

ДОДАТОК А
ГРАФІЧНИЙ МАТЕРІАЛ

Плакат 1

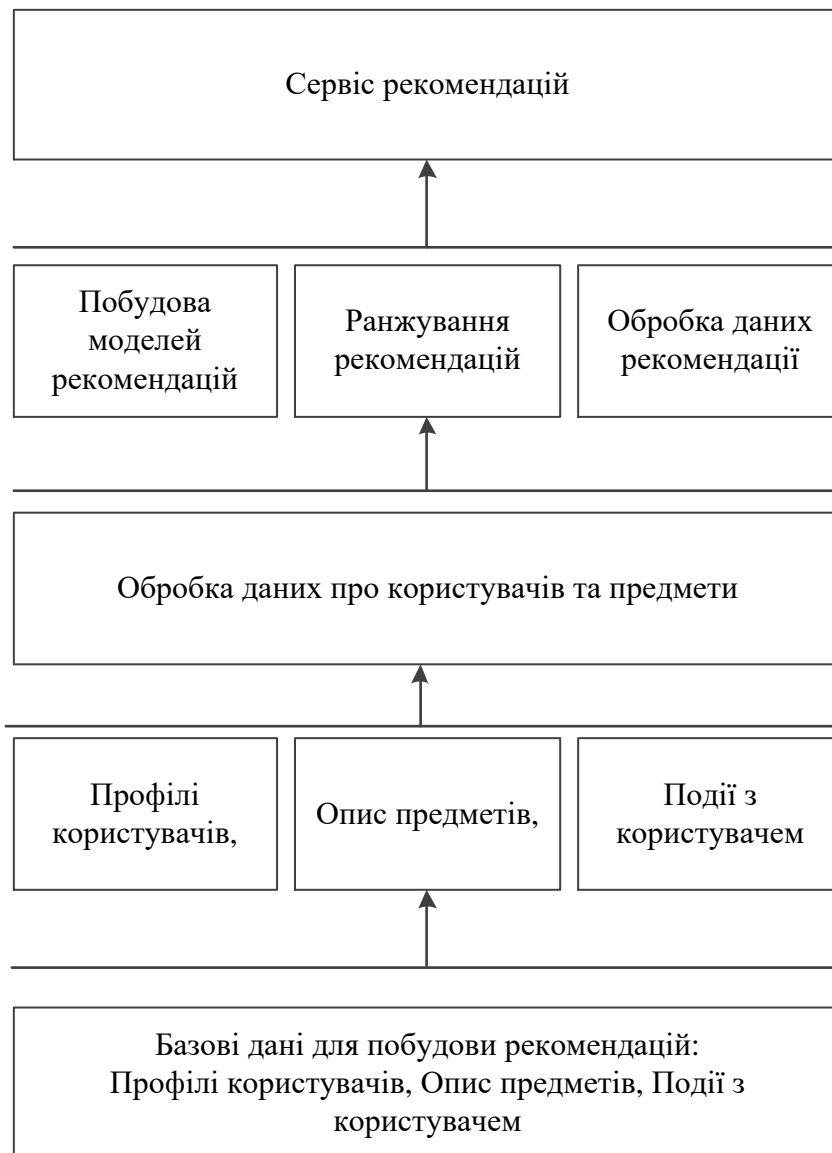
ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність	Існуючі гібридні методи побудови рекомендацій, як правило, не враховують зміни і вимогах та потребах користувачів з часом. Тому такі методи використовуються в офлайн-режимі. Однак це створює проблеми при побудові рекомендацій в режимі онлайн. Такі рекомендації мають менші вимоги до точності, але вони мають бути швидко адаптовані до поточних потреб користувача.
Об'єкт дослідження	Процес формування рекомендацій у ІТ- проектах рекомендаційних систем
Предмет дослідження	Методи формування рекомендацій пояснень у проектах рекомендаційних систем
Мета роботи	Дослідження гібридних методів формування рекомендацій для побудови рекомендацій в онлайн-режимі
Наукова новизна	Удосконалено гібридний метод побудови рекомендацій на основі попередньої фільтрації даних щодо змін вибору користувачів. Метод додатково враховує найбільш популярні товари, що дає можливість формувати рекомендації в режимі онлайн для нових користувачів.
Практичні результати	Побудова рекомендацій з урахуванням змін вибору користувачів
Задачі дослідження	<ul style="list-style-type: none"> – дослідження характеристик ІТ- проектів рекомендаційних систем; – аналіз методів побудови рекомендацій; – дослідження проблем побудови рекомендацій при використанні методів фільтрації; – дослідження гібридних методів побудови рекомендацій; – удосконалення методу гібридного методу формування рекомендацій для побудови рекомендацій в режимі онлайн; – планування проекту розробки гібридного методу побудови рекомендацій; – експериментальна перевірка удосконаленого гібридного методу побудови рекомендацій в режимі онлайн у проектах рекомендаційних систем.

