

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний університет «Запорізька політехніка»

Факультет комп'ютерних наук і технологій
Кафедра системного аналізу та обчислювальної математики

Пояснювальна записка
до дипломного проєкту (роботи)
магістра
(ступінь вищої освіти)

на тему: Модель рекомендаційної системи для прогнозування об'єму продажів власної
торгової марки аптечної мережі

A model of a recommendation system for forecasting the sales volume of a pharmacy's own
brand

Виконав: студент 2 курсу, групи КНТ 810 м

Спеціальності 124 – Системний аналіз
(код і найменування спеціальності)

Освітня програма (спеціалізація)
«Інтелектуальні технології та прийняття рішень в
складних системах»

Керівник Бакурова А.В. Супрун О. С.
(підпис та ініціали)
Рецензент Очеретін Д.В.
(підпис та ініціали)

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний університет «Запорізька політехніка»
(повне найменування закладу вищої освіти)

Інститут, факультет Факультет комп'ютерних наук і технологій
Кафедра Кафедра системного аналізу та обчислювальної математики
Ступінь вищої освіти магістр
Спеціальність 124 – Системний аналіз
(код і найменування)
Освітня програма (спеціалізація) «Інтелектуальні технології та прийняття рішень в складних системах»

ЗАТВЕРДЖУЮ
Завідувач кафедри Т.В. Корнієнко
« 11 » листопада 2021 року

ЗАВДАННЯ
НА ДИПЛОМНИЙ ПРОЄКТ (РОБОТУ) СТУДЕНТА

Супрун Олексія Сергійовича
(прізвище, ім'я, по батькові)

1. Тема проєкту (роботи) Модель рекомендаційної системи для прогнозування об'єму продажів власної торгової марки аптечної мережі

керівник проєкту (роботи) Бакурова А.В., д.е.н., професор,
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом закладу вищої освіти від «10» грудня 2021 року №490-к

2. Строк подання студентом проєкту (роботи) _____

3. Вихідні дані до проєкту (роботи) _____

Дані про продажі товарів в аптечній мережі

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити) _____

Дослідити сучасну методику прогнозування об'єму продажів при виході на ринок

ВТМ аптечної мережі

Виконати компаративний (порівняльний) аналіз існуючих алгоритмів

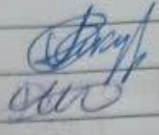
колаборативної фільтрації

Обґрунтувати вибір моделі для розробки рекомендаційної системи для аптечної мережі.

Виконати ряд експериментів з генерації рекомендацій на основі розробленої рекомендаційної системи на об'єм продажів при виході на ринок ВТМ та об'єм виробництва для подальшого планування діяльності мережі.

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)
не має

6. Консультанти розділів проєкту (роботи)

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	прийняв виконав завдання
1, 2, 3	Бакурова А.В. проф		
~~4, 5, 6, 7~~	~~Шулькевич О.В. ст. вик.~~		

7. Дата видачі завдання « 30 » вересня 2021 року.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проєкту (роботи)	Строк виконання етапів проєкту (роботи)	Примітки
1	Отримання завдання	30 вересня 2021	
2	Збір інформації та аналіз літератури	3 жовтня 2021	
3	Проведення огляду існуючих методів колаборативної фільтрації	10 жовтня 2021	
4	Реалізація алгоритмів колаборативної фільтрації та їх тестування.	15 жовтня 2021	
5	Реалізація алгоритмів колаборативної фільтрації та їх тестування.	26 жовтня 2021	
6	Оформлення дипломної роботи	30 жовтня 2021	
7	Отримання допуску до захисту та подача роботи в ДЕК	20 грудня	

Студент

(підпис)

Супрун О.С.

(прізвище та ініціали)

Керівник проєкту (роботи)

(підпис)

Бакурова А.В.

(прізвище та ініціали)

РЕФЕРАТ

ПЗ: 57 с., розділів 3, рис. 12 , табл. 18, джерел 21

Об'єкт дослідження – процес виведення на ринок товарів власної торгової марки централізованої аптечної мережі.

Предмет дослідження – моделі рекомендаційних систем для аналізу і прогнозування продажів товарів, зокрема для аптечних мереж.

Мета диплома: розробка рекомендаційної системи для аптечної мережі, що дозволить прогнозувати об'єм продажів при виході на ринок власної торгової марки (ВТМ) на основі даних про схожі товари та аналоги, які вже є на ринку

В дипломі проведено компаративний аналіз існуючих моделей рекомендаційних систем та програмного забезпечення для їхньої розробки. Обґрунтовано вибір моделі для розробки рекомендаційної системи для аптечної мережі та проведено ряд експериментів з генерації рекомендацій на основі розробленої рекомендаційної системи на об'єм продажів при виході на ринок ВТМ та об'єм виробництва для подальшого планування діяльності мережі.

Розрахунки виконано на основі даних мережі аптек, розраховано за існуючою методикою на підприємстві та за розробленою рекомендаційною системою, передбачається, що доробки рекомендаційної системи можна буде використати для покращення прогнозування. Науковою новизною роботи є модель рекомендаційної системи для централізованої аптечної мережі, що дозволяє вирішувати задачу з виведення товарів власної торгової марки на ринок, а також виконання функції для прогнозування аптекама об'ємів продажів та залишків.

АПТЕЧНА МЕРЕЖА, КОМПАРАТИВНИЙ АНАЛІЗ, РЕКОМЕНДАЦІЙНА СИСТЕМА, ПРОГНОЗ

ЗМІСТ

Завдання	2
РЕФЕРАТ.....	4
ВСТУП.....	6
РОЗДІЛ 1. Сучасні методи розробки рекомендаційних систем	8
1.1 Компаративний аналіз існуючих моделей рекомендаційних систем	8
1.2 Компаративний аналіз існуючих алгоритмів колаборативної фільтрації.....	9
1.3 Сучасна методика прогнозування об'єму продажів при виході на ринок ВТМ аптечної мережі. Постановка завдання	18
РОЗДІЛ 2. Обґрунтування вибору моделі для розробки рекомендаційної системи для аптечної мережі.	28
1.1 Огляд існуючих моделей рекомендаційних систем для аптечної мережі.....	28
2.2 Обґрунтування моделі для розробки рекомендаційної системи для виходу на ринок ВТМ	36
РОЗДІЛ 3. Експерименти з генерації рекомендацій	41
3.1 Генерація рекомендацій на основі розробленої рекомендаційної системи на об'єм продажів при виході на ринок ВТМ	41
3.2 Порівняння результатів експериментів з результатами прогнозування за існуючою методикою	50
ВИСНОВКИ.....	53
ПЕРЕЛІК ПОСИЛАНЬ	55

ВСТУП

Проблема створення рекомендацій для користувача є актуальною для будь-яких систем, що базуються на продажу або наданні доступу до контенту кінцевому користувачу. Ця галузь почала бурхливо розвиватися із зростанням кількості популярних інтернет сервісів в останні роки.

Раніше для формування рекомендацій використовували список найбільш популярних продуктів, але з часом такі рекомендації почали витіснятися таргетованими (цільовими) пропозиціями: користувачам пропонуються не просто популярні пропозиції, а ті продукти, що напевно мають сподобатися саме їм. Для них життєво важливо надавати можливі варіанти, які або доповнюють товари (у випадку магазинів), або ті, що мають схожу тематику або відгуки, тому що від цього безпосередньо залежить їх прибуток.[1]

Методи фільтрації на основі вмісту засновані на описі елементів та профілі переваг користувача, що будується на попередніх взаємодіях користувача з елементами. Ця взаємодія зберігається у матрицю оцінок елементів, на основі якої потім будуються уподобання користувача. Основна ідея полягає у тому щоб пропонувати користувачу ті елементи, які найбільше відповідають його персональним уподобанням.[2]

Зважаючи на визначну користь рекомендаційних систем у різних областях, у дипломній роботі було обрано дослідження алгоритмів рекомендаційних систем та реалізацію колаборативної фільтрації для мережі аптек, для продажу власних торгових марок.

Мета диплома: розробка рекомендаційної системи для аптечної мережі, що дозволить прогнозувати об'єм продажів при виході на ринок власної торгової марки (ВТМ) на основі даних про схожі товари та аналоги, які вже є на ринку

Задачі диплома:

Дослідити сучасну методику прогнозування об'єму продажів при виході на ринок ВТМ аптечної мережі

Виконати компаративний (порівняльний) аналіз існуючих алгоритмів колаборативної фільтрації

Обґрунтувати вибір моделі для розробки рекомендаційної системи для аптечної мережі.

Виконати ряд експериментів з генерації рекомендацій на основі розробленої рекомендаційної системи на об'єм продажів при виході на ринок ВТМ та об'єм виробництва для подальшого планування діяльності мережі.

Практична реалізація на даних аптечної мережі АНЦ. Розрахунки виконувались на основі даних мережі аптек, розраховано за існуючою методикою на підприємстві та за розробленою рекомендаційною системою, передбачається, що доробки рекомендаційної системи можна буде використати для покращення прогнозування.

Науковою новизною роботи є модель рекомендаційної системи для централізованої аптечної мережі, що дозволяє вирішувати задачу з виведення товарів власної торгової марки на ринок, а також виконання функції для прогнозування аптекам об'ємів продажів та залишків.

РОЗДІЛ 1.

СУЧАСНІ МЕТОДИ РОЗРОБКИ РЕКОМЕНДАЦІЙНИХ СИСТЕМ

1.1 Компаративний аналіз існуючих моделей рекомендаційних систем.

Рекомендаційні системи поділяються на дві основні категорії: системи з алгоритмами на основі пам'яті та системи із алгоритмами на основі моделі. Алгоритми на основі пам'яті запам'ятовують всі значення оцінок і роблять рекомендації на основі співвідношення користувач-товар з рештою матриці. В алгоритмах на основі моделі модель з параметрами спочатку налаштовують на основі матриці оцінок, а потім дають рекомендації, виходячи з моделі.

Двома найпопулярнішими алгоритмами на основі пам'яті є колаборативна фільтрація на основі користувачів і колаборативна фільтрація на основі товарів. Ці алгоритми є прикладами алгоритмів на основі найближчих сусідів, що використовують оцінку схожих користувачів або товарів і дають рекомендації, беручи до уваги підсумок оцінок найбільш подібних користувачів/товарів.

Метод колаборативної фільтрації на основі користувачів побудований на припущенні, що якщо двоє користувачів мають схожі вподобання стосовно деякого товару, то вони матимуть подібні вподобання щодо інших товарів. Аналогічно для методу колаборативної фільтрації на основі товарів, але з точки зору товарів. Використовуючи алгоритми на основі моделі будують модель і рекомендують товар з найвищою оцінкою, яка прогнозується моделлю. Наприклад, метод SlopeOne передбачає використання набору лінійних регресійних моделей (одна модель для кожної пари товарів) з константою. Відповідно, ця константа відображає середню різницю між оцінками двох товарів. Використання цього алгоритму дає змогу досягти високої точності і швидкості обчислення [1].

Інший приклад такого методу – апроксимація на основі SVD розкладу. Таким способом матрицю оцінок розкладають на частини за допомогою сингулярного розкладу, а потім реконструюють, зберігаючи тільки перших r

найбільш значимих власних значень. Це дає можливість спрогнозувати пропущені значення в матриці оцінок.

Найбільш популярний товар. Використовуючи метод найбільш популярного товару (середня оцінка товару), обчислюють середню оцінку для кожного товару на основі наявних оцінок і прогнозують кожну невідому оцінку як середню для товару. В результаті, пропущені оцінки кожного товару будуть однакові для всіх користувачів [2]. За результатами порівняння сформовано таблицю 1.1

Таблиця 1.1 – Результати порівняння моделей рекомендаційних систем

Категорія	Підкатегорія	Алгоритм
На основі пам'яті	Базисні підходи	Найбільш популярний товар
	На основі подібності	Колаборативна фільтрація на основі користувачів
		Колаборативна фільтрація на основі товарів
На основі моделі	Лінійна регресія	Slope One
		Факторизація матриці з допомогою стохастичного градієнтного спуску
	Факторизація матриці	Апроксимація на основі SVD розкладу

1.2 Компаративний аналіз існуючих алгоритмів колаборативної фільтрації

1.2.1 Алгоритм колаборативної фільтрації Item-Base

Метод КФ, який базується на тому, що користувачу сподобаються товари, схожі на ті, які він вже обирає.

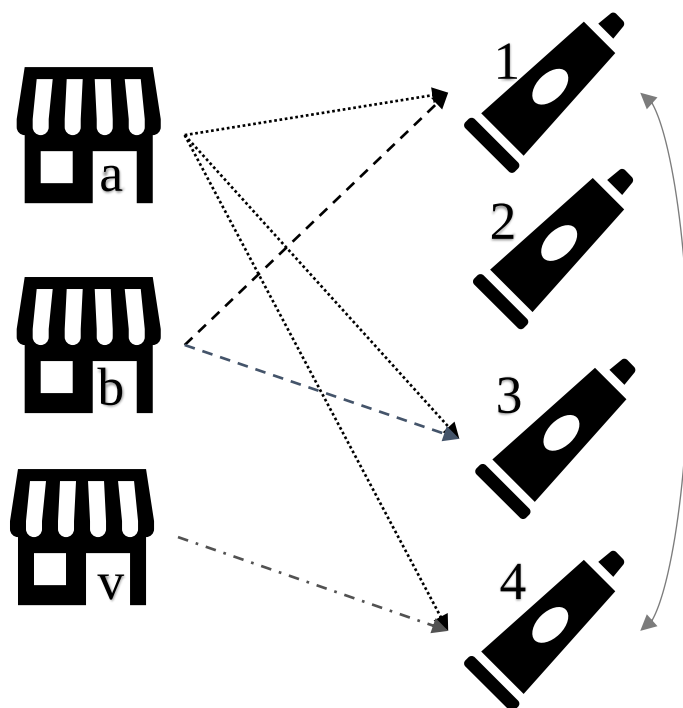


Рисунок 1.1 – Item-Base алгоритм

Користувач А характеризується об'єктами, які він переглянув або оцінив. Для кожного об'єкта з обраних/оцінених визначається m об'єктів-сусідів, тобто знаходяться m найбільш схожих об'єктів з точки зору переглядів / оцінок користувачів. При побудові РС для фільмів, m приймає значення від 10 до 30. Всі об'єкти-сусіди об'єднуються в колекцію, з якої виключаються об'єкти, переглянуті або оцінені користувачем А. А з останків колекції будується топ n рекомендацій. Таким чином, при item-based підході у створенні рекомендацій беруть участь всі користувачі, яким сподобався той чи інший об'єкт з колекції [3].

