило n_{jm} описує збільшення або зменшення обсягів продажу товару i_j протягом пари часових інтервалів (K_m, K_M).

Загальний набір правил $N_j = \{n_{jm}\}$ описує зміну обсягу продажу товарів за певний проміжок часу K .

Вхідними даними є журнал продажів L , період часу K , рівень грануляції часу k , підмножина пунктів I рекомендації, для якої потрібно пояснення.

Записи в рядку u журналу U продажів являються такими параметрами, як користувач (u_k), елемент (i_j), кількість придбаних предметів (n_j) та позначка часу (K_n). Період часу K залежить від категорії товару, для яких формуються рекомендації.

Розглянемо етапи даного способу.

Етап 1. Вибір підмножини L_k подій для заданого періоду часу K та для заданого товару I .

Етап 2. Підготовка вхідних даних.

Крок 2.1 Визначення кількості продажів n_j по інтервалам часу K п .

$$L_{norm} = \left\{ (u_k, i_j, n'_j, \tau_n) : n'_j = \frac{n_j}{\max_j(n_j)} \right\}, \quad (2.2)$$

Крок 2.2 Нормалізація кількості продажів:

Етап 3. Побудова темпоральних правил.

$$N_j = \{n_{m,j} : \forall j n_{m,j} = (n_{jM} F n_{jM})\}, \quad (2.3)$$

Етап 4. Розрахунок ваг правил:

$$W_j = \{w_{m,j} : \forall m w_{m,j} = n_{jM} - n_{jm}\}, \quad (2.4)$$

Етап 5. Розрахунок сумарної ваги правил, як пояснення

У порівнянні з іншими підходами, мій удосконалений метод для проектів РС виконує більш повний огляд пояснень в галузі проектів рекомендацій. Саме вдосконалення включає в себе відбір товарів, що знаходяться в певному географічному регіоні, та враховуючи рейтинги товарів від реальних споживачів протягом конкретного часового інтервалу. Це сприяє збільшенню довіри користувача при виборі та купівлі товарів або послуг. Результатом праці є метод пояснень, який доповнює рекомендації числовою оцінкою темпоральної динаміки переваг користувачів за конкретний періоду часу та регіону.

3 ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ПОБУДОВИ ПОЯСНЕНЬ

3.1 Опис інформаційної технології

Після докладного аналізу алгоритмів та методів, які використовуються для надання пояснень у проектах рекомендаційних систем, було виявлено суттєвий недолік: всі ці методи не враховують зміни в уподобаннях користувачів з часом. Це стає особливо важливим, оскільки існуючі алгоритми не в змозі надавати обґрунтовані пояснення до рекомендацій в точний момент часу при зміні вподобань користувача. Проблема стає ще складнішою, коли відсутні вхідні дані щодо користувача системи. У зв'язку з цим прийшло рішення розробити метод, який буде незалежно аналізувати еволюцію вподобань споживача з часом і, використовуючи оцінку рейтингів, надає точні пояснення стосовно рекомендацій. Таким чином, було ухвалено рішення створити вдосконалений метод пояснень у проектах РС та інтегрувати його за допомогою програмного забезпечення.

Технологія використання удосконаленого методу побудови пояснень з урахуванням змін уподобань користувача з часом представлена поетапно на рисунку 3.1.