ОБ'ЄКТ ДОСЛІДЖЕННЯ

Плакат 2

Рекомендаційні системи використовуються для інформування користувача про товар, який відповідає його потребам в даний момент часу.



ПРОЦЕС ПОБУДОВИ РЕКОМЕНДАЦІЙ



Рекомендаційна система використовує два базових сценарії побудови рекомендацій:

- рекомендації на основі запитів користувача;
- потокові рекомендації на основі властивостей користувача та відповідного елемента (продукту, інформації).

ПОРІВНЯЛЬНА ХАРАКТЕРИСТИКА МЕТОДІВ ПОБУДОВИ РЕКОМЕНДАЦІЙ

Метод	Вхідні дані	Особливості
Колаборативна фільтрація	Рейтинги товарів, надані користувачами	Рекомендації на основі рейтингів схожих користувачів
Контентна фільтрація	Рейтинги товарів, надані користувачами	Рекомендація на основі рейтингів схожих предметів
Демографічна фільтрація	Демографічна інформація щодо користувачів	Відбираються користувачі, які схожі за демографічними ознаками
Побудова рекомендацій на з використанням знань	Знання щодо того, як вибрані предмети відповідають потребам користувача	Визначається зв'язок між предметами та потребами користувача

ПРОБЛЕМИ ПОБУДОВИ РЕКОМЕНДАЦІЙ, ПОВ'ЯЗАНІ ІЗ НЕПОВНОТОЮ І НЕТОЧНІСТЮ ВХІДНИХ ДАНИХ

Назва	Опис
1. Проблема рейтингів	Більшість користувачів не ставлять жодних оцінок товарам. Тому виникає проблема: як дізнатися, чи задоволені вони продуктом і наскільки.
2. Проблема холодного старту 2.1. Проблема холодного запуску користувача 2.2. Проблема холодного запуску товару 2.3. Проблема холодного запуску рекомендаційної системи	В рекомендаційній системі не знайдено інформації про користувача або об'єкт. Тому неможливо надати рекомендації, порівнюючи вибір користувачів або властивості предметів.
3. Проблема розрідженості вхідних даних	Розрідженість рейтингової матриці, оскільки користувачі не виставляють оцінки товарам. Приводить до неточних рекомендацій при використанні колаборативної фільтрації.
4. Проблема масштабованості	Сповільнення роботи рекомендаційної системи, коли кількість користувачів суттєво збільшується.
5. Проблема надмірної спеціалізації	Виникає, якщо рекомендовані об'єкти занадто схожі один на одного. Користувач може перейти до іншого сайту з більш цікавими пропозиціями.

ПРОБЛЕМИ ПОБУДОВИ РЕКОМЕНДАЦІЙ

Назва	Опис
7. Проблема зміни даних з часом	Старіння інформації щодо товарів в рекомендаційній системі приводить до неточних рекомендацій
8. Проблема зміни налаштувань користувача	Уподобання користувача змінюються з часом, що приводить до неперсоналізованих рекомендацій
9. Проблема непередбачуваних предметів	Не формуються рекомендації для нових товарів, для яких не вказано категорії з історією перегляду.

ГІБРИДНІ МЕТОДИ ПОБУДОВИ РЕКОМЕНДАЦІЙ

Метод	Опис методу
Зважений	Оцінюються та комбінуються результати кількох рекомендаційних методів для кінцевої рекомендації
Перемикання	Рекомендаційна система використовує один з методів побудови рекомендацій в залежності від поточної ситуації
Змішаний	Одночасно представляються рекомендації, які побудовані кількома різними методами
Комбінація ознак	Використовуються ознаки рекомендаційних даних, що отримані із різних джерел
Каскадний	Один метод уточнює рекомендації, які дав інший метод.
Нарощування ознак	Вихід одного методу використовується як вхід для іншого методу.
Використання метарівня	Модель рекомендацій використовується як вхідні дані для методу побудови рекомендацій

КОМБІНУВАННЯ БАЗОВИХ МЕТОДІВ ПОБУДОВИ РЕКОМЕНДАЦІЙ В ГІБРИДНИХ МЕТОДАХ

Гібридні методи побудови рекомендацій	КФ	КФ	КФ	КН	КН	КН	ДМ	ДМ	ДМ	КБ	КБ	КБ
	КН	ДМ	КБ	КФ	ДМ	КБ	КФ	КН	КБ	КФ	КН	ДМ
Зважений	+	+	+		+	+			+			
Перемикання	+	+	+		+	+			+			
Змішаний	+	+	+		+	+			+			
Комбінація ознак	+	+	Ні		+	Ні			Ні			
Каскадний	+	+	+	+	+	+	+	+	+	+	+	+
Нарощування ознак	+	+	+	+	+	+	+	+	+	+	+	+
Використання метарівня	+	Ні	+	+	Ні	+	Ні	+	+	+	+	Ні

КФ – Колаборативна фільтрація;

КН – фільтрація на основі контенту;

ДМ – демографічний метод;

КБ – побудова рекомендацій на основі використання баз знань.