1.2.2 Алгоритм колаборативної фільтрації UserBase

Метод КФ, який базується на тому, що користувачу сподобаються товари, які були обрані користувачами, схожими на нього.

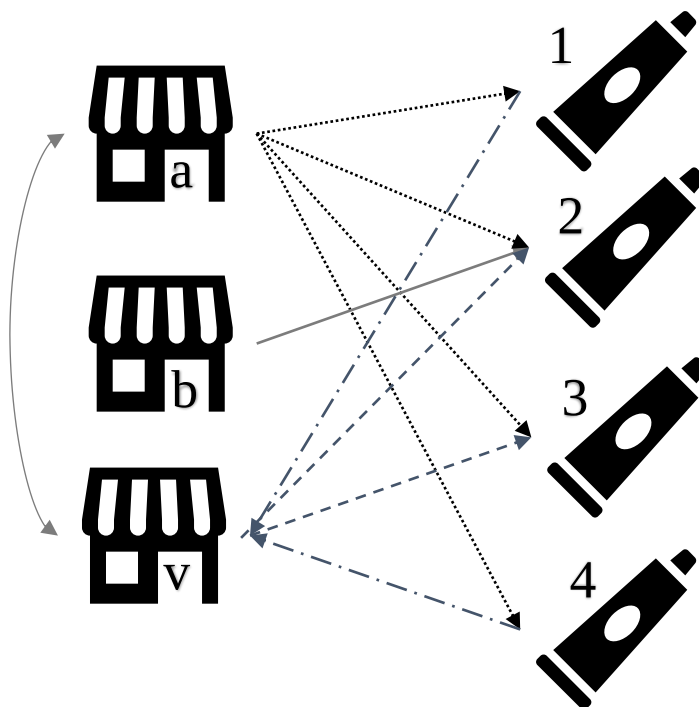


Рисунок 1.2 – User-Base алгоритм

Міра схожості $sim(u, u')$ обчислюється за матрицею оцінок R . Найбільш загальноживана метрика схожості – кореляція Пірсона і косинусна відстань рядків (стовпців) матриці [4].

1.2.3 Алгоритм колаборативної фільтрації Probabilistic matrix factorization (PMF)

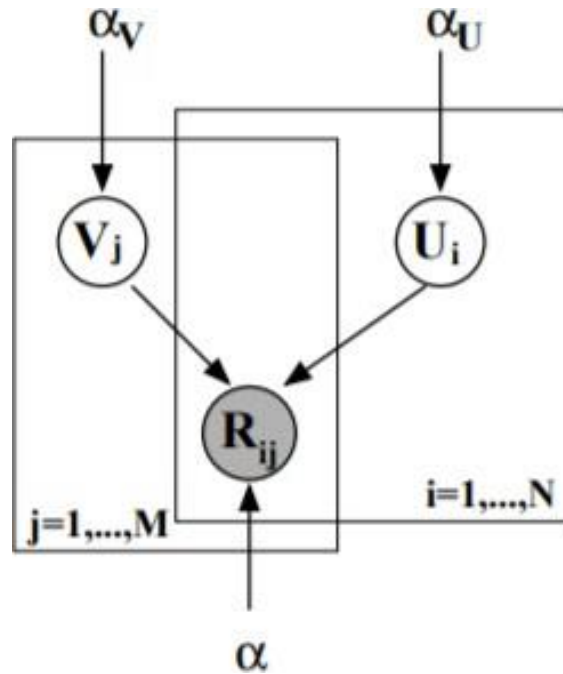


Рисунок 1.3 – PMF модель [5]

Ймовірнісна матрична факторизація (PMF) - імовірнісна лінійна модель з гаусовим спостережним шумом. Припустимо, у нас є N користувачів і M фільми. Нехай R_{ij} стане рейтинговим значенням для користувача i для фільму j , U_i та V_j представляють D -розмірні специфічні для користувача i для фільму вектори латентної функції, відповідно [5].

Умовний розподіл за спостереженими оцінками

$R \in R^{N \times M}$ (термін правдоподібності) та попередні розподіли за $U \in R^{D \times N}$ і $V \in R^{D \times M}$ визначаються за допомогою:

$$p(R|U, V, \alpha) = \prod_{i=1}^N \prod_{j=1}^M \left[\mathcal{N}(R_{ij} | U_i^T V_j, \alpha^{-1}) \right]^{I_{ij}} \quad (1.1)$$

$$p(U | \alpha_U) = \prod_{i=1}^N \mathcal{N}(U_i | 0, \alpha_U^{-1} I) \quad (1.2)$$

$$p(V|\alpha_V) = \prod_{j=1}^M \mathcal{N}(V_j|0, \alpha_V^{-1}I), \quad (1.3)$$

де $\mathcal{N}(x|\mu, \alpha^{-1})$ позначає гауссовий розподіл з середнім μ і точністю α , а I_{ij}

- індикаторна змінна, яка дорівнює 1, якщо користувач оцінив фільм j і рівний 0 в протилежному випадку. Навчання в цій моделі виконується шляхом максимізації над фільмом та функціями користувача з фіксованими гіперпараметрами (тобто дисперсією шумів спостережень та попередніми відхиленнями) [6]:

$$\ln p(U, V|R, \alpha, \alpha_V, \alpha_U) = \ln p(R|U, V, \alpha) + \\ + \ln p(U|\alpha_U) + \ln p(V|\alpha_V) + C, \quad (1.4)$$

де C - це константа, яка не залежить від параметрів. Максимізація цього поперечного розподілу по відношенню до U та V еквівалентно мінімізації функції помилок суми квадратів з членами квадратичної регуляризації

$$E = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^M I_{ij} (R_{ij} - U_i^T V_j)^2 \\ + \frac{\lambda_U}{2} \sum_{i=1}^N \|U_i\|_{\text{Fro}}^2 + \frac{\lambda_V}{2} \sum_{j=1}^M \|V_j\|_{\text{Fro}}^2 \quad (1.5)$$

де $\lambda_U = \alpha_U / \alpha$, $\lambda_V = \alpha_V / \alpha$, і $\|\cdot\|_{\text{Fro}}^2$ позначає норму Фробеніуса. Локальний мінімум об'єктивної функції, заданої у рівнянні [5].

1.2.4 Алгоритм колаборативної фільтрації Bayesian probabilistic matrix factorization (BPMF)

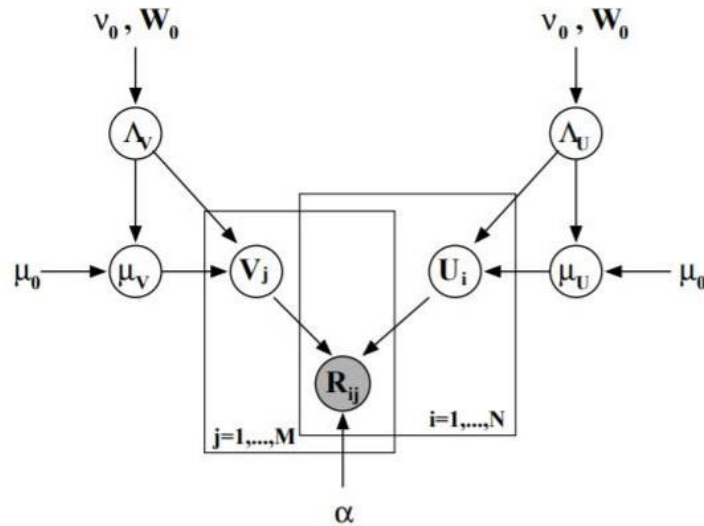


Рисунок 1.4 – BPMF модель [5]

Попередні розподіли над векторами функцій користувача та фільму вважаються гауссовськими:

$$p(U|\mu_U, \Lambda_U) = \prod_{i=1}^N \mathcal{N}(U_i|\mu_U, \Lambda_U^{-1}) \quad (1.6)$$

$$p(V|\mu_V, \Lambda_V) = \prod_{j=1}^M \mathcal{N}(V_j|\mu_V, \Lambda_V^{-1}) \quad (1.7)$$

Надалі ми розміщуємо прихильників Gaussian-Wishart на користувацькому гіперпараметрах фільму $\Theta_U = \{\mu_U, \Lambda_U\}$ і $\Theta_V = \{\mu_V, \Lambda_V\}$:

$$\begin{aligned} p(\Theta_U|\Theta_0) &= p(\mu_U|\Lambda_U)p(\Lambda_U) \\ &= \mathcal{N}(\mu_U|\mu_0, (\beta_0\Lambda_U)^{-1})\mathcal{W}(\Lambda_U|W_0, \nu_0) \end{aligned} \quad (1.8)$$

$$\begin{aligned}
p(\Theta_V|\Theta_0) &= p(\mu_V|\Lambda_V)p(\Lambda_V) \\
&= \mathcal{N}(\mu_V|\mu_0, (\beta_0\Lambda_V)^{-1})\mathcal{W}(\Lambda_V|W_0, \nu_0)
\end{aligned}
\tag{1.9}$$

Тут W - розподіл Wishart з ν_0 ступенями свободи і матриця масштабу $D \times D$ W_0 :

$$\mathcal{W}(\Lambda|W_0, \nu_0) = \frac{1}{C} |\Lambda|^{(\nu_0 - D - 1)/2} \exp\left(-\frac{1}{2} \text{Tr}(W_0^{-1} \Lambda)\right)
\tag{1.10}$$

де C - нормалізаційна константа. Для зручності ми також визначаємо $\Theta_0 = \{\mu_0, \nu_0, W_0\}$. У наших експериментах ми також встановили $\nu_0 = D$ і W_0 на матрицю тотожності як для гіперпараметрів користувача, так і для фільмів, і вибираємо $\mu_0 = 0$ за симетрією [7].

1.2.5 Алгоритм колаборативної фільтрації Singular value decomposition (SVD)

Майже всі алгоритми колаборативної фільтрації мають такі недоліки, як холодний старт, тривіальність результатів рекомендацій тощо. Одним з досить нових алгоритмів, який знижує вплив типових проблем колаборативної фільтрації, виявився SVD алгоритм, який було створено саме для покращення результатів звичайних алгоритмів [8].

SVD - це методика факторизації матриць, яка зазвичай використовується для зменшення кількості функцій набору даних шляхом зменшення розмірів простору від N до K , де $K < N$. Однак для цілей систем рекомендацій нас цікавить лише матрична факторизація, де частини зберігають таку ж розмірність. Матрична факторизація виконується на матриці рейтингів користувацьких

позицій. Матрична факторизація може розглядатися як пошук 2 матриць, добуток яких є вихідною матрицею [9].

$$expected\ rating = \hat{r}_{ui} = q_i^T p_u \quad (1.11)$$

‘ q_i ’ і ‘ p_u ’ можна знайти таким чином, що різниця квадратних помилок між їх точками продукту та відомим рейтингом у матриці користувачького елемента є мінімальною.

$$minimum(p, q) = \sum_{(u,i) \in K} (r_{ui} - q_i^T \cdot p_u)^2 \quad (1.12)$$

1.2.6 Алгоритм колаборативної фільтрації Regularized single value decomposition (RSVD)

Щоб наша модель могла добре узагальнювати і не перевищувати навчальний набір, ми запроваджуємо строк покарання для нашого рівняння мінімізації. Це представлено коефіцієнтом регуляризації λ , помноженої на квадратну суму величин векторів користувача та елементів [9].

$$minimum(p, q) = \sum_{(u,i) \in K} (r_{ui} - q_i^T \cdot p_u)^2 + \lambda(\|q_i\|^2 + \|p_u\|^2) \quad (1.13)$$

1.2.7 Алгоритм колаборативної фільтрації Slope One

Slope One - це один з найпростіших підходів до рекомендацій на основі

колаборативної фільтрації за подібністю предметів, але в той же час точність рекомендацій алгоритму порівняна з більш складними та ресурсоемними алгоритмами. [10]

У алгоритмі SlopeOne враховуються як інформація від інших користувачів, які оцінили той самий елемент, так і з інших предметів, оцінених одним і тим самим користувачем. Проте, схеми також розраховуються на точках даних, які не потрапляють ні в масив користувача, ні в масив об'єкта, але, однак, є важливою інформацією для прогнозування рейтингу [11].

Враховуючи набір тренувальних наборів χ та будь-які два пункти j та i з рейтингом u_j та u_i відповідно в деякій оцінці користувача u , ми розглянемо середнє відхилення пункту i стосовно пункту j як:

$$\text{dev}_{j,i} = \sum_{u \in S_{j,i}(\chi)} \frac{u_j - u_i}{\text{card}(S_{j,i}(\chi))} \quad (1.14)$$

Будь-яка оцінка користувача, яка не містить як u_j , так і u_i , не входить до суми. Симметрична матриця, визначена $\text{dev}_{j,i}$ може бути одночасно обчислена і швидко оновлюватися при введенні нових даних [12].

