

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ  
ІМЕНІ ВАДИМА ГЕТЬМАНА**

**Навчально-науковий інститут  
«Інститут інформаційних технологій в економіці»**

**Кафедра інформаційних систем в економіці**

**ОСВІТНЬО-ПРОФЕСІЙНА ПРОГРАМА  
«Системи штучного інтелекту»**

галузь знань	12 Інформаційні технології
спеціальність	122 Комп'ютерні науки

Форма навчання: очна (денна)

**КВАЛІФІКАЦІЙНА МАГІСТЕРСЬКА РОБОТА**

на тему: «Інтелектуальна система аналізу ринку роздрібної торгівлі»

здобувача **Сіденко Владислава Вячеславовича**

Науковий керівник: к.е.н., доцент Гордієнко І.В.

**Робота допущена до захисту перед екзаменаційною  
комісією з атестації здобувачів вищої освіти (ЕК)**

Завідувач кафедри: к.е.н., доцент Тішков Б.О.

**Київ 2024**

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ  
ІМЕНІ ВАДИМА ГЕТЬМАНА

Навчально-науковий інститут «Інститут інформаційних технологій в економіці»  
Кафедра інформаційних систем в економіці

ОСВІТНЬО-ПРОФЕСІЙНА ПРОГРАМА

«Системи штучного інтелекту»

галузь знань 12 Інформаційні технології  
спеціальність 122 Комп'ютерні науки

ПОГОДЖЕНО:

Керівник проєктної групи (гарант)  
освітньо-професійної програми

\_\_\_\_\_ Рамазанов С.К.  
« \_\_\_\_ » \_\_\_\_\_ 202\_ р.

ЗАТВЕРДЖУЮ:

Завідувач кафедри інформаційних  
систем в економіці

\_\_\_\_\_ Тішков Б.О.  
« \_\_\_\_ » \_\_\_\_\_ 202\_ р.

ІНДИВІДУАЛЬНЕ ЗАВДАННЯ

здобувача вищої освіти Сіденка Владислава Вячеславовича

(прізвище, ім'я, по батькові)

очної (денної) форми навчання

очної (денної), заочної, дистанційної

на підготовку кваліфікаційної магістерської роботи

на тему: «Інтелектуальна система аналізу ринку роздрібної торгівлі»

Тему затверджено наказом ректора Університету від «23» січня 2024 р. № 126-ст

Кваліфікаційна магістерська робота виконується на матеріалах  
ринку роздрібної торгівлі України

План кваліфікаційної магістерської роботи

Розділ I Дослідження та аналіз підходів до створення предметної області СШІ  
(назва розділу)

Розділ II Характеристика СШІ та постановка задачі  
(назва розділу)

Розділ III Розроблення проєктних рішень  
(назва розділу)

Об'єкт дослідження: сучасні тенденції розвитку ритейлу та застосування інтелектуальних систем аналізу ринку роздрібної торгівлі

Предмет дослідження: розробка інтелектуальної системи аналізу ринку роздрібної торгівлі з використанням сучасних методів та технологій

Мета кваліфікаційної магістерської роботи: розроблення інтелектуальної системи аналізу ринку роздрібної торгівлі з метою підвищення конкурентоспроможності та ефективності управління у ритейлі

**Конкретні завдання, які здобувач повинен виконати для досягнення поставленої мети:**

**У розділі I** Описати предметну галузь ринку роздрібною торгівлі. Проаналізувати існуючі інформаційні системи управління роздрібною торгівлею, зокрема, системи аналізу та інтелектуальні системи. Подати постановку проблеми та описати задачі, що підлягають розв'язанню в умовах інтелектуальної системи аналізу ринку роздрібною торгівлі. Обґрунтувати вибір підходів і технологій для проектування та розроблення інтелектуальної системи аналізу ринку роздрібною торгівлі

---



---



---



---

**У розділі II** Подати характеристику об'єкта дослідження. Розробити структуру і надати характеристику інтелектуальної системи аналізу ринку роздрібною торгівлі та її компонентів. Описати методи і моделі процесів і елементів системи штучного інтелекту для аналізу ринку роздрібною торгівлі

---



---



---



---

**У розділі III** Спроекувати базу знань, користувацький інтерфейс інтелектуальної системи аналізу ринку роздрібною торгівлі. Розробити інформаційне забезпечення СІІІ. Обґрунтувати вибір та розташування комплексу технічних засобів. Описати структуру та складові програмного та організаційно-економічного забезпечення інтелектуальної системи аналізу ринку роздрібною торгівлі

---



---



---



---

**Завдання підготував**  
**науковий керівник** \_\_\_\_\_  
(підпис)

**І.В. Гордієнко**  
(ініціали, прізвище)

« \_\_\_\_ » \_\_\_\_\_ 202\_ р.

**Завдання одержав**  
**здобувач** \_\_\_\_\_  
(підпис)

**Сіденко В.В.**  
(ініціали, прізвище)

« \_\_\_\_ » \_\_\_\_\_ 202\_ р.

## Реферат

Кваліфікаційна магістерська робота містить 72 сторінок, 10 таблиць, 18 рисунків, список використаних джерел з 45 найменувань, додатки.

### **«Інтелектуальна система аналізу ринку роздрібно́ї торгівлі»**

*Об'єктом дослідження* кваліфікаційної магістерської роботи є сучасні тенденції розвитку ритейлу та застосування інтелектуальних систем аналізу ринку роздрібно́ї торгівлі.

*Предметом дослідження* є розробка інтелектуальної системи аналізу ринку роздрібно́ї торгівлі з використанням сучасних методів та технологій.

*Мета і завдання дослідження.* Розроблення інтелектуальної системи аналізу ринку роздрібно́ї торгівлі з метою підвищення конкурентоспроможності та ефективності управління у ритейлі.

Для розроблення інтелектуальної системи необхідно виконати такі задачі:

- дослідити теоретичні відомості щодо роздрібно́ї торгівлі, визначити основні особливості ритейлу в порівнянні із звичайним магазином роздрібно́ї торгівлі;
- здійснити аналіз стану сучасних інформаційних систем, які використовують в роздрібно́ї торгівлі, проаналізувати сучасні тенденції розвитку систем штучного інтелекту в ритейлі;
- описати характеристику розвитку ритейлу та застосування інтелектуальних систем аналізу ринку у роздрібно́ї торгівлі;
- визначити методи та моделі, які необхідні для аналізу ринку роздрібно́ї торгівлі;
- здійснити моделювання бази знань для інтелектуальної системи аналізу ринку роздрібно́ї торгівлі;
- розробити інформаційне забезпечення інтелектуальної системи;
- описати програмне і технічне забезпечення, необхідне для розроблення і впровадження інтелектуальної системи аналізу ринку роздрібно́ї торгівлі.

*Теоретична, методична та практична значущість отриманих результатів.* Під час дослідження було досліджено предметну область ринку роздрібно́ї торгівлі, проаналізовано існуючі інформаційні технології в роздрібно́ї торгівлі, досліджені методи і моделі, які використовуються при управлінні асортиментом продукції в роздрібно́ї торгівлі, описана структура і характеристика інтелектуальної системи аналізу ринку роздрібно́ї торгівлі та її компонентів, змодельовано процес підготовки даних для кластеризації та здійснено кластерний аналіз. Основними практичними результатами є запропонований проєкт інтелектуальної системи аналізу ринку роздрібно́ї торгівлі, який дозволить підвищити якість прийняття управлінських рішень при плануванні асортименту продукції на підприємствах роздрібно́ї торгівлі.

Рік виконання кваліфікаційної магістерської роботи – 2024.

Рік захисту роботи – 2024.

*Ключові слова:* ритейл, асортимент продукції, кластеризація, база знань, інтелектуальна система.

## Abstract

The qualifying master's thesis contains 72 pages, 10 tables, 18 figures, a list of used sources from 45 titles, appendices.

### **"Intelligent retail market analysis system"**

*The object of research* of the qualifying master's thesis is modern trends in retail development and the use of intelligent systems for retail market analysis.

*The subject of the research* is the development of an intelligent system of retail market analysis using modern methods and technologies.

*The purpose and tasks of the research.* Development of an intelligent system of retail market analysis in order to increase competitiveness and management efficiency in retail.