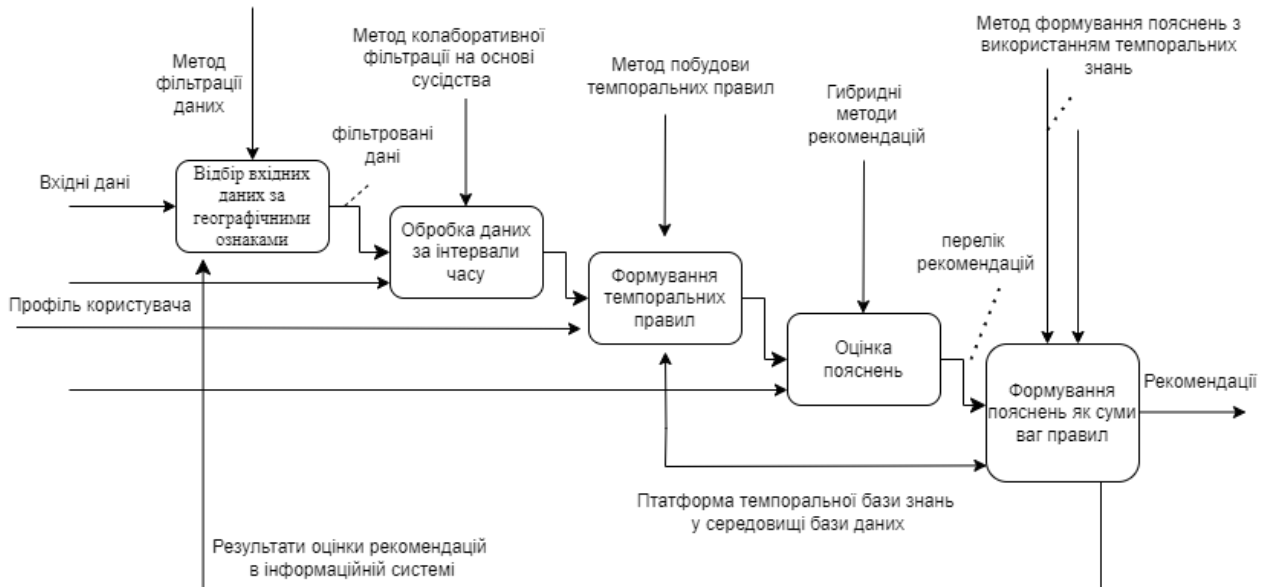


Рисунок 3.1 - Технологія проведення методу

Поетапна технологія виконання методу.

1. На першому етапі відбувається відбір підмножини подій для заданого періоду часу та для певного товару. Таким чином проводимо відбір товарів, які продавалися впродовж 1 місяця, враховуючи динаміку інтересів користувача за допомогою часових правил. Сукупність часових правил показує нам зростають або знижуються продажі у відповідному регіоні у порівнянні з теперішнім інтервалом часу.

2. На другому етапі проводиться підготовка вхідних даних, вилучення з логу конкретних товарів, які задовольняють нашим потребам.

2.1 Далі проводимо визначення кількості продажів по інтервалах дистрибуції часу. Відбираємо продажі за перші 3 дні кожного тижня впродовж місяця.

2.2 Проводимо нормалізацію кількості продаж, таким чином ми зможемо побачити їх кращу інтерпретацію шляхом приведення результатів до єдиного виду(діапазону).

3. На третьому етапі виконаємо побудову темпоральних правил. Визначимо періоди, коли певні товари найчастіше купуються або використовуються.

4. Виконаємо розрахунок ваг правил з урахуванням популярності товарів та різниці продажів між певними інтервалами часу. Показники популярності враховуємо за рахунок нормалізації на етапі 2.2.

5. Проведемо формування пояснень як суми ваг правил, згідно 4 етапу, враховуючи IP адресу кожного споживача зробимо відбір по географічному розташуванню.

Удосконалення методу формування пояснень при зміні вимог користувача проводиться шляхом врахування географічних характеристик, тобто відбору саме тих товарів, які знаходяться у певному географічному регіоні.

На рисунку 3.2 детально представлений принцип роботи удосконаленого підходу, який ми використали для нашої рекомендаційної системи товарів.



Рисунок 3.2 - Принцип удосконаленого підходу з використанням часових правил для РС товарів

3.2 Опис вхідних даних для інформаційної технології

Для експериментальної перевірки методу був використаний лог журналу продажів магазину товарів «Online Retail», який можна знайти у репозиторії UCI. Кожен запис у цьому журналі містить всю необхідну інформацію для побудови пояснення, включаючи код, назву та опис продаваного товару, а також мітку часу з датою та часом продажу. Частина логу представлена на рисунку 3.3.

95	536378	20725	LUNCH BAG RED RE	10	01.12.2010 9:37	1,65	14688	United Kingdom
96	536378	21559	STRAWBERRY LUNC	6	01.12.2010 9:37	2,55	14688	United Kingdom
97	536378	22352	LUNCH BOX WITH C	6	01.12.2010 9:37	2,55	14688	United Kingdom
98	536378	21212	PACK OF 72 RETROS	120	01.12.2010 9:37	0,42	14688	United Kingdom
99	536378	21975	PACK OF 60 DINOSA	24	01.12.2010 9:37	0,55	14688	United Kingdom
100	536378	21977	PACK OF 60 PINK PA	24	01.12.2010 9:37	0,55	14688	United Kingdom
101	536378	84991	60 TEATIME FAIRY C	24	01.12.2010 9:37	0,55	14688	United Kingdom
102	536378	84519A	TOMATO CHARLIE+L	6	01.12.2010 9:37	2,95	14688	United Kingdom
103	536378	85183B	CHARLIE & LOLA WA	48	01.12.2010 9:37	1,25	14688	United Kingdom
104	536378	85071B	RED CHARLIE+LOLA	96	01.12.2010 9:37	0,38	14688	United Kingdom
105	536378	21931	JUMBO STORAGE BA	10	01.12.2010 9:37	1,95	14688	United Kingdom
106	536378	21929	JUMBO BAG PINK VI	10	01.12.2010 9:37	1,95	14688	United Kingdom
107	536380	22961	JAM MAKING SET PR	24	01.12.2010 9:41	1,45	17809	United Kingdom
108	536381	22139	RETROSPOT TEA SET	23	01.12.2010 9:41	4,25	15311	United Kingdom
109	536381	84854	GIRLY PINK TOOL SE	5	01.12.2010 9:41	4,95	15311	United Kingdom
110	536381	22411	JUMBO SHOPPER VI	10	01.12.2010 9:41	1,95	15311	United Kingdom
111	536381	82567	AIRLINE LOUNGE,M	2	01.12.2010 9:41	2,1	15311	United Kingdom
112	536381	21672	WHITE SPOT RED CE	6	01.12.2010 9:41	1,25	15311	United Kingdom
113	536381	22774	RED DRAWER KNOB	24	01.12.2010 9:41	1,25	15311	United Kingdom
114	536381	22771	CLEAR DRAWER KNC	24	01.12.2010 9:41	1,25	15311	United Kingdom
115	536381	71270	PHOTO CLIP LINE	1	01.12.2010 9:41	1,25	15311	United Kingdom
116	536381	22262	FELT EGG COSY CHIC	1	01.12.2010 9:41	0,85	15311	United Kingdom
117	536381	22627	PICNIC BASKET RETRO	1	01.12.2010 9:41	2,55	15311	United Kingdom