Враховуючи те, що $\text{dev}_{j,i} + u_i$ є прогнозом для u_j даного u_i , розумним предиктором може бути середнє значення всіх таких прогнозів.

$$P(u)_j = \frac{1}{\text{card}(R_j)} \sum_{i \in R_j} (\text{dev}_{j,i} + u_i) \quad (1.15)$$

Можемо спростити формулу до:

$$P^{S1}(u)_j = \bar{u} + \frac{1}{\text{card}(R_j)} \sum_{i \in R_j} \text{dev}_{j,i}. \quad (1.16)$$

1.2.8 Алгоритм колаборативної фільтрації Non negative matrix factorization (NMF)

Факторизація невід'ємних матриць (NMF) - це уявлення матриці V у вигляді добутку матриць W і H , в якому всі елементи трьох матриць невід'ємні [12].

Нехай таблиця V має розмір $m \times n$. Позначимо через r ранг матриць W і H , як правило $r \ll \min(n, m)$. На відміну від точного уявлення матриці в SVD, в NMF маємо тільки наближену рівність.

$$V \approx WH. \quad (1.17)$$

Матриці W і H вибираються таким чином, щоб мінімізувати функцію втрати: $D(V, WH) \rightarrow \min$. У нашому випадку D задається на основі дивергенції Кульбака-Лейблера

$$D(A, B) = \sum_{i,j} a_{ij} \log\left(\frac{a_{ij}}{b_{ij}}\right) - a_{ij} + b_{ij}. \quad (1.18)$$

1.3 Сучасна методика прогнозування об'єму продажів при виході на ринок ВТМ аптечної мережі. Постановка завдання

Постановка задачі: в наявності є товар власної товарної марки (надалі - ВТМ), існують схожі товари та його прямі аналоги, що формують групу товарів. Потрібно дати рекомендацію на об'єм продажів при виході на ринок ВТМ та об'єм виробництва для подальшого планування на основі даних про схожі товари та аналоги, які вже є на ринку.

1.3.1 Отримання даних з внутрішніх систем

Для розрахунків використовуються вивантаження з корпоративних баз даних, наприклад, qlikview. Для прогнозу необхідні дані про продажі в упаковках товарів щомісячно, кількість точок продажу, товарні залишки в аптеках, товарні залишки на складах постачальника.

1.3.2 Створення списку товарів-конкурентів для ВТМ

Потрібно об'єднати товар ВТМ зі схожими товарами у групу товарів, об'єднання відбувається на основі схожих ключових назв та однакового застосування, наприклад, Диклофенак гель об'єднано за параметрами: в назві є «Диклофенак », форма застосування гель.

На цьому кроці також за мірою схожості виставляється частка фармацевтичної заміни у Диклофенак гель, схожість за вподобаннями чи наявність комбінацій або додаткових складників.

1.3.3 Робота з базою даних

База даних має під собою структуровану інформацію про щомісячні продажі товарів в аптеках, для роботи об'єднуються база даних та таблиця з ВТМ та їх конкурентами, у результаті отримується таблиця з щомісячними продажами ВТМ та їх товарами-конкурентами.

1.3.4 Залишки товару в аптеках та складах

Для розрахунку та планування виробництва необхідно зібрати інформацію по залишкам ВТМ в аптеках та складах постачальників за допомогою програмних комплексів, та звести у таблицю та додати за допомогою словників інформацію про код, який закріплений за позицією у внутрішній базі даних.

1.3.5 Методика прогнозу продажів та об'єму виробництва ВТМ

Вирішено надавати рекомендацію об'єму виробництва за висновком прогнозу з врахуванням коефіцієнтів, які впливають на продажі ВТМ.

Крок 1. Мережа аптек є динамічною системою, що швидко розростається, тому для прогнозу було вирішено визначити темп приросту аптек T_i , за формулою:

$$T_i = Q_i / Q_{i-1}, \quad (1.19)$$

де Q_i - кількість аптек в поточному місяці;

Q_{i-1} - кількість аптек в минулому місяці.

Наприклад: у серпні кількість аптек становила 1034, а у вересні 1044, тоді темп приросту аптек $T_i = 1044 / 1034 = 0,97\%$.

Розрахунок Темпу приросту кількості аптек за весь період наведено на рисунку 1.5



Рисунок 1.5 – Темп приросту аптек

Крок 2. Далі потрібно визначити темп приросту в продажах групи товарів, для цього побудувати таблицю з продажами щомісячно групи товарів та вирахувати відсоткове відхилення місяця поточного до місяця попереднього за вирахуванням темпу приросту.

Вирахувати темп приросту потрібно для зменшення впливу від збільшення кількості аптек на приріст продажів групи товарів для подальшого коректного розрахунку .

Є потреба очистити продажі в упаковках від впливу зміни кількості аптек для розрахунку темп приросту, для цього було вирішено розрахувати показник «Продажі ЧИЩ, уп.» $S_{\text{чищ}}$ за формулою [13]:

$$S_{i \text{ чищ}} = S_i * (1 - T_i) \quad (1.20)$$

де S_i – продажі за місяць

T_i – темп приросту аптек за обраний місяць

Розрахунки наведені на рисунку 1.6



Рисунок 1.6 – Продажі групи товарів за вирахуванням темпу приросту аптек

Крок 3. В подальшому розрахунку потрібно визначити наявність приросту об'єму продажів групи товарів. Для більшості груп товарів є період, в який продажі значно збільшуються, тому для розрахунку приросту C використано формулу [13]:

$$C_i = S_{\text{чищ } i} / S_{\text{чищ } i-1} \quad (1.21)$$

Де $S_{\text{чищ } i}$ – очищенні продажі поточного місяця

$S_{\text{чищ } i-1}$ - очищенні продажі попереднього місяця

Розрахунок приросту групи товарів наведено на рисунку 1.7

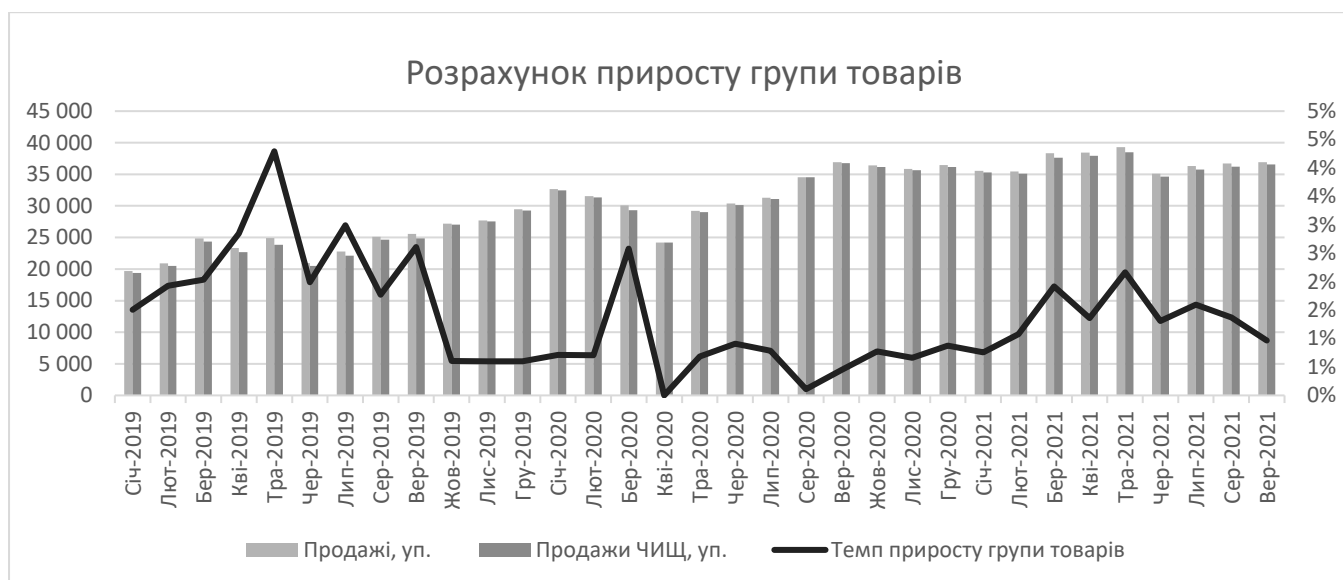


Рисунок 1.7 – Розрахунок приросту групи товарів

Занісши фактичні дані та розділивши об'єми продажів групи товарів на об'єми продажів ВТМ та товарів-конкурентів, розраховується частка ВТМ в продажах групи, в результаті побудовано рисунок 1.8, з даних якого будується прогноз.

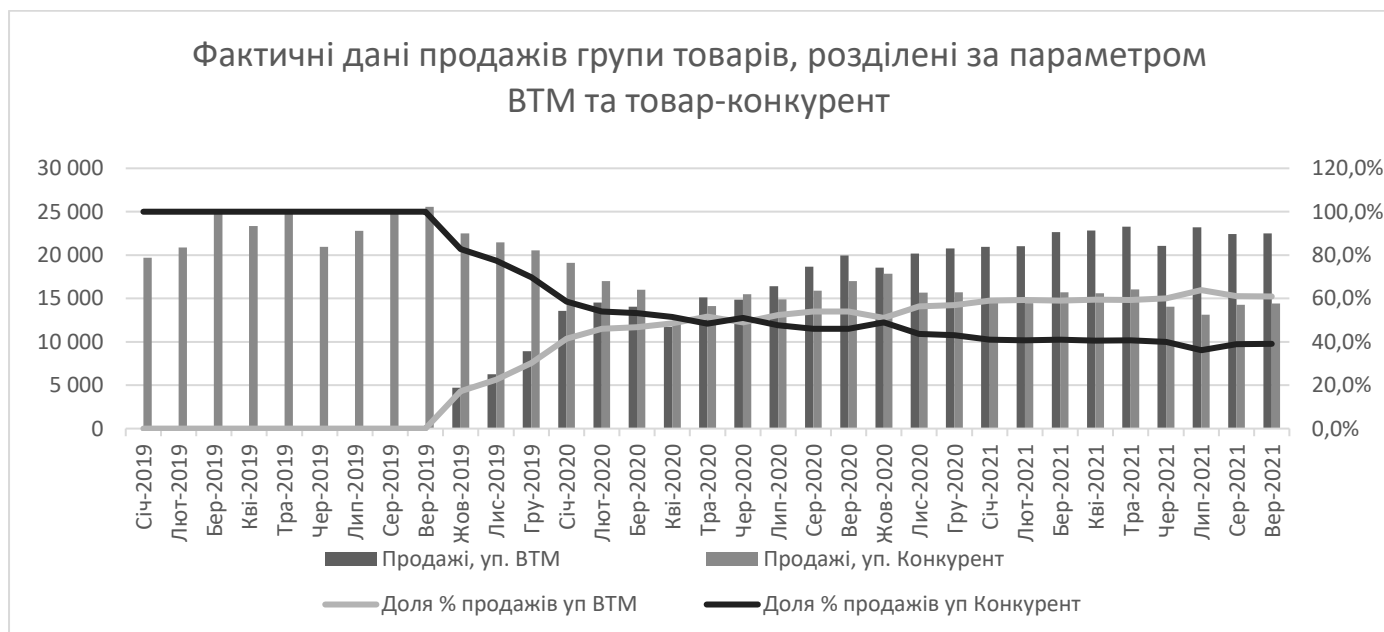


Рисунок 1.8 – Фактичні дані продажів групи товарів, розділені за параметром ВТМ та товар-конкурент

Крок 4. Розраховую прогноз групи товарів в цілому $\Pi_{\text{груп}}$ за формулою:

$$\Pi_{\text{груп}} = S_{i-1} * (1 + C_i + T_{\text{сер}}) \quad (1.22)$$

де S_{i-1} - продажі минулого місяця

C_i - приріст продажів

$T_{\text{сер}}$ - середня частка приросту аптек за останній рік

Прогноз фармацевтичної заміни ВТМ конкурентів в групі товарів розраховується як середнє значення фактичної фармацевтичної заміни за останні 2 місяці при умові що ВТМ знаходиться в продажі довше 3 місяців. Кількість місяців продажів розраховується як сума заповнених комірок в стовпці ВТМ продажі в упаковках. В даному випадку місяців продажів 31 що більше ніж 3, тому розрахунок відбудеться за середнім значення фармацевтичної заміни двох останніх місяців і дорівнює 60%.