ПРОБЛЕМА ДОСЛІДЖЕННЯ

Проведений аналіз показав, що існуючі гібридні методи комбінують традиційні методи побудови рекомендацій. Побудова рекомендацій виконується в дві фази.

На першій фазі відбувається попередній відбір рекомендованих даних, а на другій – уточняється список предметів, які будуть представлені користувачеві.

Однак існуючі методи, як правило, не враховують зміни і вимогах та потребах користувачів з часом.

Це створює проблеми при побудові рекомендацій в режимі онлайн. Такі рекомендації мають менші вимоги до точності, але вони мають бути швидко адаптовані до поточних потреб користувача.

Тому при удосконаленні гібридного методу побудови рекомендацій для роботи в режимі онлайн доцільно враховувати зміни у виборі користувачів з часом.

УДОСКОНАЛЕНИЙ ГІБРИДНИЙ МЕТОД

Фаза 1. Відбір вхідних записів про вибір користувача, що враховують зміни у часі

Фаза 2. Колаборативна фільтрація для відібраної підмножини вхідних записів

Особливості методу.

Вхідні дані: інформація про продажі та рейтинги з врахуванням часу

На першій фазі виконуються такі дії:

– дані розбиваються на підмножини по інтервалам часу;

– з кожної підмножини відбираються дані про вибір схожих користувачів;

схожими є користувачі, які:

– вибрали ті ж товари, що й цільовий користувач;

– вибрали в тій же або більше кількості, що й цільовий користувач;

– відбір виконується спочатку на останньому інтервалі, а потім на попередніх інтервалах часу;

– з множини схожих товарів відбираються найбільш популярні товари;

За результатами першої фази формуються вхідні дані для другої фази.

На другій фазі формуються рекомендації методом колаборативної фільтрації.

ЕТАПИ УДОСКОНАЛЕНОГО ГІБРИДНОГО МЕТОДУ

Етап 1. Розбиття множини вхідних даних Dt на впорядковану послідовність підмножин Dt^k . Кожна елемент Dt_s^k містить мітку часу t_s^k з інтервалу T^k :

$$Dt = \left\langle Dt^k = \{Dt_s^k\} : (\forall k) t_s^k \in T^k, k = \overline{1, K} \right\rangle. \quad (1)$$

Результатом цього етапу можуть бути, наприклад, підмножини записів про продаж за кожен день (за годину, за тиждень).

Етап 2. Відбір даних із підмножини Dt^K за умовою зміни інтересів користувачів. Запит на рекомендації від цільового α – користувача надходить на останньому відомому інтервалі часу T^K . Підмножина Dt^K містить найактуальніші дані, що виникли на інтервалі запиту рекомендацій, що є ключовим для онлайн-рекомендацій. Результатом даного етапу є підмножина записів з Dt^K , яка містить інформацію про вибір користувачів, схожих на вибір користувача Us_α^K , на останньому інтервалі. Схожість полягає в тому, що користувачі Us_i^K вибирають стільки ж або на 1 більше екземплярів об'єкта It_j^K , що і цільовий користувач Us_α^K .

Етап 3. Відбір підмножин даних Dt^{k*} з Dt^k , для схожих користувачів. Для таких користувачів виконується умова $\exists Us_i^K \in G_\alpha^K$ належності користувача Us_i^K до графу продажів для цільового користувача :

$$Dt^* = \left\langle Dt^{k*} = \{Dt_s^k\} : (\forall i \forall k) \exists Us_i^K \in G_\alpha^K \right\rangle. \quad (2)$$

Результатом даного етапу є підмножина вхідних даних Dt^* , який містить записи про вибір користувачів, схожих на α – користувача.

ЕТАПИ УДОСКОНАЛЕНОГО ГІБРИДНОГО МЕТОДУ

Етап 4. Відбір із підмножини Dt^* найбільш популярних даних на кількох послідовних інтервалах часу T^k .

$$Dt^{**} = \left\langle \{Dt^{k*}\} : (\forall k) S_{i,j}^{k*} = \max_i (S_{i,j}^k) \right\rangle. \quad (3)$$

Результатом даного етапу є підмножина вхідних даних Dt^{**} , який містить записи про вибір схожими користувачами об'єктів, популярних протягом кількох послідовних інтервалів часу T^k , включаючи інтервал T^K .