Рисунок 3.3 – Вхідні дані для проведення розрахунків

Даний метод був реалізований за допомогою мови програмування Python. Для цього ми імпортували бібліотеки, таблиці журналу продажу та завантажили вихідні дані.

Розрахунок проводився у два етапи.

На першому етапі виконувалися роботи присвячені побудові пояснень, а на другому проводилося обґрунтування вибору рівня часової грануляції (провели розрахунки за 14 днів та протягом 2х місяців).

Під час першого експерименту період часу K становив 7 днів, а час грануляції рівня $k - 1$ день. У результаті першого етапу було відібрано 12 тис. записів продажу серед усіх товарів. Крім того, були відібрані вихідні дні. У ці 7 днів було 2 вихідних. Таким чином, дані про продажі були отримані за 5 днів. Початковий список рекомендованих елементів отримуємо за допомогою спільної фільтрації для випадково вибраного користувача. Даний набір складається із 7 елементів. Для п'яти предметів було сформовано часові залежності, а для інших рекомендованих товарів розпродажі відбувалися лише в один із обраних днів, що не дозволяло зробити такі залежності.

4 ЕКСПЕРИМЕНТАЛЬНА ПЕРЕВІРКА МЕТОДУ ПОБУДОВИ ПОЯСНЕНЬ

4.1 Програмна реалізація методу побудови пояснень

Система рекомендацій буде виконана у формі клієнт-серверного додатка. У складі серверної частини будуть впроваджені два модулі: один з яких призначений для обробки рейтингів товарів, а інший - для генерації рекомендацій на основі часових критеріїв. Клієнтська частина буде запитувати списки рекомендацій зі сторони серверної частини стосовно вподобань користувача і виводити результати на екран у вигляді списку рекомендацій.

Модель, що використовується у цьому проекті, є універсальною і забезпечує можливість внесення різноманітних варіантів та розширень для поліпшення та розширення функціональності системи.

Вхідними даними служать:

- записи журналу про кількість проданого товару за проміжок часу;
- відповідні записи про рейтинги товару;
- вибрані категорії товару;
- записи про користувачів даного ресурсу;
- записи про зроблений вибір користувачів.

Після ініціювання процесу збору даних додаток очікує на відповіді від бази даних та в подальшому проводить аналіз отриманих даних протягом певного періоду часу з урахуванням історії пошуку зареєстрованих користувачів в системі. За допомогою запиту додаток виділяє поточну сесію браузера клієнта у системі та зберігає інформацію про виконані дії.

```

##import data
data_set = pd.read_csv('C:/Users/Лилия/PycharmProjects/sales/OnlineRetail.csv', parse_date

# compute user rating matrix and timestamp matrix
num_user = 943 # user id 1 to 943
num_movie = 1682 # movie id 1 to 1682

user_rating_dict = {}
# key is user id : value are rating of all movie
for user_id in range(1, num_user + 1):
    user_rating_dict[user_id] = np.array([0] * num_movie)

user_timestamp_dict = {}
for user_id in range(1, num_user + 1):
    user_timestamp_dict[user_id] = np.array([0] * num_movie)

# append rating data set to user rating dict
data_set_list = data_set.tolist()
for each_row in data_set_list:
    user_id = each_row[0]
    mmovie_id = each_row[1]
    rating = each_row[2]
    movie_index = mmovie_id - 1

```

Рисунок 4.1 – Фрагмент програмного коду

Частина програмного коду відповідає за отримання даних про останні рейтинги товарів за ключовими словами та за певний період часу. Далі за допомогою команди Flask надаємо запит до БД рейтингу товарів по категорії за ключовими словами та отримавши відповідь проводить сортування та формує оцінки для виводу рейтингу товарів за ключовими словами. Швидкість та зручність використання додатку обумовлені легкістю написаного програмного коду та широкою можливістю його впровадження в сучасних веб-системах.