Отримавши прогноз продажів групи товарів і прогноз фармацевтичної заміни, достатньо даних для здійснення прогнозу продажів ВТМ для кожного наступного місяця $\Pi_{\text{ВТМ}}$ за формулою:

$$\Pi_{\text{ВТМ}} = \Pi_{\text{груп}} * \Pi_{\text{заміни}} \quad (1.23)$$

Де $\Pi_{\text{груп}}$ - прогноз групи товарів в цілому

$\Pi_{\text{заміни}}$ - середнє значення частки продажів ВТМ за останні 2 місяці

В результаті отримую таблицю з прогнозом продажів ВТМ до кінця 2022 року, роботи прогноз на більший період немає сенсу через значне збільшення похибки (більше 5%)

Результат прогнозу наведено на рисунку 1.9

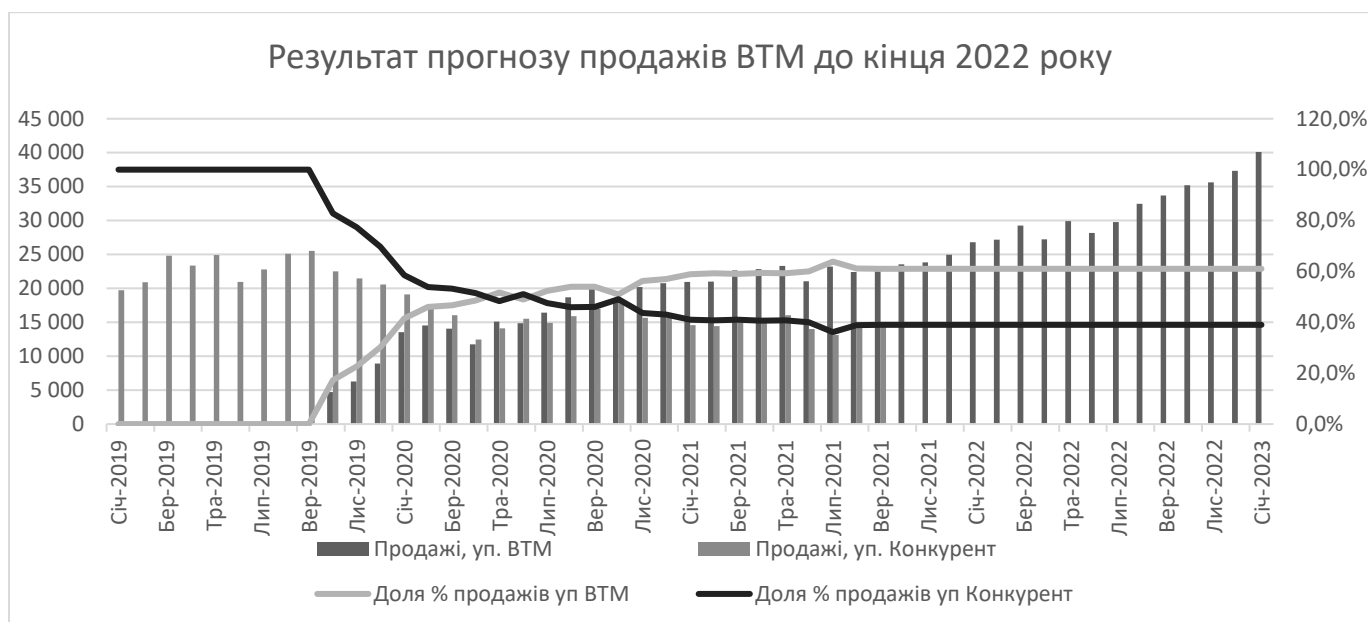


Рисунок 1.9 – Результат прогнозу продажів ВТМ до кінця 2022 року

Помаранчевим виділено період прогнозування продажів ВТМ в упаковках та фармацевтична заміна.

Після здійснення прогнозу продажів достатньо даних для надання рекомендації, щодо закупівель ВТМ та об'єму виробництва до кінця року.

Для початку потрібно визначити загальний залишок ВТМ на кінець місяця у поточному місяці, додавши залишки в аптеках, складах та заплановане надходження.

Для можливості своєчасного постачання в аптеки ВТМ визначено норматив в підтримці залишків в упаковках на шістдесят днів реалізації

Для обчислення днів реалізації залишку з запланованим надходженням $D_{\text{зал}}$ використовую формулу [13]:

$$D_{\text{зал}} = (Z + H) / S_i * D \quad (1.24)$$

де Z - Загальний залишок на кінець місяця

H - Надходження поставки на початок місяця в уп

S_i - Продажі, уп.

D - Кількість днів в місяці

Тепер можна вирахувати потребу поставки на початок місяця ВТМ та надати рекомендацію об'єму виробництва виробникам. Враховуючи мінімальну партію до замовлення якщо виробник ставить таку умову в даному випадку мінімальна партія 1000 упаковок та розмір одного ящика що виступає кратністю в даному випадку 100 упаковок Розрахунок потреби P відбувається за формулою:

$$P = \{ \text{Якщо } D_{\text{зал}} < 60, \text{ то Макс } \{ V_{\text{мін}} ; 3 / D_{\text{зал}} * (60 - 3) \} \} \quad (1.25)$$

де $D_{\text{зал}}$ - Дні реалізації залишку з запланованим надходженням

$V_{\text{мін}}$ - Мінімальна партія до замовлення

3 - Загальний залишок на кінець місяця

На рисунку 1.10 наведено результат розрахунку потреби на початок місяця в упаковках, що є рекомендацією об'єму виробництва.

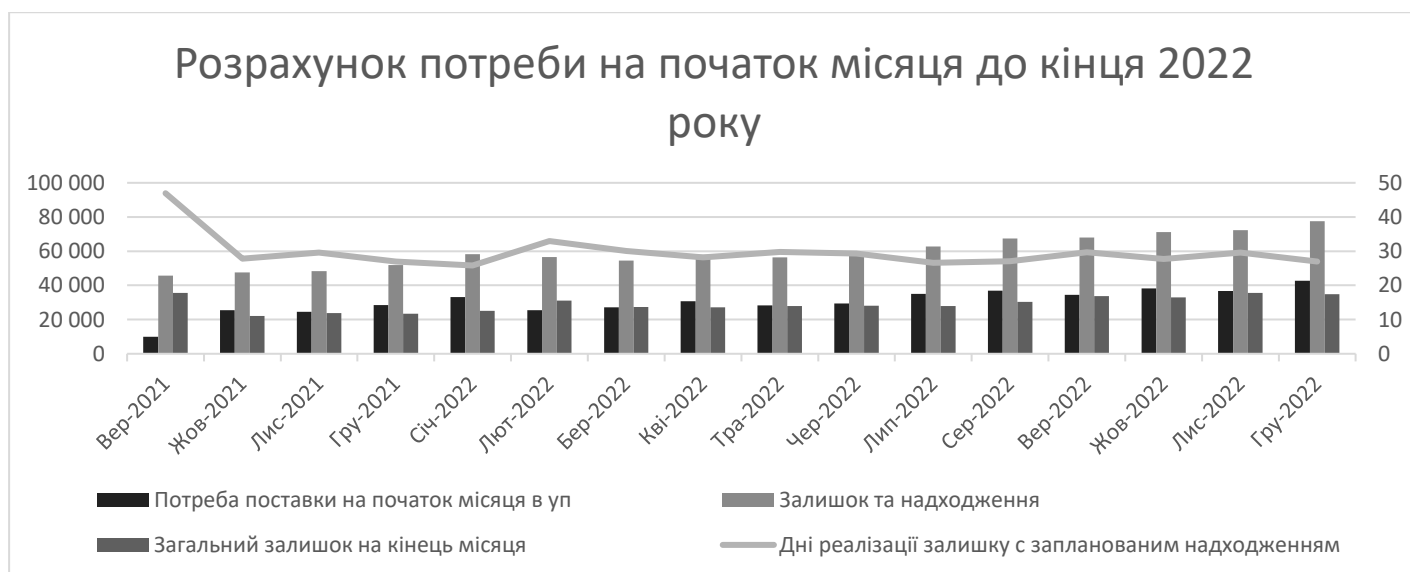


Рисунок 1.10 – Розрахунок потреби на початок місяця до кінця 2022 року

В результаті отримана рекомендація до виробництва ВТМ усього в упаковках до кінця 2022 року та у розрізі місяців. До кінця 2022 року буде потреба в 742 300 упаковок. При цьому варто зауважити, що розрахунок оновлюється кожного місяця. Похибка складає не більше 5% і в середньому становить 3%, на прогнозуванні товарів ВТМ в яких не відбувалось змін.

Висновки до розділу 1

В розділі 1 наведено теоретичні основи рекомендаційних систем та їх алгоритмів. Часто використовуваним в рекомендаційних є алгоритм колаборативної фільтрації. Це зумовлено тим, що він є простим як для аналізу ефективності так і для побудови. Зокрема його можна використовувати на найбільш розповсюджених програмних засобах Microsoft Office. Також він може давати рекомендацію на доволі доступному малому обсязі первинних даних, що може бути як оцінкою користувачів та і показниками продажів товару. Це робить його популярним вибором для побудови рекомендаційних систем на ранньому етапі та в комерційних компаніях.

Він побудований на припущенні, що якщо двоє користувачів мають схожі дії стосовно деякого товару, то вони матимуть подібні дії щодо інших товарів. Використовуючи алгоритми на основі моделі будують модель і оцінюють товар, для рекомендації іншим користувачам.

Наведено існуючу модель прогнозування на підприємстві, зібрано дані про кількість аптек в мережі та розраховано їх темпи приросту, зібрано дані про продажі ВТМ та їхніх конкурентів за останні 2 роки, виконано розрахунки темпу приросту продажів групи конкурентів.

Розрахований прогноз продажів для групи в цілому та ВТМ окремо. Зібрано дані залишків ВТМ в мережі та залишків на складах постачальника для прогнозування замовлення у виробника. В результаті отримано прогноз продажів, що може використовуватись для формування необхідних залишків на аптеках та прогноз замовлення для відправки виробнику ВТМ. На основі частини цих даних відбудуватиметься розрахунок в рекомендаційній системі.

РОЗДІЛ 2.

ОБҐРУНТУВАННЯ ВИБОРУ МОДЕЛІ ДЛЯ РОЗРОБКИ РЕКОМЕНДАЦІЙНОЇ СИСТЕМИ ДЛЯ АПТЕЧНОЇ МЕРЕЖІ.

2.1 Огляд існуючих моделей рекомендаційних систем для аптечної мережі

2.1.1 Побудова рекомендаційної системи на основі нейронної мережі

Нещодавні досягнення в глибоких нейронних мережах для графічно-структурованих даних привели до найсучаснішої продуктивності в тестах рекомендаційної системи. Адаптація цих методів до завдань рекомендацій щодо перехресного продажу аптечних товарів із мільйоном продуктів і сотнями мільйонів продажів залишається проблемою через складні медичні та юридичні властивості фармацевтичних даних. Щоб вирішити цю проблему, Franz Hell [14] розробили алгоритм згорткової мережі на графі (GCN) під назвою PharmaSage, який використовує згортки графів для створення вставок для аптечних продуктів, які потім використовуються в рекомендаційному завданні. У базовий граф автори включають, як інформацію про перехресні продажі з транзакції продажу в структуру графа, так і інформацію про продукт як функції вузла. За допомогою модифікацій вибірки, залученої до процесу оптимізації мережі, автори звернулись до поширеного явища в рекомендаційних системах, так званого упередження популярності: популярні продукти часто рекомендуються, тоді як менш популярні елементи часто нехтують і рекомендуються рідко, або взагалі не рекомендуються. Автори розгорнули PharmaSage, використовуючи реальні дані про продажі, і навчили його на 700000 статей, представлених у вигляді вузлів на графі з ребрами між вузлами, що представляють приблизно 100 мільйонів транзакцій продажу. Використовуючи властивості фармацевтичного продукту, такі як показання, інгредієнти та побічні ефекти, і поєднуючи їх із великими історіями продажів, автори досягли кращих результатів, ніж при підході, заснованому на чисто статистичних даних.

Щоб використовувати можливості згорткових нейронних мереж із графами для вивчення вбудовування продуктів, автори перетворюють дані про продажі в аптеках та інформацію про продукт у граф. Автори представляють кожен унікальний аптекний продукт у вигляді вузла, який також містить описову інформацію відповідного продукту, закодовану у мульти-горячому векторі, який має близько 15 000 записів, що представляють медичні показання, діючі речовини та побічні реакції продуктів. Дані були отримані з основної німецької комерційної фармацевтичної бази даних, яка використовується у фармацевтичних та інших програмних рішеннях для аптек, лікарів, лікарень та інших медичних закладів, наданої Pharmatechnik. Неорієнтовані зважені межі між двома вузлами потім показують, як часто відбувалися перехресні продажі для кожної пари продуктів, де «перехресний продаж» визначається як два продукти, продані в рамках однієї транзакції. Набір із приблизно 100 мільйонів транзакцій з інформацією про те, які продукти були продані разом, був наданий Pharmatechnik, використовуючи операції з продажу, які документуються через їх систему управління аптеками IXOS, яка наразі використовується в більш ніж 5000 аптеках. Перед подальшою обробкою обмежують транзакції, на яких базують показники перехресних продажів, до тих транзакцій з двома або трьома проданими товарами, що становить приблизно 25% від повного набору. Автори припускають, що, чим менше товарів в одній транзакції, тим точніше пов'язані продані продукти між собою. Ці транзакції разом включають близько 700 000 різних продуктів, але також включають багато подібних продуктів, які лише незначно (або зовсім не відрізняються один від одного). Наприклад, існує безліч пропозицій продуктів Aspirin 100 від різних виробників. Для навчання, перевірки та тестування випадковим чином вибрали 60% цих попередньо відібраних транзакцій для побудови навчального графа і 20% кожену для побудови графів перевірки та тестування. Ці графи є вхідними для етапу навчання та перевірки моделі. Однак остаточна модель, яка використовується для висновку, потім навчається на всіх включених операціях продажу. Під час навчання моделей автори використовують найпопулярніші статті з перехресних продажів (ті, що

мають найвищу вагу країв із заданим запитом) як кандидатів для агрегації та оптимізації процесу моделювання.