To develop an intelligent system, it is necessary to perform the following tasks:

—research theoretical information about retail trade, determine the main features of retail in comparison with an ordinary retail store;

—carry out an analysis of the state of sophisticated information systems used in retail trade, analyze modern trends in the development of artificial intelligence systems in retail;

—describe the characteristics of retail development and the use of intelligent market analysis systems in retail trade;

—to determine the methods and models that are necessary for the analysis of the retail market;

—carry out modeling of the knowledge base for the intelligent system of retail market analysis;

—to develop the information support of the intellectual system;

—describe the software and technical support necessary for the development and implementation of an intelligent system of retail market analysis.

*Theoretical, methodical and practical significance of the obtained results.* During the research, the subject area of the retail market was investigated, the existing information technologies in retail were analyzed, the methods and models used in the management of the range of products in retail were investigated, the structure and characteristics of the intelligent system for the analysis of the retail market and its components were described, the process was simulated data preparation for clustering and cluster analysis. The main practical results are the proposed project of an intelligent system of retail market analysis, which will improve the quality of managerial decision-making when planning the assortment of products at retail enterprises.

The year of completion of the qualifying master's thesis is 2024.

The year of job protection is 2024.

*Keywords: retail, product range, clustering, database, intelligent system.*

## В і д г у к

про кваліфікаційну магістерську роботу  
здобувача навчально-наукового інституту «Інститут інформаційних технологій в  
економіці»

освітньо-професійної програми «Системи штучного інтелекту»

Сіденка Владислава Вячеславовича

на тему «Інтелектуальна система аналізу ринку роздрібно́ї торгівлі»

1. **Актуальність теми:** Актуальність теми кваліфікаційної магістерської роботи пов'язана з тим, що ефективне управління торгівлею неможливе без урахування результатів аналізу ринкової інформації. Одним з сучасних підходів до проведення такого аналізу є використання інтелектуальних методів і засобів, що було реалізовано у проєкті інтелектуальної системи здобувача

2. **Позитивні риси кваліфікаційної магістерської роботи:** робота містить результати вивчення здобувачем предметної області роздрібно́ї торгівлі, зокрема, ритейлу, дослідження можливостей використання інтелектуальних систем у ритейлі, відображає використані здобувачем в інтелектуальній системі сучасні методи машинного навчання та кластерного аналізу, демонструє володіння інформаційно-технологічним інструментарієм розробки ІС, зокрема мовами програмування Java і Python.

3. **Наявність самостійних розробок автора:** описана структура і характеристика інтелектуальної системи аналізу ринку роздрібно́ї торгівлі та її компонентів, розроблена структура бази даних, змодельовано процес підготовки даних для кластеризації та здійснено кластерний аналіз.

4. **Цінність теоретичних висновків та практичних рекомендацій:** розроблений проєкт інтелектуальної системи аналізу ринку роздрібно́ї торгівлі дасть змогу підвищити якість прийняття управлінських рішень при плануванні асортименту продукції на підприємствах роздрібно́ї торгівлі.

5. **Наявність недоліків:** у тексті роботи у незначній кількості зустрічаються описки

6. **Загальна оцінка кваліфікаційної магістерської роботи та її допущення до захисту перед ЕК:** кваліфікаційна магістерська робота відповідає встановленим вимогам методичних вказівок щодо структури, обсягу та змісту, є самостійно виконаною роботою здобувача Сіденко В.В. і рекомендується до захисту з високою оцінкою.

Науковий керівник:

к. е. н., доцент кафедри ІСЕ

Гордієнко І.В.

« 16 » травня 2024 р.

## Рецензія

на кваліфікаційну магістерську роботу  
здобувача вищої освіти

**Сіденка Владислава Вячеславовича**

(прізвище, ім'я, по батькові)

Тема **«Інтелектуальна система аналізу ринку роздрібної торгівлі»**

Актуальність теми кваліфікаційної магістерської роботи і доцільність її розроблення: Впровадження методів і технологій штучного інтелекту є актуальним і перспективним. Використання інтелектуальної системи аналізу ринку роздрібної торгівлі підвищить ефективність управління запасами, аналізу попиту, плануванню асортименту, а також вдосконаленню процесів обслуговування клієнтів та роботи з даними. Саме тому тема Сіденка В.В. є актуальною і своєчасною.

Якість проведеного дослідження В роботі здобувач дослідив дослідження теоретичних відомостей щодо використання інтелектуальних систем в роздрібній торгівлі, проаналізував сучасний ринок інформаційних систем, що застосовують у ритейлі та виокремивши важливі задачі, які будуть вирішені за допомогою інтелектуальної системи. Аналіз проведений якісно і ґрунтовно, з використанням сучасних публікацій українських та закордонних фахівців.

Позитивні риси кваліфікаційної магістерської роботи: Запропоновані проєктні рішення щодо застосування кластерного аналізу при аналізі асортименту продукції. Запропонований авторський підхід до проєктування користувацького інтерфейсу та забезпечувальних підсистем (інформаційного, програмного та технічного забезпечення)

Зауваження: В роботі є стилістичні і граматичні помилки

Практична значимість висновків і рекомендацій Основними практичними результатами є запропонований проєкт інтелектуальної системи аналізу ринку роздрібної торгівлі, який дозволить підвищити якість прийняття управлінських рішень при плануванні асортименту продукції на підприємствах роздрібної торгівлі. Запропоновані рішення можуть бути використані при проєктуванні подібної системи.

Місце роботи та посада рецензента Український державний університет імені Михайла Драгоманова, завідувач кафедри інформаційних технологій і програмування

Науковий ступінь, учене звання: к.пед.н. доцент

Василь ЄФІМЕНКО

Підпис завідувача:

(посада, підпис)



## ЗМІСТ

<b>ВСТУП</b> .....	3
<b>РОЗДІЛ 1 ДОСЛІДЖЕННЯ ТА АНАЛІЗ ПІДХОДІВ ДО СТВОРЕННЯ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ АНАЛІЗУ РИНКУ РОЗДРІБНОЇ ТОРГІВЛІ</b> .....	5
1.1 Дослідження предметної області. Збір інформації та вивчення матеріалів з теми кваліфікаційної магістерської роботи. ....	5
1.2 Аналіз існуючих інтелектуальних систем аналізу ринку роздрібноЇ торгівлі.....	12
1.2.1 Інформаційні системи в роздрібній торгівлі .....	12
1.2.2 Цифрові технології в роздрібній торгівлі .....	18
1.2.3 Використання систем штучного інтелекту в ритейлі.....	19
1.3 Постановка проблеми та формування задач .....	23
1.4 Обґрунтування вибору підходів і технологій для проектування та створення інтелектуальноЇ системи аналізу ринку роздрібноЇ торгівлі .....	24
<b>РОЗДІЛ 2 ХАРАКТЕРИСТИКА ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ АНАЛІЗУ РИНКУ РОЗДРІБНОЇ ТОРГІВЛІ ТА ПОСТАНОВКА ЗАДАЧІ</b> .....	28
2.1. Характеристика об'єкта дослідження.....	28
2.2 Структура і характеристика інтелектуальноЇ системи аналізу ринку роздрібноЇ торгівлі та її компонентів. ....	32
2.3 Методи, моделі і моделювання процесів і елементів складних систем ...	37
2.3.1 Методи дослідження й синтезу компонент систем штучного інтелекту.....	37
2.3.2 Моделі та методи оптимізації в системах штучного інтелекту .....	39
2.3.3 Методи та моделі управління в системах штучного інтелекту .....	41
<b>РОЗДІЛ 3 РОЗРОБЛЕННЯ ПРОЕКТНИХ РІШЕНЬ</b> .....	43
3.1 Моделювання та проектування бази знань прийняття інтелектуальних рішень і управління .....	43
3.2 Розроблення користувацького інтерфейсу. Елементи та структура. ....	49
3.3 Проектування забезпечувальних підсистем СШ. Реалізація системи... ..	52
3.3.1 Інформаційне забезпечення. ....	52
3.3.2 Програмне забезпечення .....	62
3.3.3 Технічне забезпечення .....	64
3.3.4 Організаційно- економічне забезпечення .....	65
<b>ВИСНОВКИ</b> .....	66
<b>СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ</b> .....	67
<b>ДОДАТКИ</b> .....	72

## ВСТУП

**Актуальність проблеми.** У сучасній економіці роздрібна торгівля відіграє важливу роль, забезпечуючи споживачам доступ до різноманітного асортименту товарів та послуг. Крім того, роздрібні торговці є основними посередниками у прямих продажах від виробників до кінцевих споживачів, що сприяє загальному економічному розвитку. Інформаційні та інтелектуальні системи в цьому сегменті грають важливу роль, допомагаючи роздрібним торговцям у вирішенні різних завдань. Вони сприяють ефективнішому управлінню запасами, аналізу попиту, плануванню асортименту, а також вдосконаленню процесів обслуговування клієнтів та роботі з даними. Такі системи дозволяють підтримувати оптимальний рівень запасів, прогнозувати попит, адаптувати асортимент до змін у попиті та поведінці споживачів, що сприяє підвищенню конкурентоспроможності та ефективності роздрібної торгівлі.