Під час обробки наша рекомендаційна система отримує всю необхідну інформацію, що дозволяє передбачити вподобання користувача та сформувати перелік товарів, які можуть бути запропоновані користувачеві як альтернатива або доповнення. Після попередньо сформованих рекомендацій наш веб-додаток входить в дію, використовуючи інформацію з бази даних про користувача (приклад наведений у таблиці 4.1) для надання пояснень щодо рекомендаційної системи. Структура програмного додатку зображена на рисунку 4.2.

Таблиця 4.1 – Приклад даних користувача

ID	Name	Age	Gender	Countru
14	Maric	35	M	Kanada
15	Asya	30	W	Ukrainsan
16	Sonay	32	W	Polish
17	Ivan	41	M	English

Де Id є унікальним ідентифікатором користувача, Name вказує на ім'я користувача, Age - вік, Gender - стать, Country - країна проживання.

Структура програмного додатку зображена на рисунку 4.2.

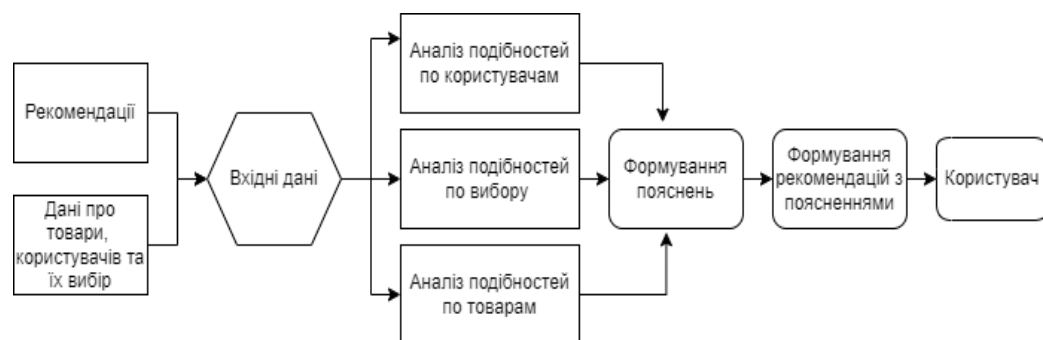


Рисунок 4.2 – Структура програмного додатку РС

Після активації алгоритму потрібно ввести початкові дані запиту від користувача. Далі відбувається обробка цих введених значень та формування рейтингу оцінок товару протягом визначеного періоду часу. Після аналізу отриманих даних запускається процес генерації рекомендацій для конкретного пошуку користувача. Далі на послідуєчому етапі починається надання пояснень до цих рекомендацій відповідно до методології роботи додатку. Після отримання користувачем рекомендацій та відповідних пояснень, настає період очікування

його реакції. Після взаємодії користувача з модулем рекомендацій та пояснень, вибір користувача запам'ятовується у його профілі, і алгоритм припиняє свою роботу. Якщо станеться відмова під час зв'язку з клієнтом, алгоритм повторює процес створення та подання пояснень.

У результаті виконання отримуємо список товарів, які можна рекомендувати новому користувачу. Ці товари були оцінені сусідніми користувачами зі схожими географічними характеристиками, мали доцільні пояснення та враховували ваги правил.

	Coef.	Std.Err.	t	P> t	[0.025	0.975]
const	50.0947	23.6149	2.1213	0.0449	1.2435	98.9459
DPI	0.6723	0.0793	8.4782	0.0000	0.5082	0.8363
ASSETS	0.1741	0.0587	2.9679	0.0069	0.0528	0.2954

Рисунок 4.3 – Фрагмент програмного коду

Таким чином виконано опрацювання логіки побудови пояснень у рекомендаціях для нових користувачів з урахуванням часових правил.

Результати розрахунків приведені у таблицях 4.1 та 4.2.

Таблиця 4.1 – Результати розрахунків методу побудови пояснень з урахуванням змін уподобань користувача

Дні	Щоденні продажі товарів n_{jm} у певному регіоні	Ваги правил W	Збільшення або зменшення попиту g_j	Використані рекомендації для пояснення
1	2	3	4	5

Кінець таблиці 4.1

1	2	3	4	5
1	15;17;28;13	0,02; -0,05; -0,35	-0,36	-
2	3;2;2;1;3	0,02; 0,06; 0,01; 0,07	0,16	+
3	7;5;24;4	0,56; 0,0; -0,65	-0,12	-
4	25;3;26;32	0,1; 0,9; 0,18	1,08	+
5	31;2;21;33	0,02; 0,7; 0,12	1,12	+

Виходячи з розрахунків, наведених у таблиці 4.1 бачимо, що в першому та третьому пункті ваги правил відображають тенденцію до припинення зростання продажів товарів, навіть з урахуванням зростання закупок в один із представлених днів. У другому та четвертому пунктах ми бачимо тенденцію зростання продажів і збільшення попиту на товари.