Автори моделі згорткової нейронної мережі з графом використовують локалізовані згортки на агрегованих векторах сусідства для створення вкладень продуктів, представлених вузлами графа. Основна ідея полягає в тому, що автори перетворюють представлення сусідів даного вузла через щільну нейронну мережу, а потім застосовують функцію агрегатора/об'єднання на результуючому наборі векторів. Цей крок агрегації забезпечує векторне представлення локальної околиці вузла. Потім автори об'єднують цей агрегований вектор сусідства (темно-сіра рамка) з поточним представленням вузлів (світло-сіра рамка) і перетворюють об'єднаний вектор через інший щільний шар нейронної мережі. Результатом роботи алгоритму є подання вузла, яке називається вбудовуванням вузла, яке включає інформацію про себе та його локальне оточення графа. Зміна, яку внесли автори, полягає в тому, що агреговано інформацію про вузли оточень лише серед «верхніх сусідів» даного вузла. Верхні сусідні вузли даного вузла запити – це вузли з найбільшою вагою краю серед усіх підключених вузлів, що представляють продукти, які продаються разом із запитом частіше, ніж очікувалося. Потім автори обчислюють рекомендації, використовуючи ці остаточні вбудовування вузлів.

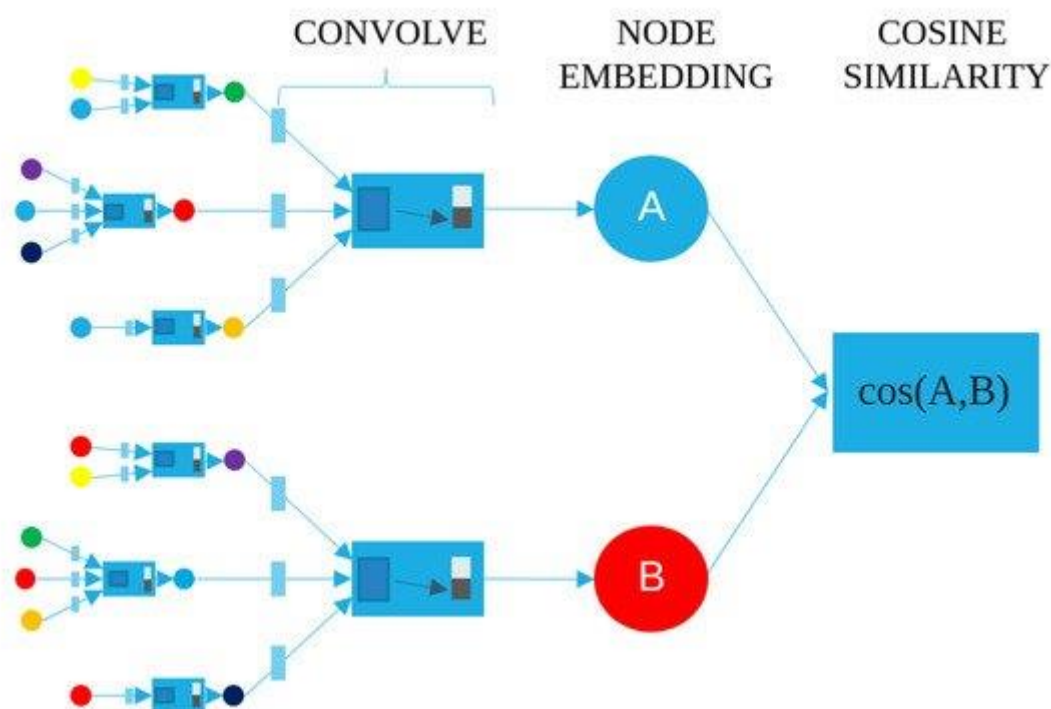


Рисунок 2.1 – Графова згортка [14]

Графово згорткова нейронна мережа використовує локалізовані згортки на агрегованих векторах сусідства для вивчення вбудовування продуктів. Тут автори показують нейронну мережу двошарового графа, яка обчислює остаточне вбудовування вузлів A і B, використовуючи попереднє представлення рівня вузлів A і B відповідно, а також їх відповідних околиць. Різні кольори позначають різні вузли. Рекомендаційний бал між двома продуктами A, B потім обчислюється за допомогою використання косинусної подібності між двома кінцевими векторами вбудовування вузлів A і B.

Остаточні вбудовування, які обчислюються нашою моделлю після навчання, потім використовуються для обчислення рекомендаційних балів для всіх продуктів. Автори отримали рекомендаційний бал, обмежений інтервалом $[0, 1]$ з використанням косинусної подібності двох векторів, $sim(A, B)$, які зазвичай використовуються в техніках пошуку інформації та інтелекту. Враховуючи показники схожості між одним продуктом запиту та всіма іншими продуктами (крім ліків за рецептом, які можна продавати лише за призначенням лікаря, і, отже, не є законними рекомендаціями щодо

перехресного продажу), вибираються продукти з найбільшою схожістю із запитом. як рекомендації до статті запиту .

Для програмного забезпечення автори використовували фреймворк Python, бібліотеку networkx для створення графа та фреймворк PyTorch для реалізації GCN. Мережа тренується на системі з процесором Intel Core i7 8750H, графічним процесором Nvidia GeForce GTX1070 і 32 ГБ оперативної пам'яті. Навчання займає приблизно 48 годин і припиняється, коли втрата триплетів, яка починається з межі 0,5, досягає порогу 0,05 для графіка навчання. Втрата підтвердження становить 0,092, а втрата при тестуванні – 0,095. Автори виявили, що емпірично якість результатів не значно покращується після цього моменту. Автори припинили навчання моделі, яка використовується для висновку на тому самому порозі.

Остаточні вбудовування, які обчислюються моделлю в статті після навчання, потім використовуються для обчислення рекомендаційних балів для всіх продуктів. Автори отримали рекомендаційний бал, обмежений інтервалом $[0, 1]$ з використанням косинусної подібності двох векторів, $\text{sim}(A, B)$, які зазвичай використовуються в техніках пошуку інформації та інтелекту. Враховуючи показники схожості між одним продуктом запиту та всіма іншими продуктами (крім ліків за рецептом, які можна продавати лише за призначенням лікаря, і, отже, не є законними рекомендаціями щодо перехресного продажу), вибираються продукти з найбільшою схожістю із запитом. як рекомендації до статті запиту .

Система наведена в англомовній статті хоч і є більш точною, проте має вади, серед яких є проблема накопичення вхідних даних, яка полягає в тому, що ринок фарм продукції зарегульований і має в свої основі велику частку патентів, що ускладнює порівняння препаратів, які мають схожу дію і склад та відрізняються не суттєво для покупця. Також є проблема в відсутності великої вибірки повних, актуальних та поновлювальних даних для навчання та удосконалення нейронних мереж, що є критичним для рекомендації саме власних торгових марок.

2.1.2 Модель рекомендаційної системи для аптечної мережі

В іншій статті [15] автори підкреслюють, що двома базовими елементами будь-якої рекомендаційної системи є суб'єкт інтересу (це аптека – в другому блоці, користувач інтернет-аптекою – в третьому блоці) та предмет інтересу (медичні товари і лікарські засоби). Відповідно, рекомендаційна система отримує від аптеки інформацію по певному товару, а потім видає йому прогнозовану оцінку (*educated guess*) інших товарів. Формат даних, які надає аптека системі - це дані по проданим товарам [15].

В першому блоці виконується кластеризація аптек мережі у просторі ознак, що характеризують геолокацію та ефективність функціонування. Для реалізації завдань блоку авторами в роботі [15] була побудована система підтримки прийняття рішень (СППР) на основі вирішення задачі лінгвістичної апроксимації. Серед аптек мережі з однаковими характеристиками розташування (геолокації) виокремлюються два класи: ефективно функціонуючих (клас A_i) та неефективних (клас B_i), де індекс i вказує певну групу геолокації.

Другий блок містить офлайн-модель розробки рекомендацій з формування оптимального портфелю для аптек класу B_i на основі досвіду ефективно функціонуючих аптек, що мають схожу з ними геолокацію. Таким чином забезпечується підвищення якості портфелю неефективних аптек, а також неявно враховано користувацькі вподобання, що знайшли відображення у складі портфелів аптек класу A_i . В основу моделі другого блоку покладено колаборативну фільтрацію методом найближчого сусіда з косинусною мірою схожості.

Варто відмітити, що ці прогнози є індивідуальними, незважаючи на те, що інформація зібрана від багатьох учасників. Тим самим даний метод відрізняється від більш простого підходу, який дає усереднену оцінку для кожного об'єкта.

Для її реалізації було побудовано таблицю аптек досліджуваної аптечної мережі з відповідними до них портфелями товарів медичного призначення та

лікарських засобів, в якій аптеки діляться за результатом кластеризації першого блоку на два класи A_i та B_i (показано на рис. 2.2).

Місто	Товар Адреса аптеки	Клас												
			Адреналіна гідрохлориду розчин 0,1% р-р д/ин. 0,1 %	Андрокур табл. 50 мг фл. 20 кор.	Аспаркам табл. уп/контурн.б/яч. 10	Аспаркам табл. уп/контурн.б/яч. 20	Аспірин табл. 500 мг бл. 10	Аспірин табл. 100 мг бл. 10	ацетилсаліцилова кислота табл. 0,5 г уп/контурн.б/яч. 10	АЦЦ 100 таблшип. 100 мг туба 20 кор.	АЦЦ 200 таблшип. 200 мг туба 20 кор.	АЦЦ лонг таблшип. 600 мг туба 10 кор.	Беродуал р-р д/інгал. фл.-кап. 20 мл кор.	...
ХЕРСОН	гв. Херсонської дивізії, 10	А								66	6		3	...
КИЇВ	А.Ахматової, 50	В									1	1		...
ПОЛТАВА	Шевченка, 50а	А			5				1	75	2	10	5	...
КИЇВ	Бальзака, 80	А	2		26	1					3	5	4	...
ВАСИЛЬКІВ	Грушевського, 22/2	А									1	2	1	...
КИЇВ	Давидова, 2/7	А			21		1				3	5	10	...
ОДЕСА	Канатна, 72	А		6				1		80	1		4	...
ЖИТОМИР	Київська, 39	В						2	1	20		1	3	...
КИЇВ	Константиновська, 25	А									1	3	5	...
			...	...	...	...	...	...	...	...	...	...	...	...

Рисунок 2.2 – Фрагмент матриці суб'єкт (аптека)-об'єкт (товари) [15]

За цією інформацією в середині кожної групи автори виділили найпопулярніші товари медичного призначення, які аптеки продають частіше всього. Отже, кожні учасники певної групи рекомендують популярні товари.

Автори ввели позначення для вихідних даних [15]:

U – множина торгівельних точок (аптек) $U = \{u_1, u_2, \dots, u_m\}$;

I – множина товарів (лікарських засобів або товарів медичного призначення), $I = \{i_1, i_2, \dots, i_n\}$

Y – простір опису транзакцій;

D – транзакційні події, $D = (u_t, i_t, y_t)_{t=1}^m \in U \times I \times Y$.

Агреговані дані:

$R = \|r_{ui}\|$ – матриця розміру $|U| \times |I|$,

де $r_{ui} = \text{aggr}\{(u_t, i_t, y_t)\} \in D$ $u_t = u, i_t = i$

R_{ui} – відповідає події: аптека u продала товар i .

Всі наявні продажі від усіх аптек створюють матрицю $U \cdot I$.

Системи колаборативної фільтрації зазвичай застосовують двоступеневу схему:

- знаходять тих, хто розділяє оціночні судження активного користувача;
- використовують оцінки користувачів, що мислять подібно активному користувачу для обчислення прогнозу

Автори розглянули окремо аптеку a класу V_i та ті аптеки класу A_i , що продали товар j . Алгоритм найближчого сусіда з обраною мірою схожості включає в себе три кроки:

- для кожної аптеки u визначимо, наскільки її інтереси співпадають з інтересами аптеки a ;
- після цього оберемо множину аптек, найбільш близьких до a ;
- визначимо прогнозну оцінку на основі оцінок об'єкта j «сусідами» з попереднього кроку.

Кожній аптеці в матриці відповідає один рядок. Тож, треба обчислити близькість векторів-рядків.

Для обчислення міри близькості між векторами в моделі було обрано косинусну міру. Результатом другого блоку є пріоритети товарів для формування портфелів аптек класу V_i . Побудовану в другому блоці офлайн-модель можна також використовувати для формування портфелю товарів нової аптеки, що немає ще власної історії продаж [15].

Третій блок має на меті використання онлайн-моделі, що ґрунтується на транзакційній історії інтернет-аптек мережі та даних про вподобання певних користувачів.

Дві головні переваги моделі із статті - це збільшення точності і можливості більш гнучкого налаштування на різні групи клієнтів. До недоліків слід віднести меншу інтерпретованість та більшу складність реалізації і підтримки. Наведена

модель є хорошою але складність інтерпретації не дозволяє використати її для вирішення задачі.

2.2 Обґрунтування моделі для розробки рекомендаційної системи для виходу на ринок ВТМ

Постановка задачі.

На відміну від статті [14] в даній дипломній роботі проблема маркетингу мережі аптек полягає не в максимізації перехресних продажів, а в виведенні на конкурентний ринок товарів власної торгової марки. Кожен з таких товарів може мати на ринку аналоги, які грають роль товарів-конкурентів. Для спрощення будемо далі вважати, що такі товари – ВТМ і його аналоги - не будуть продаватися в одній транзакції, хоча в практиці такі випадки можливі за різних обставин.

Задано:

Множина товарів власної торгової марки $ВТМ = \{ ВТМ_i \}$, $i=1, \dots, n$;

Множина аптек централізованої аптечної мережі $A = \{ A_j \}$, $j=1, \dots, m$;

Множину товарів-конкурентів для i -го товару $ВТМ_i$ позначимо, як $K_i = \{ k_{is} \}$, $i=1, \dots, n$; $s=1, \dots, p_i$.