Етап 5. Формування рекомендацій за допомогою колаборативної фільтрації на основі підмножини вихідних даних Dt^{**} .

Результатом цього етапу є впорядкований набір рекомендованих об'єктів, який враховує інтереси цільового користувача, і навіть найбільш популярні нині об'єкти схожих користувачів.

ПРОЕКТ РОЗРОБКИ МЕТОДУ ГІБРИДНОЇ ФІЛЬТРАЦІЇ

Форма використання задач

Название задачи	Трудозатраты	Длительность	Начало	Окончание
[-] Постановка задачі	8 ч	1 день	Пн 08.11.21	Пн 08.11.21
<i>Менеджер проекту</i>	8 ч		Пн 08.11.21	Пн 08.11.21
[-] Написання ТЗ	48 ч	6 днів	Вт 09.11.21	Вт 16.11.21
<i>Менеджер проекту</i>	48 ч		Вт 09.11.21	Вт 16.11.21
[-] Розробка макету	24 ч	3 днів	Ср 17.11.21	Пт 19.11.21
<i>Програміст</i>	24 ч		Ср 17.11.21	Пт 19.11.21
[-] Розробка модулю оцінк	32 ч	4 днів	Пн 22.11.21	Чт 25.11.21
<i>Програміст</i>	32 ч		Пн 22.11.21	Чт 25.11.21
[-] Тестування	16 ч	2 днів	Чт 25.11.21	Пт 26.11.21
<i>Тестувальник</i>	16 ч		Чт 25.11.21	Пт 26.11.21
[-] Розробка модулю гібри,	48 ч	6 днів	Пт 26.11.21	Пт 03.12.21
<i>Програміст</i>	48 ч		Пт 26.11.21	Пт 03.12.21
Тестування	0 ч	2 днів	Чт 02.12.21	Пт 03.12.21
[-] Інтеграція модулів	32 ч	2 днів	Вт 07.12.21	Ср 08.12.21
<i>Програміст</i>	16 ч		Вт 07.12.21	Ср 08.12.21
<i>Тестувальник</i>	16 ч		Вт 07.12.21	Ср 08.12.21
[-] Розробка звіту	8 ч	1 день	Пт 10.12.21	Пт 10.12.21
<i>Менеджер проекту</i>	8 ч		Пт 10.12.21	Пт 10.12.21

ЕКСПЕРИМЕНТАЛЬНА ПЕРЕВІРКА УДОСКОНАЛЕНОГО МЕТОДУ

Приклади рекомендацій

Код товару Опис товару

222963 JAM JAR WITH GREEN LID

023167 SMALL CERAMIC TOP STORAGE JAR

623293 SET OF 12 FAIRY CAKE BAKING CASES

722978 PANTRY ROLLING PIN

823295 SET OF 12 MINI LOAF BAKING CASES

922962 JAM JAR WITH PINK LID

123165 LARGE CERAMIC TOP STORAGE JAR

323294 SET OF 6 SNACK LOAF BAKING CASES

422980 PANTRY SCRUBBING BRUSH

523296 SET OF 6 TEA TIME BAKING CASES

Код товару Опис товару

023167 SMALL CERAMIC TOP STORAGE JAR

123165 LARGE CERAMIC TOP STORAGE JAR

422980 PANTRY SCRUBBING BRUSH

523296 SET OF 6 TEA TIME BAKING CASES

ЕКСПЕРИМЕНТАЛЬНА ПЕРЕВІРКА УДОСКОНАЛЕНОГО МЕТОДУ

Таблиця 4. 1 – Результати оцінки рекомендацій з урахуванням динаміки інтересів користувачів

Кількість рядків після відбору даних для колаборативної фільтрації, тис. записів	4	8	12
AUC	0,63	0,71	0,73

Таблиця 4.2 – Результати порівняльної оцінки рекомендацій

Кількість рядків, тис. записів	4	8	12
AUC	0,6	0,68	0,71

У разі обліку динаміки інтересів користувачів оцінки отриманих рекомендацій підвищуються

ВИСНОВКИ

Досліджено проблему побудови рекомендацій при зміні потреб користувачів.

Удосконалено гібридний метод побудови рекомендацій на основі попередньої фільтрації даних щодо змін вибору користувачів. Метод додатково враховує найбільш популярні товари, що дає можливість формувати рекомендації в режимі онлайн для нових користувачів.

Виконано експериментальну перевірку гібридного методу на наборі даних із оптових продажів товарів.