При виділенні рівня грануляції $k=1$ день виявилось, що встановити тимчасові правила на такий короткий термін не вдалося. Ця причина обумовлена тим, що ці товари були у продажу дуже рідко, наприклад один раз на місяць. Тому у наступному етапі досвіду для одного і того ж товару, який ми вибрали у переліку були обрані значення грануляції часу 14 днів та 2 місяці.

Таблиця 4.2 – Результати розрахунків методу побудови пояснень для різних варіантів часової грануляції

Рівень грануляції часу	Продажі товарів за період, n_{jm} у певному регіоні	Вага правил $w_{m,j}$	Збільшення або зменшення попиту g_j	Використані рекомендації для пояснення
1	2	3	4	5
14 днів	88; 21; 8; 2; 2; 3; 1; 2; 36; 45; 87; 154; 145; 112; 45; 11; 78; 87; 96; 254	-0,17; 0,09; 0,21; 0,26; -0,05; 0,01; 0,17; 0,18; -0,06; 0,1; 0,05; 0,43; 0,65; 0,32; 0,41; 0,89; 0,23; 0,51;	-3,14	-

Кінець таблиці 4.2

1	2	3	4	5
		0,32; -0,21		
2 місяці	121; 1; 2; 3; 66; 99; 356; 521; 485	0,89; 0,98; 0,99; 0,97; 0,88; 0,69; 0,54; -0,02; 0,01	7,45	+

Результати розрахунків згідно таблиці 4.2 показують, що рівень грануляції часу значно впливає на результуюче пояснення. Ми бачимо динаміку зростання показника n_{jM} протягом 2х місяців, що свідчить про тенденцію зростання даного продукту за рік в даному географічному регіоні.

4.2 Експериментальна перевірка удосконаленого методу

На першому етапі ми отримали дані про рейтинг товарів у категорії "Мобільні телефони", який складається з оцінок попередніх користувачів. Логічно припускати, що ми зацікавлені в продуктах з найвищим рейтингом за весь час їх існування. Таким чином, додаток сортує товари, використовуючи ідентифікатор рейтингу (rating) і, завдяки дизайну використаного екрану пояснень, відображає три списки товарів на сторінці. Рекомендації на екрані супроводжуються поясненнями, наприклад, "Попит на ці конкретні види товарів зріс на 28% в категорії мобільні телефони за тиждень". В результаті потенційний покупець, обираючи телефон, може звернути увагу на ті пункти, які були придбані раніше та взяли найвищі відгуки від користувачів. Таким чином можна прийняти своє рішення в сторону перевірених товарів. Приклад бачимо на рисунку 4.4, який показує екран пояснень із зображенням темпоральної залежності.



МОБІЛЬНИЙ ТЕЛЕФОН XIAOMI
REDMI 12 4/128GB MIDNIGHT BLACK

Ціна за одиницю:
1501\$

[купити](#)

ОПИС
Баланс між стилем та потужністю
Xiaomi Redmi 12 із преміальною скляною
панеллю зачаровує з першого погляду.
Він оснащений 6.79-дюймовим екраном FHD+,
на якому контент виглядає привабливо за
рахунок високої деталізації та великої
палітри кольорів. Продуктивність забезпечує
потужний чипсет MediaTek Helio G88.
Камера з потрійним об'єктивом має роздільну
здатність 50 Мп та підтримку штучного
інтелекту.

Попит на ці товари підвищився на 28%
в категорії мобільні телефони за тиждень

Сформоване пояснення

APPLE IPHONE 13
128GB

APPLE IPHONE 14
128GB

SAMSUNG GALAXY M34
5G

Рисунок 4.4 – Зразок моделі з темпоральною залежністю товару

На другому етапі ми отримали дані, які відображають рейтинги одиниць товарів у групі "Мобільні телефони", які були розраховані протягом наступного інтервалу часу. На зображенні 4.5 відображено пояснення до рекомендацій товарів у даній категорії, що вказує на те, що протягом цього періоду попит і рейтинги товарів, сформованих минулого тижня, зазнали змін і зросли на 50%. Це може бути зумовлено сезонністю формування пояснень, наприклад, перед Новорічними святами. Отже, користувач отримує інформацію про різке зростання попиту на ці товари та може відстежувати рейтинги оцінок від користувачів, які вже придбали чи замовили їх раніше. Ураховуючи це, надання пояснень із вказівкою на тимчасові зміни може позитивно вплинути на вибір кінцевого користувача. Детальне пояснення до рекомендацій стимулює цікавість споживача і підвищує ймовірність того, що споживач придбає товари, які були рекомендовані навіть за вищою ціною.



МОБІЛЬНИЙ ТЕЛЕФОН XIAOMI
REDMI 12 4/128GB MIDNIGHT BLACK

Ціна за одиницю:
1501\$

[Купити](#)

ОПИС
Баланс між стилем та потужністю
Xiaomi Redmi 12 із преміальною скляною
панеллю захарує з першого погляду.
Він оснащений 6.79-дюймовим екраном FHD+,
на якому контент виглядає привабливо за
рахунок високої деталізації та великої
палітри кольорів. Продуктивність забезпечує
потужний чипсет MediaTek Helio G88.
Камера з потрійним об'єктивом має роздільну
здатність 50 Мп та підтримку штучного
інтелекту.