Саме тому актуальним є питання створення інтелектуальної системи аналізу ринку роздрібної торгівлі.

**Об'єктом дослідження** кваліфікаційної магістерської роботи є сучасні тенденції розвитку ритейлу та застосування інтелектуальних систем аналізу ринку роздрібної торгівлі.

**Предметом дослідження** є розробка інтелектуальної системи аналізу ринку роздрібної торгівлі з використанням сучасних методів та технологій.

**Мета і завдання дослідження.** Розроблення інтелектуальної системи аналізу ринку роздрібної торгівлі з метою підвищення конкурентоспроможності та ефективності управління у ритейлі.

Для розроблення інтелектуальної системи необхідно виконати такі **завдання**:

- дослідити теоретичні відомості щодо роздрібної торгівлі, визначити основні особливості ритейлу в порівнянні із звичайним магазином роздрібної торгівлі;

- здійснити аналіз стану сучасних інформаційних систем, які використовують в роздрібній торгівлі, проаналізувати сучасні тенденції розвитку систем штучного інтелекту в ритейлі;
- описати характеристику розвитку ритейлу та застосування інтелектуальних систем аналізу ринку у роздрібно́ї торгівлі;
- визначити методи та моделі, які необхідні для аналізу ринку роздрібно́ї торгівлі;
- здійснити моделювання бази знань для інтелектуальної системи аналізу ринку роздрібно́ї торгівлі;
- розробити інформаційне забезпечення інтелектуальної системи<sup>4</sup>
- описати програмне і технічне забезпечення, необхідне для розроблення і впровадження інтелектуальної системи аналізу ринку роздрібно́ї торгівлі.

**Матеріал дослідження.** Робота виконана на основі матеріалів наукових публікацій та публікацій з мережі інтернет.

**Теоретична, методична та практична значущість отриманих результатів.** Під час дослідження було досліджено предметну область ринку роздрібно́ї торгівлі, проаналізовано існуючі інформаційні технології в роздрібно́ї торгівлі, досліджені методи і моделі, які використовуються при управлінні асортиментом продукції в роздрібно́ї торгівлі, описана структура і характеристика інтелектуальної системи аналізу ринку роздрібно́ї торгівлі та її компонентів, змодельовано процес підготовки даних для кластеризації та здійснено кластерний аналіз. Основними практичними результатами є запропонований проєкт інтелектуальної системи аналізу ринку роздрібно́ї торгівлі, який дозволить підвищити якість прийняття управлінських рішень при плануванні асортименту продукції на підприємствах роздрібно́ї торгівлі. Запропоновані рішення можуть бути використані про проєктуванні подібно́ї системи.

**Структура роботи.** Робота містить 72 сторінки, 10 таблиць, 18 рисунків, список використаних джерел з 45 найменувань, додатки.

## РОЗДІЛ 1 ДОСЛІДЖЕННЯ ТА АНАЛІЗ ПІДХОДІВ ДО СТВОРЕННЯ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ АНАЛІЗУ РИНКУ РОЗДРІБНОЇ ТОРГІВЛІ

### 1.1 Дослідження предметної області. Збір інформації та вивчення матеріалів з теми кваліфікаційної магістерської роботи.

Роздрібна торгівля в сучасній економіці відіграє важливу роль, надаючи споживачам доступ до широкого асортименту товарів та послуг. Крім того, роздрібні торговці є ключовими посередниками у прямих продажах від виробників до кінцевих споживачів, що сприяє загальному економічному розвитку [1].

У сучасних бізнес-колах найчастіше для роздрібної торгівлі використовують термін **ритейл**. **Ритейл** – це запозичене з англійської мови слово, яке служить для визначення процесу продажу продукції або послуг кінцевим споживачам [2].

Ритейл (від англ. retail - "роздрібна торгівля") - це сектор бізнесу, який включає в себе продаж товарів або послуг кінцевим споживачам через мережу магазинів, інтернет-магазинів, торгових центрів та інші форми роздрібної торгівлі [1].

Ритейлер — це роздрібний торговець, який продає товари або послуги кінцевому споживачеві. Тобто ритейлер - це останнє ланцюгове з'єднання між виробником і покупцем. Важливо зазначити, що розмір торгової точки не має значення. Це означає, що ритейлером може бути як маленький продуктовий магазин або прилавок поблизу будинку, так і великий супермаркет. Крім того, під цим визначенням також потрапляють люди, які займаються продажами, не маючи певного місцевого розташування торгової точки, такі як представники мережевого маркетингу та комівояжери [2].

Ринок роздрібної торгівлі є ключовим елементом економіки будь-якої країни та забезпечує великий обсяг робочих місць.

Окрім прямих продажів, він включає в себе різноманітні сегменти бізнесу, включаючи малі, середні та великі підприємства. Крім функції прямих продажів, ритейлери є великим сектором бізнесу, який забезпечує зайнятість великої кількості працівників. Тобто ритейл охоплює практично всі рівні бізнесу: від малого і середнього до великого масштабу.

Вважається, що поняття «роздріб» і «ритейл» відрізняються.

Ритейл передбачає доступ до широкого асортименту товарів та послуг у одному магазині. За допомогою спеціальних продажних прийомів працівники магазину можуть продати більше товарів, ніж планував придбати покупець. Це приносить велику користь клієнтам, адже завдяки цьому вони можуть обирати серед товарів та послуг та знаходити саме те, що їм потрібно [3].

Термін «ритейл» прийшов до нас із заходу, фактично описує діяльність роздрібною торгівлі, спрямовану на широке коло споживачів. Це не просто магазин для купівлі продуктів харчування чи одягу. Його головною особливістю є пропозиція широкого асортименту товарів, зібраних на одній торговій площадці.

Переваги такого формату для споживачів включають:

- великий вибір товарів;
- можливість придбати продукцію різних брендів і категорій в одному місці;
- легке порівняння цін;
- доступність товарів за акціями та програмами лояльності.

Ця галузь бізнесу працює на оборот і обсяг. Вона використовує безліч маркетингових заходів, щоб привернути клієнтів і мотивувати їх повертатися знову. Сюди входять акції, розпродажі, програми лояльності, бонуси та інші способи взаємодії з клієнтами. Ціни на товари тут зазвичай нижчі, ніж у невеликих роздрібних точках, завдяки індивідуальним контрактам з постачальниками, яким вигідно постачати товари великими партіями [4].

Тобто у роздрібній торгівлі не кожна точка може бути названа ритейлом. Цю назву можуть отримати компанії, які відповідають таким характеристикам:

- великі обсяги продажів;
- високий рівень відвідування покупцями;
- програми лояльності;
- автоматизація процесів;
- використання технологічних новинок;
- розвинена маркетингова стратегія.

Для організації повноцінного управління всіма процесами ритейлових мереж часто потрібні власні офіси та торгові площадки. У роздрібній торгівлі, як правило, обходяться невеликим штатом співробітників.

Ще однією особливістю є випуск продукції під власними торговими марками. Це дозволяє знизити витрати на виробництво та отримувати прибуток завдяки внутрішнім оборотам. Наприклад, мережа "Сільпо" має торгові марки "Премія" та "Повна Чаша", а "АТБ" - "Своя лінія" [4].