Для кожного товару з множин $ВТМ$ та K_i відомо обсяг закупок кожною аптекою на фіксовану дату та залишки:

для обсягу товарів $ВТМ$ i -го виду в аптеці A_j введемо позначення - V_{ij} ;

для залишків товарів $ВТМ$ i -го виду в аптеці A_j введемо позначення - Z_{ij} ;

через V_{ijs} позначимо обсяг s -го конкурента для $ВТМ$ i -го виду в аптеці A_j .

Необхідно надати рекомендації для мережі аптек з питань обсягів закупок товарів, за яких можна максимізувати обсяг продажів товарів $ВТМ$. Рекомендація розраховується окремо за трьома моделями які наведені нижче:

Модель 1. Вихідні дані формуються у вигляді матриці обсягів продажів V_{ij} , де на перетині рядків і стовпців вихідної таблиці знаходиться обсяг товарів V_{ij} .

Модель 2. Вихідні дані формуються у вигляді матриці залишків Z_{ij} . В моделі, де на перетині рядків і стовпців вихідної таблиці знаходиться обсяг залишків Z_{ij} .

Модель 3. Для вирішення задачі з урахуванням групи конкурентів введемо відносний показник ступеня унікальності U_{ij} ВТМ і-го виду в аптеці A_j як відношення обсягів закупок:

$$U_{ij} = \frac{V_{ij}}{V_{ij} + \sum_s V_{ijs}}. \quad (2.1)$$

Якщо у ВТМ немає конкурентів в даній аптеці, то ступінь його унікальності дорівнює 1. В протилежному випадку U_{ij} приймає значення на проміжку (0;1).

Рекомендаційна оцінка обсягів закупки ВТМ формується для мережі аптек за косинусною мірою схожості аптек:

Для двох заданих векторів ознак, A та B , косинус подібності, $\cos(\theta)$, представляється за допомогою скалярного добутку та довжини, як

$$\text{подібність} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}, \quad (2.2)$$

де A_i та B_i — координати вектору A та B відповідно.

В розрахунку буде розглянутий метод так званий «користувач-користувач». Для створення нових рекомендацій для користувача цей метод намагається приблизно ідентифікувати користувача із найбільш схожим

«профілем взаємодій» (найближчі сусіди) для пропонування продуктів що є найбільш популярними серед цих сусідів (та ті що є новими для користувача). Цей метод фокусується на користувачах тому, що він представляє користувачів в залежності від їх взаємодій з продуктами та обчислює дистанцію між ними [16]. Представимо що ми хочемо створити рекомендацію для обраного користувача. По перше, кожний користувач може бути представлений як його вектор взаємодій з різними продуктами (його ліній у матриці взаємодій). Потім, ми можемо обчислити меру схожості між обраним користувачем та кожним іншим. Ця міра схожості полягає у тому, що два користувача зі схожими діями на тих самих продуктах мають бути визнані як схожі. Коли міри схожості для кожного користувача обчислені система може зберігати k -схожих-сусідів для користувачів та пропонувати їм найбільш популярні продукти серед них, обираючи лише з тих, з якими обраний користувач ще не робив взаємодій [17]. Треба зауважити, що при обчисленні міри схожості кількість «спільних дій» має бути обрана обережно, бо у багатьох ситуаціях, наприклад, коли якийсь користувач має лише одну спільну дію з обраним буде мати 100% схожість і буде обраним ближчим, ніж інший, що має 100 спільних взаємодій і погоджуються «лише» у 98%. Таким чином, два користувача вважаються схожими якщо вони взаємодіяли з багатьма продуктами схожим чином [18].

Метод користувач-користувач базується на пошуку схожих користувачів на основі взаємодій з продуктами. Якщо кожен користувач взаємодіяв лише з декількома продуктами, це робить метод достатньо чутливим до будь-яких збережених взаємодій (висока дисперсія). З іншого боку, якщо фінальна рекомендація базується лише на збережених взаємодіях користувачів із схожими інтересами, це дозволяє отримати більш персональний результат (низька похибка) [19].

Найбільш поширеною метрикою дистанції, що використовується у методах, оснований на пам'яті є «косинусна схожість». Ця метрика являє собою геометричне представлення рядку для користувача або колонки для продукту як

вектору. Для підходу що базується на користувачах схожість двох будь-яких з них вимірюється як косинус між відповідними їм векторами

Завдяки цьому можна передбачити рейтинг користувача u_i для i -того продукту на основі зваженої суми рейтингу всіх інших користувачів, де значення схожості між обраним користувачем та всіма іншими використовується як вага

Також має сенс зробити нормалізацію отриманих значень рейтингу за допомогою сумарної кількості всіх оцінок [20] Для початку було обрано групу ВТМ які входять в одну групу продажів а саме ті, що запускалися в обіг в один проміжок часу та мали схожий маркетинг [21].

Для генерації рекомендації як користувачі було обрано аптеки за високою схожістю з продажами між собою в гривні та упаковках.

З метою вирішення проблеми холодного старту користувачів, тобто ситуації, коли дані про нового користувача відсутні, було вирішено обрати впорядкування за популярністю. Популярність розглядається як середнє арифметичне всіх продаж в аптеках , даного ВТМ.

Висновки до розділу 2

У розділі зроблено огляд статей рекомендаційних систем для аптечних мереж. Система наведена в англійській статті хоч і є більш точною, проте має вади, серед яких є проблема накопичення вхідних даних, проблема в відсутності великої вибірки повних, актуальних та поновлювальних даних для навчання та удосконалення нейронних мереж, що є критичним для рекомендації саме власних торгових марок.

Переваги моделі із другої статті - це збільшення точності і можливості більш гнучкого налаштування на різні групи клієнтів. До недоліків слід віднести меншу інтерпретованість та більшу складність реалізації і підтримки.

Формалізовано постановку задачі у вигляді трьох моделей:

Модель 1. Вихідні дані формуються у вигляді матриці обсягів продажів V_{ij} .

Модель 2. Вихідні дані формуються у вигляді матриці залишків Z_{ij} .

Модель 3. Для вирішення задачі прогнозування продажів з урахуванням групи конкурентів

РОЗДІЛ 3.

ЕКСПЕРИМЕНТИ З ГЕНЕРАЦІЇ РЕКОМЕНДАЦІЙ

3.1 Генерація рекомендацій на основі розробленої рекомендаційної системи на об'єм продажів при виході на ринок ВТМ

Для проведення експериментів з моделлю рекомендаційної системи будемо використовувати табличний процесор Excel.

Для моделі 1, було відібрано групу аптек як найбільш подібних за продажами товарів ВТМ. Далі сформовано групу ВТМ як регулярно продаваних та які поступили в продаж водночас. В таблиці 3.1 наведено групу товарів та аптек з продажами товарів в упаковках.

Таблиця 3.1 – Модель 1, Продажі ВТМ в упаковках

Код товару	Продажі Втм аптеках	Код Аптеки									
	Товар	1237	2343	1518	1417	1105	1617	2507	1418	836	515
1060216	Витамин С		294	65	122	40	175	106	139	141	116
1056209	ДИКЛОФЕНАК ГЕЛЬ		40	32	47	44	44	46	60	28	40
1054522	УГОЛЬ АКТИВИР		239	626	413	705	406	440	467	82	501
1057206	КАЛЕНДУЛЫ ЦВЕТКИ		9	8	6	5	3	0	3	4	5
1058623	НИФУРОКСАЗИД СУСП		1	2	2	6	1	5	4	8	1
1059437	ПЛАСТЫРЬ ПЕРЦОВЫЙ		23	39	15	0	32	35	16	21	24
1059878	ТЕСТ Д/ОПР.БЕРЕМ.		19	5	22	21	19	21	20	1	35
1060327	ПЕРЕКИСЬ ВОДОРОДА		12	23	5	16	18	2	21	15	17
1061292	ФИТОЧАЙ ПОЧЕЧНЫЙ		1	6	1	2	3	6	7	10	5
1061293	ФИТОЧАЙ СЛАБИТЕЛЬНЫЙ		7	9	10	0	6	4	7	11	17
1061298	ШИПОВНИКА ПЛОДЫ		5	0	6	2	4	1	4	3	3
1061301	ЛИСТЬЯ СЕННЫ		1	1	1	1	3	1	0	0	0
1061466	ГЕПАЙС КАПС		3	0,25	2,5	0	4	3,5	16	27,38	4
1061618	МАГНИЙ В6		0,2	1	1	4	1,2	7,2	4,4	8,6	11
вектор			382,6	632	434,3	708,1	446,3	456,9	492,5	170,8	518,3

Червоний стовбець виокремлює аптеку в бік якої відбуватиметься розрахунок, сірим виділено рядки ВТМ для яких буде згенерована рекомендація.

За умовою аптека сприймається як нова для якої відсутні дані продажів ВТМ. Постає питання про вирішення проблеми холодного старту, для цього розраховую середнє арифметичне всіх продаж в аптеках , даного ВТМ, результати наведено в таблиці 3.2.

Таблиця 3.2 – Модель 1, Вирішення проблеми холодного старту для продажів

Продажі Втм аптеках		Код Аптеки									
Код товару	Товар	1237	2343	1518	1417	1105	1617	2507	1418	836	515
1060216	Витамин С	133,1	294	65	122	40	175	106	139	141	116
1056209	ДИКЛОФЕНАК ГЕЛЬ		40	32	47	44	44	46	60	28	40
1054522	УГОЛЬ АКТИВИР	431	239	626	413	705	406	440	467	82	501
1057206	КАЛЕНДУЛЫ ЦВЕТКИ	4,778	9	8	6	5	3	0	3	4	5
1058623	НИФУРОКСАЗИД СУСП		1	2	2	6	1	5	4	8	1
1059437	ПЛАСТЫРЬ ПЕРЦОВЫЙ	22,78	23	39	15	0	32	35	16	21	24
1059878	ТЕСТ Д/ОПР.БЕРЕМ.	18,11	19	5	22	21	19	21	20	1	35
1060327	ПЕРЕКИСЬ ВОДОРОДА	14,33	12	23	5	16	18	2	21	15	17
1061292	ФИТОЧАЙ ПОЧЕЧНЫЙ		1	6	1	2	3	6	7	10	5
1061293	ФИТОЧАЙ СЛАБИТЕЛЬНЫЙ	7,889	7	9	10	0	6	4	7	11	17
1061298	ШИПОВНИКА ПЛОДЫ		5	0	6	2	4	1	4	3	3
1061301	ЛИСТЬЯ СЕННЫ	0,889	1	1	1	1	3	1	0	0	0
1061466	ГЕПАЙС КАПС		3	0,25	2,5	0	4	3,5	16	27,38	4
1061618	МАГНИЙ В6	4,289	0,2	1	1	4	1,2	7,2	4,4	8,6	11
вектор		452,4	382,6	632	434,3	708,1	446,3	456,9	492,5	170,8	518,3

З використанням методу «користувач-користувач» розраховано рекомендацію в таблиці 3.3.

Таблиця 3.3 – Модель 1, Рекомендація

Рекомендація	
Вітамін С	130,40
ДИКЛОФЕНАК ГЕЛЬ	38,33
УГОЛЬ АКТИВИР	443,74
КАЛЕНДУЛЫ ЦВЕТКИ	4,72
НИФУРОКСАЗИД СУСП	2,88
ПЛАСТЫРЬ ПЕРЦОВЫЙ	22,86
ТЕСТ Д/ОПР.БЕРЕМ.	18,62
ПЕРЕКИСЬ ВОДОРОДА	14,34
ФИТОЧАЙ ПОЧЕЧНЫЙ	3,98
ФИТОЧАЙ СЛАБИТЕЛЬНЫЙ	7,83
ШИПОВНИКА ПЛОДЫ	2,76
ЛИСТЬЯ СЕННЫ	0,91
ГЕПАЙС КАПС	5,50
МАГНИЙ В6	4,24

Наразі потрібно перевірити метод відносно фактичних значень продажів, в таблиці 3.4 порівняно розрахунок холодного старту та фактичні значення.

Таблиця 3.4 – Модель 1, Точність при холодному старті

Точність при холодному старті		
Вітамін С 500мг №30 жев.таб. апельсин СТМ 43705	130,40	-2%
УГОЛЬ АКТИВИР	443,74	3%
КАЛЕНДУЛЫ ЦВЕТКИ	4,72	-1%
ПЛАСТЫРЬ ПЕРЦОВЫЙ	22,86	0%
ТЕСТ Д/ОПР.БЕРЕМ.	18,62	3%
ПЕРЕКИСЬ ВОДОРОДА	14,34	0%
ФИТОЧАЙ СЛАБИТЕЛЬНЫЙ	7,83	-1%
ЛИСТЬЯ СЕННЫ	0,91	3%
МАГНИЙ В6	4,24	-1%

Найбільша похибка 3%, що є великим, але задовільним значенням, що свідчить про виявлений зв'язок методу розрахунку та фактичних значень, далі перевірено похибку самої рекомендації в таблиці 3.5.

Таблиця 3.5 – Модель 1, Похибка рекомендації

Рекомендація		похибка
ДИКЛОФЕНАК ГЕЛЬ	43	12%
НИФУРОКСАЗИД СУСП	3	4%
ФИТОЧАЙ ПОЧЕЧНЫЙ	4	1%
ШИПОВНИКА ПЛОДЫ	3	9%
ГЕПАЙС КАПС	5,75	5%

Найбільша похибка 12%, що є за великим значенням, його не можна сприймати як вірну рекомендацію, при цьому 4 значення з похибкою менше 10% є задовільними.