Попит на ці товари підвищився на 50%
в категорії [мобільні телефони](#) за тиждень

Сформоване пояснення

		
APPLE IPHONE 13 128GB	APPLE IPHONE 14 128GB	SAMSUNG GALAXY M34

Рисунок 4.5 – Зразок моделі пояснення з урахуванням змін попиту
споживача

Таким чином на рисунку 4.6 представлено кінцевий результат роботи додатку, що відображений в РС проєктів. Діаграма відображає кількість замовлень у категорії "Мобільні телефони" без вживання розробленого додатку та в порівнянні з результатами його включення у рекомендаційну систему.

У такому випадку раціональним є враховувати темпоральну динаміку вподобань користувачів, яка визначається набором правил, відображаючи зміни у продажах цільового товару протягом двох неспільних часових інтервалів, наприклад, для двох різних днів. Метод включає в себе етапи створення невиважених темпоральних правил, розрахунок ваг для цих правил і формування пояснень у вигляді суми ваг для кожного предмета.

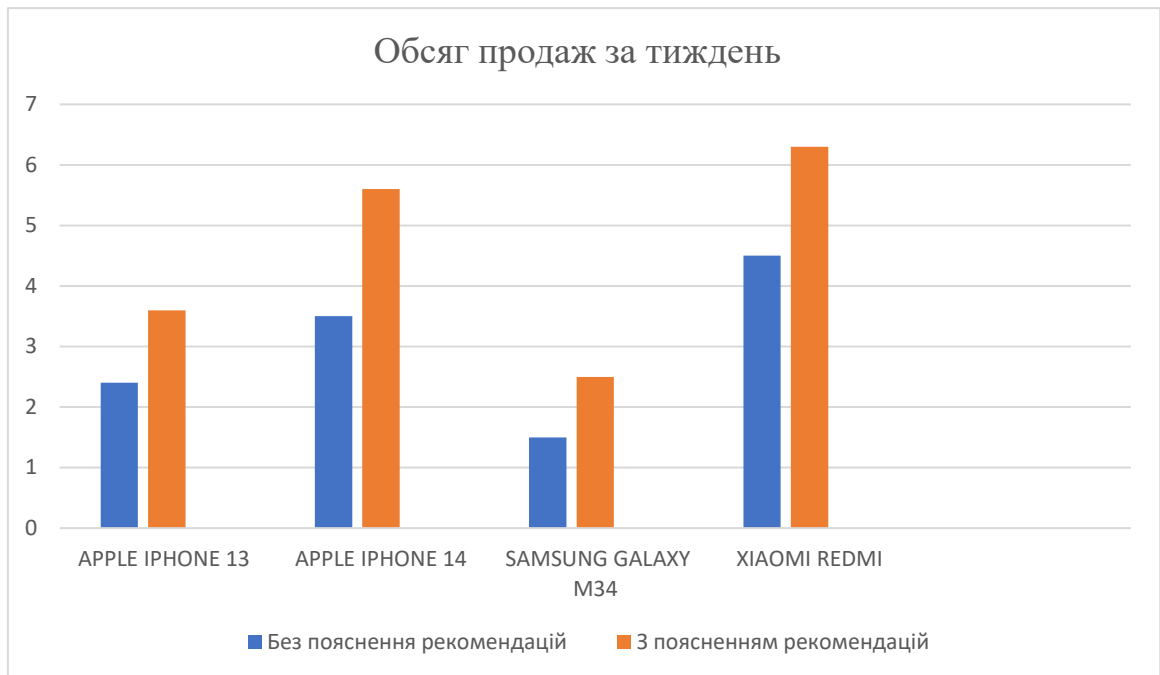


Рисунок 4.6 – Обсяг продаж мобільних телефонів за тиждень

Цей підхід відрізняється тим, що він надає пояснення у вигляді єдиного показника, який об'єднує зміни у продажах для кожного рекомендованого товару протягом визначеного періоду часу.

Запропонований метод сприяє підвищенню довіри користувача до рекомендацій завдяки очевидності використаного показника. Це в свою чергу може позитивно вплинути на ефективність роботи РС.

Як видно з результатів, кількість замовлень зросла внаслідок впровадження пояснень, в порівнянні з версією без пояснень. Користувачі краще розуміють цілі та мету, з якими надається рекомендація для їхнього поточного вибору. Оскільки кількість замовлень зросла, з графіку видно, що якість виконаних рекомендацій практично зросла вдвічі. Результат конверсії збільшився з 2,3% до 2,7. В наслідок цього можна зробити висновок, що введення пояснень до рекомендацій працює і призводить до позитивних результатів щодо користувача .