Завдання у роздрібній торгівлі зазвичай менш глобальні, тому не потребують значних кадрових ресурсів.

До ритейлових мереж відносяться торгові мережі, гіпермаркети та супермаркети. Кожен з них має свої особливості та переваги. Конкуренція мотивує їх до постійного розвитку, що є додатковим плюсом для звичайного покупця.

Основні особливості ритейлу в порівнянні із звичайним магазином роздрібною торгівлі [3, 5]:

1. Розташування магазину на вигідній території, яка ретельно підбирається. Наприклад, ювелірні магазини розташовуються поруч із центром міста або будинками забезпечених клієнтів.

2. Легкість знаходження необхідного товару. У гіпермаркетах товари розміщуються за конкретними групами та виробниками, що спрощує пошук необхідної категорії продуктів. Для онлайн-магазинів застосовується зручна фільтрація або можливість знаходження товару за ключовими словами у пошуковому рядку.

3. Використання відповідного торгового або холодильного обладнання, яке виконує кілька функцій. Наприклад, холодильники використовуються не лише для підтримки низької температури, але й як охолоджувальні вітрини.

4. Застосування маркетингових прийомів, таких як акції та знижки, що допомагають привернути увагу покупців.

5. Можливість самообслуговування клієнтами, що сприяє швидкому здійсненню покупок та не вимагає великої кількості працівників для обслуговування клієнтів.

6. Автоматизація зберігання товарів, закупок та обліку за допомогою комп'ютерних програм, що значно спрощує логістичні та бізнес-процеси.

Існують різні види ритейлу, які відрізняються за форматом торгівлі, особливостями товарів та локацією [4, 5]:

– **Продуктовий ритейл**: торгівля продуктами харчування на великих торгових площах. В основному це супермаркети, що є найбільш затребуваними серед аудиторії будь-який час.

– **Непродуктовий** або **non-food ритейл**: торгівля непродовольчими товарами в широкому асортименті, такими як взуття, цифрова техніка, одяг, косметика та інше.

– **Вуличний або стріт-ритейл**: класична форма торгівлі на перших поверхах будівель вулиць з великою прохідністю. Головна умова успішної роботи - відсутність близьких великих торгових центрів та конкурентів.

– **Мережевий ритейл**: мережі магазинів, які об'єднані загальним стилем і концепцією, з одним власником та єдиною системою логістики, закупівлі та управління.

– **Інтернет-ритейл**: торгівля через інтернет, яка здобуває все більшу популярність та обороти.

У кожному з цих напрямків існують свої особливості, але всі вони мають спільну мету - отримання прибутку за рахунок обороту та масових продажів товарів роздрібною торгівлі.

Зазвичай, ланцюжок роздрібних постачань складається з таких етапів:

- виробник, що виготовляє продукцію;
- гуртовики або дистриб'ютори, які закупають товари безпосередньо у виробників;
- роздрібні торговці, які купують товари у гуртовиків та потім реалізують їх споживачам.

На кожному з цих етапів є власна маржа прибутку, врахована у вартості товару. Виробники визначають вартість виготовлення продукту, додаючи до неї відсоток прибутку, за яким продають товар гуртовикам. Гуртовики роблять те саме, додаючи відсоток прибутку до тієї ціни, за яку закупили продукт. Після цього роздрібні торговці додають свою власну маржу прибутку до вартості товару перед його продажом кінцевому споживачу [2].

**Бізнес-процеси у ритейлі.** Роздрібна торгівля, спрямована на масового споживача, визначається складними бізнес-процесами, які враховують особливості свого ринку. Кожна компанія у цьому сегменті має кілька напрямів діяльності та значний персонал, для якого розробляються чіткі правила та бізнес-процеси для забезпечення розвитку компанії.

Зростання у роздрібній торгівлі залежить від таких факторів, як:

- впровадження сучасних технологій;
- використання якісного обладнання;
- розробка стратегії розвитку;
- поетапне виконання бізнес-плану;
- наявність кваліфікованого персоналу.

Кожна компанія має свої бізнес-процеси, які потребують індивідуального підходу та уважного вивчення. Чим більший масштаб проекту, тим складніше та ширше будуються процеси впровадження.

Важливі технології у роздрібній торгівлі включають:

- класифікація товарів за категоріями;

- використання різних методів продажу;
- автоматизація процесів (облік, зберігання, закупівля);
- робота з великими обсягами товарів у роздріб.

Успіх торгової точки значною мірою залежить від сучасного обладнання, особливо якщо йдеться про продуктові магазини. Сучасні технології дозволяють створювати індивідуальний стиль, підбираючи обладнання за різними характеристиками, включаючи габарити, енергоефективність, функціональність та дизайн. Технології є ключовим показником успішності торгової точки, тому важливо використання сучасної інформаційної системи з елементами штучного інтелекту в ритейлі [3, 4].

Особливості роботи у сегменті ритейла можуть бути дуже практичними та динамічними. У цій галузі основна увага приділяється продажам товарів та послуг кінцевим споживачам, а також виконанню багатьох інших завдань, включаючи ведення складського обліку, організацію логістики, проведення маркетингових кампаній і т.п. Однак ведення бізнесу у ритейлі може стикнутися з рядом викликів, які потрібно враховувати для досягнення успіху:

1. Конкуренція. У ритейлі існує жорстка конкуренція, і кожен бізнес повинен знати, як виділитися серед інших компаній та залучити більше клієнтів.

2. Змінна споживча хвиля. Клієнти постійно змінюють свої потреби та вподобання. Ритейлери повинні слідкувати за змінами в поведінці та вподобаннях споживачів, щоб оптимізувати свій бізнес та задовольнити потреби клієнтів.

3. Вартість та окупність. Забезпечення низьких цін та хорошої окупності є основним викликом для будь-якої компанії у ритейлі. Компанії повинні працювати над підвищенням ефективності своїх бізнес-процесів, щоб знизити витрати та підвищити прибутковість.

4. Зростання споживчих очікувань. Компанії повинні забезпечувати все більш високий рівень сервісу та якості продуктів і послуг, щоб утримати своїх клієнтів та приваблювати нових [1].

Незважаючи на складності, ритейл також надає велику кількість можливостей для бізнесу:

1. Використання нових технологій. Ритейл може використовувати сучасні технології, такі як штучний інтелект, інтернет речей та аналітика даних, для поліпшення бізнес-процесів та підвищення ефективності.

2. Різноманітність каналів продажу. Ритейлери можуть продавати свої товари через різні канали, включаючи магазини, Інтернет, соціальні мережі, ринки тощо, що може збільшити охоплення клієнтів та прибутковість.

3. Нові категорії продуктів. Ритейлери можуть вводити нові категорії товарів та послуг, щоб розширити свій асортимент і поліпшити конкурентоспроможність свого бізнесу.

4. Глобалізація. Ритейлери можуть розширювати свій бізнес за межі своєї країни, що надає багато можливостей для зростання і розвитку.

Отже, робота у сегменті ритейла надає як виклики, так і можливості для бізнесу. Розуміння цих факторів може допомогти ритейлерам розробляти стратегії для збільшення прибутковості та задоволення потреб клієнтів.

## 1.2 Аналіз існуючих інтелектуальних систем аналізу ринку роздрібно́ї торгівлі

### 1.2.1 Інформаційні системи в роздрібно́ї торгівлі

Торгівля як сфера людської діяльності однієї з перших стала впроваджувати інформаційні технології. Торгівля, як одна з найстаріших сфер людської діяльності, першою почала використовувати інформаційні технології. Можна стверджувати, що саме тут вперше з'явилася «автоматизована система обліку».

У минулому податкова інспекція не мала контролю над касовою стрічкою, тому облік був необхідний самому підприємцю для відстеження роботи найманого продавця. Підприємець проходив через свої торгові точки, збирав виручку і перевіряв її відповідність касовій стрічці. У випадку розходження сум, він вирішував питання з продавцями.

З плином часу, із розвитком торгівлі, також розвивалися інформаційні технології. Звісно, деякі особи і сьогодні задовольняються касовими апаратами, які стали «розумнішими», простішими та зручнішими, проте суть процесу мало змінилася. Масштаби торгівлі, асортимент товарів, потік покупців - все це наразі значно перевищує обсяги минулих століть, і керувати торговельним підприємством без автоматизації стало дуже важко, а часом і неможливо.