Далі відбувається розрахунок моделі 2, рекомендації для залишку який необхідно підтримувати в аптеці для задоволення продажів, первинні залишки ВТМ наведені нижче в таблиці 3.6.

Таблиця 3.6 – Модель 2, Залишки ВТМ в аптеках

Залишки ВТМ на аптеках											
Код товару	Товар	1237	2343	1518	1417	1105	1617	2507	1418	836	515
1060216	Витамин С		164	50	89	10	87	26	35	43	24
1056209	ДИКЛОФЕНАК ГЕЛЬ		32	16	30	20	30	10	34	19	12
1054522	УГОЛЬ АКТИВИР		124	209,5	236	0	367	151	141	57	4644
1057206	КАЛЕНДУЛЫ ЦВЕТКИ		7	6	6	7	6	9	4	6	5
1058623	НИФУРОКСАЗИД СУСП		6	3	2	10	3	3	4	9	5
1059437	ПЛАСТЫРЬ ПЕРЦОВЫЙ		37	18	11	7	55	36	14	16	5
1059878	ТЕСТ Д/ОПР.БЕРЕМ.		20	13	13	8	27	8	5	8	24
1060327	ПЕРЕКИСЬ ВОДОРОДА		15	8	8	0	36	11	15	15	4
1061292	ФИТОЧАЙ ПОЧЕЧНЫЙ		3	4	4	4	7	5	8	6	2
1061293	ФИТОЧАЙ СЛАБИТЕЛЬНЫЙ		5	5	5	3	14	3	7	12	8
1061298	ШИПОВНИКА ПЛОДЫ		5	3	4	3	5	3	4	4	3
1061301	ЛИСТЬЯ СЕННЫ		3	3	3	2	4	3	4	4	2
1061466	ГЕПАЙС КАПС		3,12	3,37	4,62	4	4,25	10,13	4,62	10	6,62
1061618	МАГНИЙ В6		3,2	3,6	3	3	5,6	2,4	4	7,8	8,8
вектор			213,2	217,5	255	28,72	385,5	159,1	151,4	80,58	4644

Також за попередньою умовою аптека сприймається як нова для якої відсутні дані продажів ВТМ. Для вирішення проблеми холодного старту розраховую середнє арифметичне всіх залишків в аптеках , даного ВТМ, результати наведено в таблиці 3.7.

Таблиця 3.7 – Модель 2, Вирішення проблеми холодного старту для залишків

Залишки Втм на аптеках											
Код товару	Товар	1237	2343	1518	1417	1105	1617	2507	1418	836	515
1060216	Витамин С	116	164	50	89	10	87	26	35	43	24
1056209	ДИКЛОФЕНАК ГЕЛЬ		32	16	30	20	30	10	34	19	12
1054522	УГОЛЬ АКТИВИР	501	124	209,5	236	0	367	151	141	57	4644
1057206	КАЛЕНДУЛЫ ЦВЕТКИ	5	7	6	6	7	6	9	4	6	5
1058623	НИФУРОКСАЗИД СУСП		6	3	2	10	3	3	4	9	5
1059437	ПЛАСТЫРЬ ПЕРЦОВЫЙ	24	37	18	11	7	55	36	14	16	5
1059878	ТЕСТ Д/ОПР.БЕРЕМ.	35	20	13	13	8	27	8	5	8	24
1060327	ПЕРЕКИСЬ ВОДОРОДА	17	15	8	8	0	36	11	15	15	4
1061292	ФИТОЧАЙ ПОЧЕЧНЫЙ		3	4	4	4	7	5	8	6	2
1061293	ФИТОЧАЙ СЛАБИТЕЛЬНЫЙ	17	5	5	5	3	14	3	7	12	8
1061298	ШИПОВНИКА ПЛОДЫ		5	3	4	3	5	3	4	4	3
1061301	ЛИСТЬЯ СЕННЫ	0	3	3	3	2	4	3	4	4	2
1061466	ГЕПАЙС КАПС	4	3,125	3,375	4,625	4	4,25	10,13	4,625	10	6,625
1061618	МАГНИЙ В6	11	3,2	3,6	3	3	5,6	2,4	4	7,8	8,8
вектор		516,7	213,2	217,5	255	28,72	385,5	159,1	151,4	80,58	4644

З використанням методу «користувач-користувач» розраховано рекомендацію в таблиці 3.8.

Таблиця 3.8 – Модель 2, Рекомендація

Рекомендація	
Вітамін С	60,36
ДИКЛОФЕНАК ГЕЛЬ	19,78
УГОЛЬ АКТИВИР	779,66
КАЛЕНДУЛЫ ЦВЕТКИ	6,11
НИФУРОКСАЗИД СУСП	3,71
ПЛАСТЫРЬ ПЕРЦОВЫЙ	23,37
ТЕСТ Д/ОПР.БЕРЕМ.	14,64
ПЕРЕКИСЬ ВОДОРОДА	13,65
ФИТОЧАЙ ПОЧЕЧНЫЙ	4,31
ФИТОЧАЙ СЛАБИТЕЛЬНЫЙ	7,27
ШИПОВНИКА ПЛОДЫ	3,35
ЛИСТЬЯ СЕННЫ	3,21
ГЕПАЙС КАПС	5,61
МАГНИЙ В6	4,77

Також потрібно перевірити метод відносно фактичних значень залишків, в таблиці 3.9 нижче порівняно розрахунок холодного старту та фактичні значення.

Таблиця 3.9 – Модель 2, Точність при холодному старті

Точність при холодному старті		
Вітамін С	60,36	3%
УГОЛЬ АКТИВИР	779,66	18%
КАЛЕНДУЛЫ ЦВЕТКИ	6,11	-2%
ПЛАСТЫРЬ ПЕРЦОВЫЙ	23,37	6%
ТЕСТ Д/ОПР.БЕРЕМ.	14,64	5%
ПЕРЕКИСЬ ВОДОРОДА	13,65	9%
ФИТОЧАЙ СЛАБИТЕЛЬНЫЙ	7,27	6%
ЛИСТЬЯ СЕННЫ	3,21	3%
МАГНИЙ В6	4,77	4%

Найбільша похибка 18%, що є за великим значенням, його не можна сприймати як вірну рекомендацію, при цьому 8 значення з похибкою менше 10% є задовільними. Даля в таблиці 3.10 розрахована похибка рекомендації.

Таблиця 3.10 – Модель 2, Рекомендація та її похибка

Рекомендація		похибка
ДИКЛОФЕНАК ГЕЛЬ	43	117%
НИФУРОКСАЗИД СУСП	3	-19%
ФИТОЧАЙ ПОЧЕЧНЫЙ	4	-7%
ШИПОВНИКА ПЛОДЫ	3	-9%
ГЕПАЙС КАПС	5,75	2%

Найбільша похибка 117%, що є за великим значенням, його не можна сприймати як вірну рекомендацію, при цьому 3 значення з похибкою менше 10% є задовільними.

Далі відбувається розрахунок моделі 3, рекомендації засновані на продажах ВТМ у відношенні до групи конкурентів, результатом буде прогноз фармацевтичної заміни, що є ключовим показником ефективності маркетингу та управлінських рішень, що відображає яку долю ринку ВТМ займає у аптечній мережі в групі схожих товарів. Первинні дані розраховані за формулою в таблиці 3.11.

Таблиця 3.11 – Модель 3, Вхідні дані

Продажі Втм аптеках		Код Аптеки									
Код товару	Товар	1237	2343	1518	1417	1105	1617	2507	1418	836	515
1060216	Витамин С		0,96	0,70	0,83	0,14	0,90	0,68	0,84	0,84	0,56
1056209	ДИКЛОФЕНАК ГЕЛЬ		0,70	0,65	0,75	0,47	0,70	0,72	0,83	0,85	0,60
1054522	УГОЛЬ АКТИВИР		1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,42	1,00
1057206	КАЛЕНДУЛЫ ЦВЕТКИ		1,00	1,00	0,86	0,56	1,00	0,00	1,00	1,00	0,56
1058623	НИФУРОКСАЗИД СУСП		1,00	0,50	0,67	0,26	0,50	0,71	0,33	0,53	0,11
1059437	ПЛАСТЫРЬ ПЕРЦОВЫЙ		0,77	0,78	0,42	0,00	0,67	0,92	0,55	0,95	0,51
1059878	ТЕСТ Д/ОПР.БЕРЕМ.		0,86	0,83	0,81	0,48	0,90	0,88	0,74	0,17	0,85
1060327	ПЕРЕКИСЬ ВОДОРОДА		1,00	0,72	1,00	0,34	0,90	0,18	0,62	0,75	0,63
1061292	ФИТОЧАЙ ПОЧЕЧНЫЙ		0,20	0,35	0,08	0,04	0,33	0,43	0,50	0,43	0,21
1061293	ФИТОЧАЙ СЛАБИТЕЛЬНЫЙ		0,88	0,90	0,63	0,00	0,60	0,36	0,50	0,79	0,85
1061298	ШИПОВНИКА ПЛОДЫ		0,45	0,00	0,75	0,15	0,50	0,25	0,44	0,60	0,18
1061301	ЛИСТЬЯ СЕННЫ		0,07	0,05	0,17	0,02	0,14	0,13	0,00	0,00	0,00
1061466	ГЕПАЙС КАПС		0,04	0,00	0,04	0,00	0,04	0,08	0,26	0,25	0,05
1061618	МАГНИЙ В6		0,02	0,07	0,06	0,07	0,05	0,22	0,21	0,22	0,30
вектор			2,788	2,434	2,501	1,41	2,522	2,14	2,357	2,381	2,066

Також за умовою аптека сприймається як нова для якої відсутні дані ВТМ. Для вирішення проблеми холодного старту розраховую середнє арифметичне всіх показників в аптеках , даного ВТМ, результати наведено в таблиці 3.12 .

Таблиця 3.12 – Модель 3, Вирішення проблеми холодного старту для фармацевтичної заміни

Продажі Втм аптеках		Код Аптеки									
Код товару	Товар	1237	2343	1518	1417	1105	1617	2507	1418	836	515
1060216	Витамин С	0,72	0,96	0,70	0,83	0,14	0,90	0,68	0,84	0,84	0,56
1056209	ДИКЛОФЕНАК ГЕЛЬ		0,70	0,65	0,75	0,47	0,70	0,72	0,83	0,85	0,60
1054522	УГОЛЬ АКТИВИР	0,94	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,42	1,00
1057206	КАЛЕНДУЛЫ ЦВЕТКИ	0,77	1,00	1,00	0,86	0,56	1,00	0,00	1,00	1,00	0,56
1058623	НИФУРОКСАЗИД СУСП		1,00	0,50	0,67	0,26	0,50	0,71	0,33	0,53	0,11
1059437	ПЛАСТЫРЬ ПЕРЦОВЫЙ	0,62	0,77	0,78	0,42	0,00	0,67	0,92	0,55	0,95	0,51
1059878	ТЕСТ Д/ОПР.БЕРЕМ.	0,73	0,86	0,83	0,81	0,48	0,90	0,88	0,74	0,17	0,85
1060327	ПЕРЕКИСЬ ВОДОРОДА	0,68	1,00	0,72	1,00	0,34	0,90	0,18	0,62	0,75	0,63
1061292	ФИТОЧАЙ ПОЧЕЧНЫЙ		0,20	0,35	0,08	0,04	0,33	0,43	0,50	0,43	0,21
1061293	ФИТОЧАЙ СЛАБИТЕЛЬНЫЙ	0,61	0,88	0,90	0,63	0,00	0,60	0,36	0,50	0,79	0,85
1061298	ШИПОВНИКА ПЛОДЫ		0,45	0,00	0,75	0,15	0,50	0,25	0,44	0,60	0,18
1061301	ЛИСТЬЯ СЕННЫ	0,06	0,07	0,05	0,17	0,02	0,14	0,13	0,00	0,00	0,00
1061466	ГЕПАЙС КАПС		0,04	0,00	0,04	0,00	0,04	0,08	0,26	0,25	0,05
1061618	МАГНИЙ В6	0,13	0,02	0,07	0,06	0,07	0,05	0,22	0,21	0,22	0,30
вектор		2,788	2,434	2,501	1,41	2,522	2,14	2,357	2,381	2,066	

З використанням методу «користувач-користувач» розраховано рекомендацію в таблиці 3.13 .

Таблиця 3.13 – Модель 3, Рекомендація фармацевтичної заміни

Рекомендація	
Витамин С	0,72
ДИКЛОФЕНАК ГЕЛЬ	0,61
УГОЛЬ АКТИВИР	0,94
КАЛЕНДУЛЫ ЦВЕТКИ	0,79

Продовження таблиці 3.13

Рекомендація	
НИФУРОКСАЗИД СУСП	0,45
ПЛАСТЫРЬ ПЕРЦОВЫЙ	0,62
ТЕСТ Д/ОПР.БЕРЕМ.	0,74
ПЕРЕКИСЬ ВОДОРОДА	0,69
ФИТОЧАЙ ПОЧЕЧНЫЙ	0,25
ФИТОЧАЙ СЛАБИТЕЛЬНЫЙ	0,62
ШИПОВНИКА ПЛОДЫ	0,33
ЛИСТЬЯ СЕННЫ	0,06
ГЕПАЙС КАПС	0,07
МАГНИЙ В6	0,13

Потрібно перевірити метод відносно фактичних значень залишків, в таблиці 3.14 порівняно розрахунок холодного старту та фактичні значення.