ВИСНОВКИ

У магістерській роботі проведено дослідження щодо проблеми формування пояснень, звертаючи увагу на зміни вимог користувачів з часом. Проаналізовано існуючі методи створення рекомендацій в інтелектуальних системах. Результат аналізу показує, що вживання пояснень сприяє збільшенню довіри користувачів до отриманих рекомендацій та підвищенню конверсії в електронній комерції(ЕК).

У магістерській кваліфікаційній роботі було досліджено проблему формування пояснень з урахуванням змін вимог користувачів з часом. Проаналізовано існуючі методи формування рекомендацій в інтелектуальних системах. Аналіз результатів вказує на те, що використання пояснень сприяє підвищенню довіри користувачів до отриманих рекомендацій і збільшує конверсію в системі електронної комерції.

Вивчення методу формування пояснень призвело до висновку про значущість врахування зміни інтересів споживача з часом та їхніх вподобань стосовно об'єктів інтересу. У ході дослідження було вдосконалено підхід до забезпечення пояснень у проектах рекомендаційних систем за допомогою врахування рейтингів і оцінок, а також врахування темпоральної динаміки споживачів. Дослідження процесу формування пояснень призвело до висновку про важливість врахування зміни інтересів користувачів з часом, їхні переваги щодо об'єктів інтересу. У роботі був удосконалений метод надання пояснень у проектах рекомендаційних систем за допомогою введення врахування рейтингів та зміни темпоральної динаміки споживачів.

Для реалізації дослідження було розроблено клієнт-серверний додаток, який реалізовує два модулі: модуль обробки оцінки рейтингів товарів та модуль

формування рекомендацій, враховуючи часові норми та надає пояснення до рекомендаційної системи.

Було впроваджено модуль створення пояснень, який оцінює вхідні дані щодо користувачів, товарів і інформації про продажі. Додаток генерує пояснення у формі списку товарів, впорядкованих за темпоральною динамікою.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Методичні вказівки до передатестаційної практики для студентів усіх форм навчання спеціальності 122 – Комп’ютерні науки, освітньо-професійної програми "Інформаційні управляючі системи та технології" / Упоряд.: Чалий С.Ф., Євланов М. В., Чала О. В. - Харків: ХНУРЕ, 2021

2. Chalyi S., Leshchynskiy V. Knowledge Representation in the Recommendation System Based on the White Box Principle. Сучасні інформаційні системи. 2019. Т. 3, № 3 С. 82-86.

3. Jones M. Recommender systems, Part 1. Introduction to approaches and algorithms. Learn about the concepts that underlie web recommendation engines / M. Jones – 2013. [Електронний ресурс] – Режим доступу: https://www.ibm.com/developerworks/opensource/library/osrecommender1/index.html?s_tact=105agx99&s_cmp=cp

4. Xiaoyuan Su and Taghi M. Khoshgoftaar A Survey of Collaborative Filtering Techniques A Survey of Collaborative Filtering Techniques // Hindawi Publishing Corporation, Advances in Artificial Intelligence archive, USA : журнал. — 2009. — P. 1 - 19.

5. Меньшикова Н.В. Обзор рекомендательных систем и возможностей учета контекста при формировании индивидуальных рекомендаций / Н. В. Меньшикова, И.В. Портнов, И.Е. Николаев. // АCADEMY. – 2016. – №6. – с. 20–22.

6. Пономарев А. В. Обзор методов учета контекста в системах коллаборативной фильтрации // Труды СПИИРАН, 2013. № 7 (30), С. 169-188.

7. Мелешко Є.В. Дослідження методів побудови рекомендаційних систем в мережі Інтернет / Є.В. Мелешко, Г.С. Семенов В.Д. Хох. // Збірник наукових

праць "Системи управління, навігації та зв'язку". Випуск 1(47). – Полтава: ПНТУ ім. Ю. Кондратюка. – 2018. – С. 131-136

8. Chalyi S., Levykin I., Biziuk A., Vovk A., Bogatov I. Development of the technology for changing the sequence of access to shared resources of business processes for process management support. *Eastern-European Journal of Enterprise Technologies*. 2020. P. 22-29.

9. Xiang L., Yuan Q. Temporal Recommendation on Graphs via Long-and Short-term Preference Fusion. *KDD'10 of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2010,723-732.

10. Чалий С.Ф., Лещинський В.О., Лещинська І.О., Концепція формування пояснень в рекомендаційних системах за принципом білого ящика. *Збірник наукових праць*. – Полтава: ПНТУ, 2019. – Т. 3 (55). – С. 156-160.

11. Gunes I., Kaleli C., Bilge A., Polat H. Shilling attacks against recommender systems: a comprehensive survey. *Artificial Intelligence Review*, 2014. 42, 767–799.