Сучасну торгівлю можна структурувати за основними сегментами діяльності, як показано на рисунку 1.1 [6].

Розглянемо основні задачі, які вирішують інформаційні технології на кожній ділянці ритейлу.

**Оптова торгівля, дистрибуція.** Оптова торгівля та дистрибуція виступають як ключові ланки у ланцюгах постачання, забезпечуючи зв'язок між виробником і споживачем, які часто знаходяться на великій відстані один від одного, навіть у різних країнах.

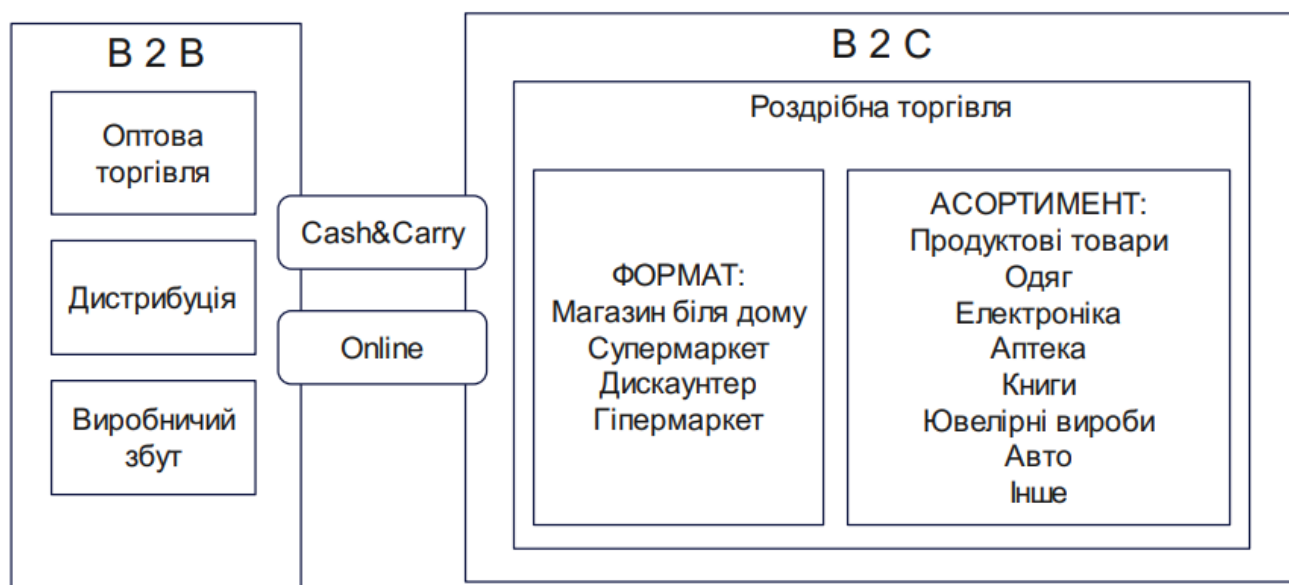


Рисунок 1.1 – Сегментування торгової галузі

Джерело: [6]

Лише найбільші глобальні виробники можуть організувати свою систему постачання так, щоб уникнути посередників. Діяльність оптовика або дистриб'ютора тепер включає в себе не лише простий процес купівлі та продажу, а складний набір завдань:

- вивчення та аналіз потреб клієнтів;
- приваблення нових та збереження постійних клієнтів;
- оптимізація та управління запасами на складі;
- управління замовленнями;
- координація поставок;
- контроль внутрішньоскладської та транспортної логістики.

Ці завдання вирішуються за допомогою товарно-облікових систем та торговельних модулів ERP-систем. Управління складською логістикою здійснюється за допомогою систем управління складами (WMS), а транспортної - за допомогою систем управління транспортом (TMS). Для зовнішнього спілкування використовуються спеціалізовані системи електронного обміну даними (EDI). Завдання підтримки відданості клієнтів та маркетингу вирішуються

засобами товарно-облікових систем або відповідних модулів ERP-систем та систем управління взаємовідносинами з клієнтами (CRM).

Тенденції розвитку оптової торгівлі та дистрибуції обумовлені як ринковими умовами, так і розвитком інформаційних технологій. Зростаюча конкуренція призводить до розширення спектру послуг та програм лояльності, при цьому цінова конкуренція поступається на другий план.

В таких умовах якість обслуговування клієнтів стає вирішальною конкурентною перевагою. На перший план виходять послуги, що покращують якість обслуговування, такі як:

- системи автоматизованого замовлення товару;
- інтеграція інформаційних систем продавця та клієнта;
- оперативне повідомлення клієнта про виконання його замовлення.

Інформаційні технології активно розвиваються в мобільному секторі, що означає надання працівникам доступу до автоматизованих робочих місць через планшети та смартфони. Інший важливий тренд - розвиток хмарних сервісів, які значно спрощують інтеграцію та розв'язання завдань інтеграції.

**Виробничий збут.** Збутові підрозділи виробничих підприємств, фактично, виступають першими ланками у ланцюгу постачання і, отже, розв'язують практично ті ж самі завдання, що й оптові та дистриб'юторські компанії. Основна відмінність полягає у тому, що вони співпрацюють не з зовнішніми постачальниками, а з власним підприємством. Тому для того, щоб виробництво могло планувати свою роботу на основі актуальних замовлень, робота виробничого збуту має бути автоматизована в рамках єдиної ERP-системи підприємства.

Вирішення, чи використовувати додатково WMS, TMS і CRM-системи, чи вбудовані модулі до ERP-систем зі спрощеними функціями, в порівнянні зі спеціалізованими рішеннями, залежить від конкретних обставин кожного випадку.

З серед значущих тенденцій розвитку збутової діяльності можна виділити прагнення виробників до управління своїми ланцюгами постачання, зниження витрат на них та максимізації власного прибутку. Тому інформаційні системи повинні вирішувати ряд інтеграційних завдань у всьому ланцюзі постачання.

**Роздрібна торгівля.** Роздрібна торгівля в ланцюгах постачання виявляється найбільш обмеженим сегментом. Жоден магазин не може вмістити весь асортимент товарів, що його поставляють або виробляють. Основне завдання ефективного ритейлера - розмістити на полицях саме ті товари, які приносять йому максимальний прибуток.

Роздрібні підприємства, завдяки своєму стратегічному положенню у ланцюгу постачання, встановлюють свої власні правила гри для всіх учасників ринку. Наприклад, у магазинах зазвичай не розраховуються одразу за поставлений товар. Відстрочка оплати принаймні на 40 днів є практикою, що фіксується майже у кожному договорі на постачання до роздрібною мережі. Розуміється, існують спеціальні служби, що відповідають за укладення таких угод та контроль за взаєморозрахунками за ними.

Дуже важливою для роздрібною торгівлі є організація логістики та управління нею. Частина логістичних завдань делегують постачальникам, а частину ритейлери вирішують самостійно, будуючи свої власні розподільні центри та транспортні служби.

У роздрібною торгівлі застосовуються три класи інформаційних систем [6]:

- фронт-офіс (Front-office) - вони займаються обслуговуванням покупців та працюють на спеціалізованих робочих місцях з необхідним обладнанням або на POS-терміналах.

- управління магазином (In-store solution) - це системи, що відповідають за облік товарів і грошей, управління цінами, запасами, замовленнями, персоналом, маркетинговими акціями та лояльністю покупців.

- бек-офіс (Back-office) - це системи, які вирішують широкий спектр завдань з обліку та управління торговельним підприємством, часто належать до класу ERP-систем. Вони також відповідають за управління взаємовідносинами з постачальниками та управління асортиментом у роздрібною мережі.

Роздрібна торгівля ставить досить високі вимоги до надійності інформаційних систем, оскільки багато магазинів працюють цілодобово, сім днів на тиждень. Час у цьому сегменті справді дорогоцінний, і кожен ритейлер розуміє,

що навіть кілька годин простою касової лінії чи транспорту може коштувати дорого. Інформаційні системи аналізують і прогнозують прибутковість товарів, надають інформацію для прийняття рішень про асортимент та управління витратами.