Таблиця 3.14 – Модель 3, Точність при холодному старті для фармацевтичної заміни

Точність при холодному старті		
Витамин С	0,72	1%
УГОЛЬ АКТИВИР	0,94	1%
КАЛЕНДУЛЫ ЦВЕТКИ	0,79	2%
ПЛАСТЫРЬ ПЕРЦОВЫЙ	0,62	0%
ТЕСТ Д/ОПР.БЕРЕМ.	0,74	1%
ПЕРЕКИСЬ ВОДОРОДА	0,69	2%
ФИТОЧАЙ СЛАБИТЕЛЬНЫЙ	0,62	2%
ЛИСТЬЯ СЕННЫ	0,06	0%
МАГНИЙ В6	0,13	-1%

Найбільша похибка 2% що є за задовільними значенням, його можна сприймати як вірне. Далі в таблиці 3.15 розрахована похибка рекомендації.

Таблиця 3.15 – Модель 3, Рекомендація фармацевтичної заміни та її похибка

Рекомендація		похибка
ДИКЛОФЕНАК ГЕЛЬ	0,65	6%
НИФУРОКСАЗИД СУСП	0,60	33%
ФИТОЧАЙ ПОЧЕЧНЫЙ	0,50	98%
ШИПОВНИКА ПЛОДЫ	0,38	15%
ГЕПАЙС КАПС	0,08	12%

Найбільша похибка 98%, що є за великим значенням, його не можна сприймати як вірну рекомендацію, при цьому лише одне значення з похибкою менше 10% є задовільними.

3.2 Порівняння результатів експериментів з результатами прогнозування за існуючою методикою.

У підсумку було побудовано рекомендаційну модель за декількома параметрами, результати розрахунків зібрано у таблицю 3.16.

Таблиця 3.16 – Результати розрахунків

Аптека Код 1237	Залишки Модель1	Факт	Продажі Модель1	Факт	Продажі Модель2	Факт
ВитаминС	60	139	130	139	0,67	0,72
ДИКЛОФЕНАКГЕЛЬ	20	43	38	43	0,65	0,61
УГОЛЬАКТИВИР	780	445	444	445	1,00	0,94
КАЛЕНДУЛЫЦВЕТКИ	6	5	5	5	1,00	0,79
НИФУРОКСАЗИДСУСП	4	3	3	3	0,60	0,45
ПЛАСТЫРЬПЕРЦОВЫЙ	23	25	23	25	0,68	0,62
ТЕСТД/ОПР.БЕРЕМ.	15	19	19	19	0,63	0,74
ПЕРЕКИСЬВОДОРОДА	14	15	14	15	0,65	0,69
ФИТОЧАЙПОЧЕЧНЫЙ	4	4	4	4	0,50	0,25
ФИТОЧАЙСЛАБИТЕЛЬНЫЙ	7	8	8	8	0,80	0,62
ШИПОВНИКАПЛОДЫ	3	3	3	3	0,38	0,33
ЛИСТЬЯСЕННЫ	3	1	1	1	0,05	0,06
ГЕПАЙСКАПС	6	6	6	6	0,08	0,07
МАГНИЙВ6	5	4	4	4	0,16	0,13

Далі результати рекомендаційної системи необхідно порівняти з фактичними даним, результати розрахунків буди округлені до 1 упаковки для залишків та продажів, та до двох знаків після коми для моделі фармацевтична заміна, розрахунок зібрано в таблицю 3.17.

Таблиця 3.17 – Відхилення рекомендації від фактичних значень

Аптека Код 1237	Відхилення модель Залишки 1	Відхилення модель Продажі 1	Відхилення модель Фармацевтична заміна
ДИКЛОФЕНАК ГЕЛЬ	115%	13%	-6%
НИФУРОКСАЗИД СУСП	-25%	0%	-25%
ФИТОЧАЙ ПОЧЕЧНЫЙ	0%	0%	-50%
ШИПОВНИКА ПЛОДЫ	0%	0%	-13%
ГЕПАЙС КАПС	0%	0%	-13%

За першою моделлю для трьох з п'яти ВТМ відхилення 0%, що є хорошим результатом, два значення з великим відхиленням є невірними. За другою моделлю чотири з п'яти значень з відхиленнями 0%, що свідчить про високу точність моделі для прогнозування продажів, відхилення третьої моделі у відсоткових значеннях є завеликими для впевненого прогнозування лише у Диклофенаку похибка 6% , Шипшини плоди та Гепайс мають похибку в 13%, інші значення є незадовільними для сприйняття як вірних. При цьому відхилення у третій моделі відбувається лише у бік завищення прогнозних значень відносно фактичних у порівнянні з прогнозом за існуючим методом на всі аптеки в середньому є фармацевтична заміна по Диклофенак гелю в 40% від групи товарів. Проте відібрана аптека отримала рекомендацію в 65% фармацевтичної заміни від групи товарів і фактично виконала 61% фармацевтичної заміни. Це означає що для аптеки розрахунок рекомендаційної системи був більш точним ніж середній розрахунок за всіма аптеками.

Висновки до розділу 3

У розділі на основі зібраних даних було розраховано рекомендацію для аптеки за різними параметрами, такими як залишки втм, продажі втм, фармацевтична заміна втм в групі товарів. Отримані результати були зібрані та порівняні з фактичними значеннями. В цілому рекомендаційна модель є достатньо точною для того, щоб вважати прогноз достовірним. Також було порівняно результати рекомендації з існуючим методом прогнозування на підприємстві за одним з ВТМ а саме Диклофенак гель. Рекомендація надана рекомендаційною системою при перевірці навіть попри відхилення від фактичних даних, є більш точною ніж існуючий прогноз. Хоч побудова рекомендаційної системи буде значно ресурсо-затратна ніж існуючий прогноз, це збільшить точність розрахунків ключових параметрів, таких як продажі, залишки та фармацевтична заміна.

ВИСНОВКИ

В магістерському дипломі розглянуто теоретичні основи рекомендаційних систем та їх алгоритмів.

В розділі 1 виконано аналіз існуючих алгоритмів колаборативної фільтрації. Серед сучасних систем особливе місце займають алгоритми колаборативної фільтрації, тому їх основним різновидам та їх підґрунтя приділено значну частину розділу. Розглянуто ідеї, закладені у основі user-user, item-item та matrix factorization алгоритмів колаборативної фільтрації.

Зібрано дані за кількість аптек в мережі та розраховано їх темп приросту, зібрано дані про продажі ВТМ та його конкурентів за достатній для прогнозування період, виконано розрахунки темпу приросту продажів групи конкурентів та ВТМ.

За існуючою моделлю прогнозування на підприємстві розраховано прогноз продажів для групи в цілому та ВТМ окремо. Зібрано дані залишків ВТМ в мережі та залишків на складах постачальника для прогнозування замовлення у виробника. В результаті отримано прогноз продажів, що може використовуватись для формування необхідних залишків на аптеках та прогноз замовлення для відправки виробнику ВТМ.

У розділі 2 наведено огляд сучасних рекомендаційних систем для аптек. Обґрунтовано вибір моделі «користувач-користувач» в розробці рекомендаційної системи для аптечної мережі АНЦ. Розглянуто і пояснено причини впровадження алгоритмів «користувач-користувач» у роботу рекомендаційної системи.

У розділі 3 на основі зібраних даних було виконано ряд експериментів з генерації рекомендації на основі розробленої рекомендаційної системи, розраховано рекомендацію для аптеки за різними параметрами, такими як: залишки ВТМ, продажі ВТМ, фармацевтична заміна ВТМ в групі товарів. Отримані результати були зібрані та порівняні з фактичними значеннями.

Порівняно результати рекомендації з існуючим методом прогнозування на підприємстві за одним з ВТМ, а саме Диклофенак гель.

Практична реалізація виконана на даних аптечної мережі АНЦ.

Запропонована модель рекомендаційної системи для централізованої аптечної мережі, дозволяє вирішувати задачу з виведення товарів власної торгової марки на ринок, а також виконання функції для прогнозування аптекам об'ємів продажів та залишків.

ПЕРЕЛІК ПОСИЛАНЬ

1. **Єрошенко О. С.** Рекомендаційні системи в дистанційному навчанні – [Електронний ресурс] – Режим доступу: <https://nauka-online.com/ua/publications/informatsionnye-tehnologii/2018/6/rekomendatelnye-sistemy-v-distantionnom-obuchenii/>
2. **Негрей М.В.** Компаративний аналіз ефективності рекомендаційних систем в маркетингу [Текст] / Негрей М.В., Гнот Т.В // Вісник Хмельницького національного університету – 2017 – № 5 – С. 278-286.
3. **Codina V. A** Система рекомендацій для семантичного WEB [Електронний ресурс] – Victor Codina and Luigi Ceccaroni // Матеріали Міжнародного симпозиуму з розподілених обчислень та штучного інтелекту – 2010. – Режим доступу: <http://www.lsi.upc.edu/~vcodina/papers/dcai10>.
4. Методи на основі моделі для рекомендаційних систем. [Текст] / Ю. Стех, М. Лобур, В.Арцибасов. – Пер. с англ. // Міжнародний журнал передових досліджень у галузі обчислювальної техніки та технологій – 2015. – т.4. – С.3061-3066.
5. **Салахутдінов Р.** Імовірнісна матрична факторизація. У досягненнях в системах обробки нейронної інформації. [Текст] / Р. Салахутдінов і А. Мніх. – Пер. с англ. 2008
6. **Vozalis E.** Аналіз рекомендаційних систем [Електронний ресурс] – Emmanouil Vozalis, Konstantinos G. Margaritis –Пер. с англ //9 Всегрецька конференція з інформатики. – 2003. – Режим доступу: <https://pdfs.semanticscholar.org/6543/1c3a27f01e70bb5e4b1e68fc6a91249f6114.pdf>.
7. **Hnot T.** Порівняння системи рекомендацій: найкращий алгоритм [Електронний ресурс] Hnot T. , Gladkikh T. – Режим доступу : <https://www.softserveinc.com/en-us/tech/blogs/recommender-comparison-algorithm/>

8. **Сегаран. Т.** Программируем коллективный разум. [Текст] / Т. Сегаран. – Пер. с англ. – СПб: Символ-Плюс, 2008. – 368 с.
9. **Meleshko, Ye.** Комп'ютерна модель віртуальної соціальної мережі з рекомендаційною системою, – [Електронний ресурс] // Інноваційні технології та наукові рішення для галузей – 2019 – Режим доступу: <https://doi.org/10.30837/2522-9818.2019.8.080> .
10. **Adomavicius G.** Системи рекомендацій з урахуванням контексту [Текст] / G. Adomavicius, A. Tuzhilin — Пер. с англ. // Довідник з рекомендаційних систем. – 2015. – Р. 191–226.
11. Коллаборативная фильтрация [Електронний ресурс] /– Режим доступу: <http://www.machinelearning.ru/wiki/images/archive/9/95/2014041318417%21Voron-ML-CF.pdf>.
12. **Глибовець М. М.** Розробка агентної рекомендаційної системи віртуального університету засобами JADE [Текст] М. М. Глибовець, О. В. Конюшенко // Проблеми програмування. – 2015. – № 1. – С. 104–115.
13. Сайт довідник з маркетингу [Електронний ресурс] – Режим доступу до ресурсу: https://www.marketch.ru/marketing_dictionary/o/inventory_turnover/.
14. **Franz H.** Граф згортка нейронної мережі для аптечної системи рекомендацій крос-продажу [Електронний ресурс] – Franz Hell, Yasser Taha, Gereon Hinz, Sabine Heibei, Harald Müller, Alois Knoll // Режим доступу: <https://www.mdpi.com/2078-2489/11/11/525>
15. **Бакурова А.В.** Модель рекомендаційної системи для аптечної мережі [Текст] / Бакурова А.В., Ропало Г.М. // Збірник матеріалів 7 Міжнародна науково-технічна конференція «ІНФОРМАЦІЙНІ СИСТЕМИ ТА ТЕХНОЛОГІЇ ІСТ-2018»
16. **Ekstrand M. D.** Системи рекомендацій для спільної фільтрації [Електронний ресурс] / Michael D. Ekstrand, John T. Riedl, Joseph A. Konstan. – Режим доступу: <http://files.grouplens.org/papers/FnT%20CF%20Recsys%20Survey.pdf>. – Title from the screen.

17. **Джонс Т.** Рекомендательные системы: Введение в подходы и алгоритмы. [Электронный ресурс] – Режим доступа: <http://www.ibm.com/developerworks/ru/library/os-recommender1/index.htm>.
18. **Валько Д.В.** Рекомендательная система на основе интеллектуального анализа наукометрического профиля исследователя. [Текст] / Д.В. Валько // Программные продукты и системы. – 2018. – Т 31. – № 3. – 275 с
19. **Городецкий В.И.** Онтологии и персонификация профиля пользователя в рекомендующих системах третьего поколения [Текст] / В.И. Городецкий, О.Н. Тушканова // Онтология проектирования – 2014 – № 3 (13) – Самара : Предприятие «Новая техника» – С. 7–31.
20. **Глибовець А. М.** Агенти для рекомендацій у колаборативних середовищах [Текст] А. М. Глибовець, С. С. Гороховський, А. А. Піка // Наукові праці МДУ ім. Петра Могили. Комп'ютерні технології. – 2010. – Вип. 121, т. 34. – С. 142–151.
21. **Noy N. F.** Розробка онтології 101: Посібник зі створення вашої першої онтології [Електронний ресурс] / Natalya F. Noy and Deborah L. McGuinness // Технічний звіт лабораторії Stanford Knowledge Systems KSL-01-05 і технічний звіт Stanford Medical Informatics SMI-2001-0880 – 2001. – Режим доступу: http://protege.stanford.edu/publications/ontology_development/ontology101.html.