12. Chala O., Novikova L., Chernyshova L., Method for detecting shilling attacks in e-commerce systems using weighted temporal rules. *EUREKA: Physics and Engineering*. 2019. Vol. 5, 29-36. DOI: 10.21303/2461-4262.2019.00983

13. Tintarev N., Masthoff J. A survey of explanations in recommender systems. *23rd International Conference on Data Engineering Workshop*, 2007, 801– 810.

14. Tintarev N., Masthoff J. Designing and evaluating explanations for recommender systems, in Ricci, F. Rokach, L. Shapira, B. and Kantor, P. (Eds.), *Recommender systems handbook*. Springer, Dordrecht, 2010, 479–510.

15. Чалий С. Ф. Моделювання пояснень щодо рекомендованого переліку об'єктів з урахуванням темпорального аспекту вибору користувача / Чалий С. Ф., Лещинський В. О., Лещинська І. О. // *Системи управління, навігації та зв'язку*. 2019. Т. 6. № 58. С. 97-101.

16. Чалий С. Ф. Ситуаційна модель користувацького вибору в рекомендаційній системі / С. Ф. Чалий, І. Б. Прібильнова // *Системи управління,*

навігації та зв'язку. - 2019. - Вип. 2. - С. 159-163. - Режим доступу: http://nbuv.gov.ua/UJRN/suntz_2019_2_34.

17. Чалий С.Ф., Лещинський В.О., Лещинська І.О. (2018). Моделювання контексту в рекомендаційних системах. Проблеми інформаційних технологій, 1(023), 21-26.

18. Chalyi, S., Leshchynskyi, V., & Leshchynska, I. (2019). Концепція формування пояснень в рекомендаційних системах за принципом білого ящика. Системи управління, навігації та зв'язку. Збірник наукових праць, 3(55), 156-160.

19. Types of Recommendation Systems & Their Use Cases [Електронний ресурс] // Types of Recommendation Systems & Their Use Cases - Режим доступу: <https://medium.com/mlearning-ai/what-are-the-types-of-recommendation-systems-3487cbafa7c9>, вільний.

20. Matrix Factorization For Recommendation Systems [Електронний ресурс] // Matrix Factorization For Recommendation Systems - Режим доступу: <https://medium.com/@melih.kacaman/matrix-factorization-for-recommendation-systems-284f95c79ef6>, вільний.

21. Документація Weaviate [Електронний ресурс] // Документація Weaviate - Режим доступу: <https://weaviate.io/developers/weaviate/quickstar>, вільний.

22. DARPA 's explainable AI (XAI) program: A retrospective [Електронний ресурс] // DARPA 's explainable AI (XAI) program: A retrospective - Режим доступу: https://www.researchgate.net/publication/356781652_DARPA_'s_explainable_AI_XAI_program_A_retrospective, вільний.

23. Linden G. Amazon.com recommendations: Item-to-item collaborative filtering / G. Linden, B. Smith, J. York // IEEE Internet Computing, vol. 7, no. 1, pp. 76–80, 2003.

24. Chalyi S. Доповнення вхідних даних рекомендаційної системи в ситуації циклічного холодного старту з використанням темпоральних обмежень

типу «next» / S. Chalyi, V. Leshchynskyi, I. Leshchynska // Системи управління, навігації та зв'язку. Збірник наукових праць. – Полтава: ПНТУ, 2019. – Т. 4 (56). – С. 105-109. – doi:<https://doi.org/10.26906/SUNZ.2019.4.105>.

25. Melville P., Sindhvani V. Recommender systems. Encyclopedia of Machine Learning. 2010. p. 30

26. Chala O. Models of temporal dependencies for a probabilistic knowledge base. ECONTECHMOD: An International Quarterly Journal on Economics of Technology and Modelling Processes. 2018. Polska Akademia Nauk. Oddział w Lublinie PAN. Т. 7. № 3. С. 53-58.

27. Чала О.В. Поліпшення процесів системи управління якістю на основі коригувальних і попереджувальних дій. Вісник економіки транспорту і промисловості: Зб. наук. праць. Харків. 2006. Вип. 15–16. С. 118-121.

28. Методичні вказівки щодо розробки та оформлення кваліфікаційної роботи (для студентів усіх форм навчання другого (магістерського) рівня програми "Інформаційні управляючі системи та технології") / Упоряд.:Петров К.Е., Левикін В.М., Чалий С.Ф., Євланов М.В., Саєнко В.І., Міхнов Д.К., Міхнова А.В., Чала О.В. – Харків: ХНУРЕ, 2021. – 30с.

29. ДСТУ 3008:2015. Інформація та документація. Звіти у сфері науки і техніки. Структура і правила оформлювання. – Чинний від 22.06.2015. – Київ: ДП «УкрНДНЦ», 2016. – 31 с.

30. ДСТУ 8302:2015. Інформація та документація. Бібліографічні посилання. Загальні положення та правила складання. – Чинний від 04.03.2016. – Київ: ДП «УкрНДНЦ», 2016. – 20 с.