У роздрібній торгівлі є специфічні вимоги до інформаційних систем, залежно від сегменту. Наприклад, для аптек важливий облік серій лікарських препаратів та контроль їх якості, а для магазинів з будівельних матеріалів - облік мірного матеріалу і точна інформація про запаси.

Функціональне наповнення інформаційних систем для роздрібної торгівлі зазнало невеликих змін за останні 30 років, але їхні можливості розширюються завдяки розвитку технологій, включаючи мобільність та аналітичну обробку даних в реальному часі.

Важливо зазначити, що конкурентна боротьба в секторі роздрібної торгівлі стає все більш жорсткою, і малий бізнес повинен знайти своє місце в цьому середовищі. Інформаційні системи для малого бізнесу грають важливу роль у допомозі малим магазинам залишатися конкурентоспроможними, забезпечуючи їхню ефективну роботу за доступну ціну.

**Cash&Carry** (оптово-роздрібна торгівля). Інформаційні технології в роздрібній торгівлі об'єднують складні завдання як з сегменту B2B, так і з B2C. Оптові продажі в цьому сегменті можуть бути дрібними для малого бізнесу, і вони часто мало автоматизовані. Такі оптові продажі мало відрізняються від роздрібних продажів, але можуть мати свої особливості, які потребують функціональності товарооблікових систем для оптової торгівлі.

Роздрібні продажі відрізняються тим, що вони можуть включати персоналізований облік і угоди за договорами купівлі-продажу, а не лише продажі відповідно до публічних ofert, як у роздрібних магазинах. Тому інформаційні системи для Cash&Carry повинні мати функціональність, що відповідає системам автоматизації роздрібної торгівлі, а також здатність витримувати навантаження, подібні до оптових і роздрібних систем, і працювати безперервно.

Умови роботи в цьому сегменті можуть бути дуже екстремальними, і тому вимоги до надійності і відмовостійкості інформаційних систем такі ж високі, як і до систем роздрібної торгівлі, з працездатністю 24/7.

Онлайн-торгівля. Рішення про купівлю, що приймаються на підставі інформації, поданої в Інтернеті, відносяться до онлайн-торгівлі. Це може бути як покупка через інтернет-магазин, так і покупка, оформлена за допомогою телефонного дзвінка до продавця. Онлайн-торгівля практикується як у сегменті B2C, так і в сегменті B2B, проте їхні вимоги до інформаційних систем різняться.

У сегменті B2B важливим є управління взаєминами з клієнтами. Використання CRM-систем, інтегрованих з корпоративним сайтом, дозволяє контролювати весь цикл продажу, аналізувати ефективність рекламних кампаній та розраховувати вартість залучення клієнтів.

У сегменті B2C продавці зазвичай зосереджуються на пошуковій оптимізації та просуванні своїх сайтів, а також на підвищенні зручності для користувачів. Багато сайтів використовують готові системи управління вмістом сайту (CMS) з модулями для онлайн-торгівлі.

Незалежно від сегменту, інформаційні системи повинні бути інтегрованими та забезпечувати автоматизацію бек-офісу, а також оптимізовані для управління доставкою.

Ринок **онлайн-торгівлі** розвивається швидко, зокрема завдяки мобільному сектору Інтернету та популярності мобільних додатків. Кожен сегмент торгівлі має свої вимоги до інформаційних систем, але ринок програмного забезпечення пропонує рішення, що задовольняють ці вимоги.

## 1.2.2 Цифрові технології в роздрібній торгівлі

У сучасному світі роздрібні торговці активно використовують цифрові технології для оптимізації процесів продажів, збільшення клієнтської бази та розвитку власних брендів

Цифрові технології все більше впливають на ритейл, змінюючи існуючі бізнес-моделі та збільшуючи ефективність та прибутковість компаній. Ми можемо виділити кілька суттєвих тенденцій впровадження нових технологій у ритейл:

**Використання Big Data.** Технології аналізу великих обсягів даних можуть застосовуватися для підвищення ефективності магазинів та поліпшення якості обслуговування клієнтів. Big Data дозволяє збирати дані про звички, вподобання, покупки та інші параметри клієнта, що дозволяє точніше реагувати на їх запити та підвищити якість обслуговування.

**Використання мобільних додатків.** Мобільні додатки дозволяють спростувати процес покупки для клієнтів та підвищувати лояльність до бренду. З їх допомогою можна надавати клієнтам персональні знижки, промокоди та інші бонуси, що сприяє збільшенню кількості продажів та прибутку.

**Розширення онлайн-продажів.** Онлайн-продажі стають все більш популярними серед клієнтів, тому багато ритейлерів розширюють свій онлайн-асортимент та поліпшують користувацький інтерфейс для полегшення процесу покупки.

**Автоматизація процесів.** Цифрові технології дозволяють ритейлерам автоматизувати багато бізнес-процесів, такі як управління складом, облік працівників та розклад поставок, що дозволяє скоротити витрати та підвищити ефективність виробництва.

Отже, цифрові технології відіграють велику роль у ритейлі, змінюючи існуючі бізнес-моделі та відкриваючи нові можливості. Ритейлери можуть використовувати ці технології для підвищення ефективності та прибутковості

бізнесу, покращення якості обслуговування клієнтів та створення конкурентних переваг [1].

### **1.2.3 Використання систем штучного інтелекту в ритейлі**

Сьогодні в технологічну епоху штучний інтелект стає невід'ємною частиною різноманітних сфер, включаючи роздрібну торгівлю. Діджитал трансформує наше сприйняття шопінгу, а компанії активно використовують ШІ-технології на кожному етапі розповсюдження товарів. Штучний інтелект дозволяє ритейлерам ефективно взаємодіяти з клієнтами, забезпечуючи персоналізований підхід, передбачення попиту та автоматизовану комунікацію. Впровадження ШІ-рішень стає ключовим фактором успіху для великих роздрібних мереж.

Розглянемо найбільш вдалі приклади використання штучного інтелекту в роздрібній торгівлі:

- персоналізація;
- допомога у виборі;
- передбачення попиту
- динамічне ціноутворення
- чатботи та автоматизована комунікація.

#### **Персоналізація та допомога у виборі [7]**

За часів великих маркетплейсів, коли споживачеві треба обрати товар серед сотні тисяч пропозицій, допоможуть технології персоналізованого підбору. Покупець хоче, щоб бренди, з якими він взаємодіє, розуміли та передбачали його вподобання. Якщо споживачу складно зробити вибір, він може відмовитися від купівлі або перейти на іншу платформу.

Наприклад, на сайті INTERTOP впровадили персоналізований показ товарів з каталогу, відображення схожих речей на сторінці товару, а також підбір луків під особисті запити клієнта.

Персоналізація у каталозі працює на основі взаємодії користувача із сайтом (запити, кліки, перегляд сторінок, додавання у кошик), сервіс аналізує інформацію про всіх користувачів, які підтвердили збір куків, та адаптує показ товарів під їхні запити. Завдяки сервісу користувач може одразу побачити і перейти на сторінку схожого товару, якщо обраного немає або він йому не підходить. Обидва рішення дозволяють генерувати додатковий дохід. Коефіцієнт конверсії для сервісу з підбору схожих товарів становить приблизно 20%.

Наприклад, проблемою онлайн-платформ у сфері моди є відмінності розмірних сіток та підбір речей під власні параметри. Висока частка повернень для ритейлера — це зайві витрати ресурсів, а для клієнта — незадоволеність шопінгом і марнування часу. До того ж поширеною є проблема брекетингу (з англ. bracketing — практика, за якої споживач замовляє один і той самий товар у різних розмірах або кольорах для подальшої примірки та повернення).

Тому продавці пропонують сервіс підбору розмірів за фото. У картках товарів 20 найпопулярніших брендів з'явиться кнопка «Підібрати розмір». Щоб отримувати рекомендації, користувачеві достатньо один раз зазначити свою вагу та зріст і зробити два фото. За фото сервіс формує 3D-модель тіла, після чого порівнює отримані виміри з розмірною сіткою товару і підбирає максимально релевантний розмір. Таким чином споживач відкидає зайві сумніви щодо розміру та оминає небажану процедуру повернення.

### **Передбачення попиту та динамічне ціноутворення.**

Сучасні роздрібні мережі активно використовують технології штучного інтелекту для оптимізації різних процесів, зокрема аналізу цін та реагування на попит. Наприклад, в INTERTOP штучний інтелект застосовується для попереднього прогнозування попиту та оптимізації цін. Замість довготривалого процесу ручного аналізу, ШІ-сервіс враховує різноманітні фактори, включаючи аналіз ринку та історію продажів, і швидко визначає оптимальну ціну для кожного

товару. Такий підхід дозволяє ефективно виконувати завдання менеджера та досягати високих показників у продажах та прибутку.

Shein, китайський фешнритейлер, також використовує штучний інтелект для аналізу попиту та прогнозування модних тенденцій. За допомогою ШІ-рішень компанія отримує інформацію про попит на нові продукти з використанням даних з пошукових систем та соціальних мереж. Це дозволяє Shein швидко реагувати на зміни в попиті та пристосовувати свій ланцюг постачання до актуальних тенденцій. Наприклад, більше 800 дизайнерів Shein негайно реагують на популярність товарів та надсилають замовлення постачальникам через систему управління ресурсами підприємства (ERP) для виробництва та продажу. Цей підхід дозволяє компанії швидко задовольняти потреби клієнтів та мінімізувати витрати на операції.

#### **Чатботи та автоматизована комунікація.**

В сфері B2C-бізнесу чатботи стають одним з найпоширеніших інструментів ШІ. Вони беруть на себе функції менеджера під час спілкування з клієнтами, швидко надають базові консультації та допомагають у замовленні 24/7. Зараз область застосування чатботів постійно розширюється. Наприклад, вони використовуються для автоматизації комунікаційних процесів в HR, управління потоками інформації всередині компанії та прискорення підбору кандидатів на вакансії. Впровадження чатботів призвело до економії робочого часу, що становить понад 1000 годин на місяць.

На глобальному ринку спостерігається безліч цікавих прикладів використання чатботів, що виходять за рамки рутинних завдань менеджерів. Наприклад, Zalando та Google декілька років тому представили чатбота, який допомагає користувачам з вибором подарунків. Пошук подарунків через Google Assistant викликає запитання щодо віку та інтересів отримувача, а потім пропонує варіанти подарунків, які можна придбати на Zalando. Такі нові функції відображають зовсім інший рівень взаємодії користувача з онлайн-платформою.

Застосування інноваційних технологій, зокрема штучного інтелекту, у сфері роздрібного продажу продуктів є актуальним напрямом розвитку. Різні торгові

мережі активно впроваджують ці технології для оптимізації бізнес-процесів та покращення взаємодії з клієнтами [8].

Наприклад, Auchan Україна використовує ІІІ для першої візуалізації ідей та покращення досвіду покупців. Одним із прикладів є впровадження системи iBeacon, яка є інноваційною технологією в сфері роздрібною торгівлі. Ця система датчиків, розташованих всередині магазину, надсилає персоналізовані сповіщення клієнтам щодо акцій та пропозицій, залежно від їх місця перебування у магазині. Незважаючи на успішне впровадження цієї системи у Європі та США, Auchan Україна визнає, що український ринок може вимагати іншого підходу до персоналізованих пропозицій.

Також, мережа VARUS використовує ІІІ для різних цілей, зокрема, для оптимізації веб-сайту та створення контенту. Застосування ІІІ дозволяє генерувати текстові описи товарів та створювати ілюстрації для них, що ефективно оптимізує робочі процеси та зменшує витрати компанії.

Компанія Fozzy також активно використовує ІІІ для покращення своєї діяльності. Зокрема, вони розробили систему Kissa AI, яка використовує комп'ютерний зір для розпізнавання продуктів у ресторанному бізнесі. Ця система дозволяє автоматизувати процес замовлення та обслуговування гостей, що сприяє збільшенню швидкості обслуговування та покращенню досвіду клієнтів.

Хоча ІІІ має великий потенціал для застосування в різних сферах бізнесу, деякі ритейлери, такі як NOVUS, ще вивчають ці технології та не поспішають з їх впровадженням. Однак з розвитком цих технологій, вони можуть стати важливим елементом у покращенні бізнес-процесів та взаємодії з клієнтами у майбутньому.

**Перспективи розвитку ритейлу з використанням штучного інтелекту [9].**

Згідно з опитуванням Honeywell [10], за кордоном майже 6 із 10 ритейлерів очікують впровадження штучного інтелекту, машинного навчання (МН) та технологій комп'ютерного бачення (КБ) протягом наступного року з метою покращення досвіду купівель онлайн та в магазині.

Лише 3% опитаних сказали, що взагалі не використовують ці нові технології.

Майже половина (48%) респондентів назвали ІІ, МН і КБ найкращими технологіями, які, як очікується, матимуть значний вплив на галузь роздрібно́ї торгівлі протягом наступних трьох-п'яти років.

Основні напрямки розвитку ІІ в роздрібно́ї торгівлі:

Планування: штучний інтелект можна використовувати для аналізу даних із різних джерел (таких як дані про клієнтів, дані про продажі та дані з соціальних мереж), щоб робити прогнози щодо поведінки клієнтів і тенденцій на ринку. Це допоможе роздрібним мережам приймати більш обґрунтовані рішення про те, які продукти зберігати, як на них встановлювати ціну та як їх продавати.

Управління запасами: ІІ можна використовувати для оптимізації управління запасами. Наприклад, для прогнозування попиту на продукти, відстеження рівня запасів у режимі реального часу та автоматизації процесів зміни замовлень.

Чатботи: ІІ можна використовувати для створення чатботів, які відповідатимуть на запитання клієнтів, надаватимуть рекомендації та навіть завершуватимуть транзакції без допомоги людини», – каже Ямпольський.

Тож нові технології, зокрема штучний інтелект, допоможуть ритейлерам надавати персоналізований клієнтський досвід, оптимізувати роботу, підвищити продуктивність і ефективність, покращити управління запасами та запобігти шахрайству — а отже, покращити й фінансові показники компанії.

### **1.3 Постановка проблеми та формування задач**

У роздрібно́ї торгівлі постійно зростає конкуренція, а також змінюються уподобання споживачів та умови ринку. Це створює потребу у розробці та впровадженні інтелектуальних систем, які допомагатимуть аналізувати ринок, прогнозувати тенденції та розробляти стратегії для підприємств роздрібно́ї торгівлі. Однак, на сьогоднішній день, багато існуючих інформаційних систем не здатні забезпечити високу точність аналізу та прогнозування, що може призвести

до недооцінки потенційних можливостей або неправильного прийняття стратегічних рішень підприємствами.

Тому актуальним є створення інтелектуальної системи аналізу ринку роздрібною торгівлі, яка забезпечить підприємствам ефективний контроль над ринковою ситуацією та допоможе в прийнятті обґрунтованих стратегічних рішень.

Для побудови інтелектуальної системи сформулюємо наступні задачі:

1. Вивчити теоретичні основи роздрібною торгівлі, включаючи порівняння ритейлу з традиційними формами торгівлі, аналіз різних видів ритейлу, його властивості та бізнес-процеси.

2. Оцінити поточний стан інформаційних систем, що застосовуються в роздрібній торгівлі, та розглянути останні тенденції розвитку систем штучного інтелекту в цьому секторі.

3. Вивчити розвиток ритейлу та використання інтелектуальних систем. Визначити основні завдання, які вирішують дані системи.

4. Дослідити методи та моделі для аналізу ринку, визначити оптимальні методи та моделі, необхідні для проведення аналізу ринку в роздрібній торгівлі.

5. Створити базу знань для системи. Провести моделювання бази знань, яка буде використовуватися в інтелектуальній системі для аналізу ринку.

6. Розробити необхідне інформаційне забезпечення для інтелектуальної системи аналізу ринку роздрібною торгівлі.

7. Описати необхідне програмне та технічне забезпечення.

#### **1.4 Обґрунтування вибору підходів і технологій для проектування та створення інтелектуальної системи аналізу ринку роздрібною торгівлі**

Інтелектуальна система аналізу ринку роздрібною торгівлі є важливою складовою системи управління ритейлом, яка спрямована на покращення стратегічного та операційного процесів у роздрібній торгівлі. Її функції та можливості включають в себе комплексний аналіз ринкових тенденцій,

прогнозування попиту, визначення оптимального асортименту товарів, підвищення ефективності маркетингових стратегій та підтримку прийняття рішень на стратегічному та оперативному рівнях.

У рамках системи управління ритейлом, інтелектуальна система дозволить виконувати різноманітні завдання, зокрема, аналіз даних щодо споживчих уподобань, ринкових тенденцій та конкурентного середовища для підтримки ухвалення стратегічних рішень, таких як планування асортименту товарів, розробка маркетингових стратегій та ціноутворення.

Крім того, система забезпечить можливість виконання операційних завдань, таких як персоналізоване обслуговування клієнтів, управління запасами товарів, планування та моніторинг доставки товарів, а також аналіз ефективності продажів.

Важливо підкреслити, що інтелектуальна система є інтегрованою частиною стратегічного планування та операційного управління роздрібним торговим підприємством. Вона допомагає забезпечити взаємодію між різними підрозділами та виробляти обґрунтовані рішення на всіх рівнях управління для досягнення стратегічних та оперативних цілей бізнесу.

Існує кілька підходів до створення інтелектуальних систем аналізу ринку роздрібною торгівлі, які можуть використовуватися залежно від конкретних потреб і можливостей бізнесу [11, 12,13]:

1) Методи машинного навчання і штучного інтелекту: використання алгоритмів машинного навчання для аналізу великих обсягів даних щодо споживчих уподобань, попиту на товари, поведінки клієнтів та інших факторів, що впливають на роздрібний ринок. Цей підхід дозволяє автоматизувати процеси аналізу даних і робити прогнози з високою точністю [14].

2) Системи бізнес-аналітики і аналітичні платформи: використання спеціалізованих програмних засобів для збору, обробки та аналізу даних з метою виявлення тенденцій ринку, ідентифікації нових можливостей та визначення оптимальних стратегій продажу [15].

3) Розробка власних алгоритмів та моделей: створення власних алгоритмів та моделей аналізу ринку, які враховують специфіку бізнесу та

унікальні особливості роздрібної торгівлі. Цей підхід дозволяє налаштувати систему точно під потреби конкретного підприємства.

4) Інтеграція з вже існуючими системами управління: побудова інтелектуальної системи на основі інтеграції з вже існуючими системами управління роздрібною торгівлею, такими як системи управління складами, програмні засоби для обліку товарів та продажів, системи управління відносинами з клієнтами (CRM) тощо.

Кожен з цих підходів має свої переваги і обмеження, і вибір конкретного підходу залежить від потреб бізнесу, його ресурсів та стратегічних цілей.

Метод кластерного аналізу є відмінним інструментом для вирішення багатьох завдань у сфері роздрібної торгівлі, оскільки дозволяє систематизувати та групувати великі обсяги даних на основі їхньої схожості.

Тому ми обрали методи машинного навчання і штучного інтелекту, зокрема методи кластеризації, оскільки вони є надзвичайно потужними інструментами для аналізу великих обсягів даних та виявлення складних зв'язків і патернів. Використання методів машинного навчання дозволяє нам автоматизувати процес аналізу ринку роздрібної торгівлі і отримати цінні інсайти для прийняття управлінських рішень.

Методи кластеризації, зокрема, дозволяють нам групувати дані на основі їхньої схожості, що дозволяє виділяти сегменти ринку, ідентифікувати ключові характеристики клієнтів, товарів та конкурентів, а також виявляти потенційні можливості для розвитку бізнесу. Кластерний аналіз дозволяє нам робити персоналізовані рекомендації, прогнозувати попит та реагувати на зміни в ринкових умовах швидко та ефективно.

Кластерний аналіз є статистичним методом аналізу даних, який використовується для групування об'єктів або спостережень у взаємоподібні групи, відомі як кластери. Метою кластерного аналізу є виявлення внутрішньої структури даних та визначення груп об'єктів, які подібні за певними характеристиками, але відрізняються від об'єктів інших груп [16].

Крім того, використання методів машинного навчання дозволяє нам навчити систему адаптуватися до змін в ринкових умовах та навколишньому середовищі, що є важливим для успішного функціонування в динамічному середовищі роздрібною торгівлі. Такий підхід дозволяє нам ефективно використовувати наявні дані та робити з ними зручні та корисні висновки для управління бізнесом.

У ритейлі кластерний аналіз дозволяє виявити закономірності у споживчому попиті та групувати магазини в кластери зі схожими характеристиками споживання. Це дозволяє ефективно адаптувати асортимент та маркетингові стратегії для кожного кластеру магазинів, забезпечуючи відповідність потребам та уподобанням клієнтів. Цей метод також використовується для сегментації аудиторії та клієнтів на групи з подібними характеристиками та вподобаннями. Він дозволяє розділити покупців на різні сегменти на основі таких критеріїв, як споживчі звички, вік, доходи, географічне положення тощо [17].

## РОЗДІЛ 2 ХАРАКТЕРИСТИКА ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ АНАЛІЗУ РИНКУ РОЗДРІБНОЇ ТОРГІВЛІ ТА ПОСТАНОВКА ЗАДАЧІ

### 2.1. Характеристика об'єкта дослідження

В сучасному ринковому середовищі роздрібна торгівля (ритейл) є однією з ключових галузей економіки, де вирішальне значення має ретельне вивчення та аналіз споживчого попиту. Однак, у зв'язку з різноманітністю покупців та їхніх потреб, виникає потреба в ефективних стратегіях асортиментного управління, спрямованих на максимізацію задоволення клієнтів та підвищення рентабельності підприємств роздрібною торгівлі.

Фахівці вважають, що не так важливо, чи має магазин зручне розташування, чи є в ньому правильне розташування товарів, привабливі ціни і ввічливий персонал. Це всього лише додаткові переваги, які не можуть задовольнити основну потребу клієнта - купити бажаний товар [18].

Одна з найважливіших конкурентних переваг мережі – наявний асортимент продукції. Якщо в магазині немає потрібного товару для клієнта, він піде не задоволений до ваших конкурентів.

Оптимальний асортимент – це перелік товарів, який задовольняє потреби покупців, приносить мережі бажаний рівень прибутку, швидко реалізується і не утримує кошти в надлишках. У сьогоднішніх умовах кризи оптимізація асортименту набуває надзвичайної важливості для кожного роздрібного магазину. Вона дозволяє [18]:

- ефективно конкурувати з іншими торговими мережами;
- створювати постійний покупців;
- перетворювати випадкових покупців на постійних клієнтів;
- покращувати репутацію мережі та запускати "сарафанне радіо";

- мінімізувати залишки непроданого товару;
- уникати втрат продажів через нульові залишки або відсутність товарів у асортименті;
- грамотно планувати закупівлю та вдосконалювати ефективну логістику.

Фахівці [18] виокремлюють ряд помилок, яких допускаються різні торговельні мережі у керуванні асортиментом, які протягом тривалого періоду можуть призвести до стратегічних помилок у розвитку бізнесу. Серед них:

- недостатнє вивчення потреб покупців та непередбачення попиту;
- формування асортименту на основі власної інтуїції;
- вплив переконань та пропозицій постачальників;
- неправильне врахування сезонності окремих товарів та особливостей торгової точки;
- відсутність відстеження нових тенденцій та трендів на ринку;
- невідкоригований склад асортименту

Головним рішенням для оптимізації асортименту продукції є планування попиту, який дозволить зрозуміти портрет клієнта і його потреби. Проте, цей процес вимагає враховувати велику кількість факторів:

- зміни товарообігу LFL (Like for like);
- сезонність товарів;
- наявність та результативність промо-акцій;
- властивість гало, тобто, коли один товар допомагає продавати інший;
- властивість канібалізації, тобто коли один товар шкодить продажам іншого;
- фактори, які показують, як часто або наскільки успішно товари або бренди потрапляють на ринок.
- демографічні особливості цільової аудиторії.

Великі обсяги інформації не дозволяють врахувати вручну усі фактори, які впливають на асортимент продукції, або значно знижують точність розрахунків.

Тому для вирішення таких задач рекомендується використовувати аналітичні методи, які ґрунтуються на алгоритмах штучного інтелекту. В цьому контексті кластерний аналіз виявляється важливим інструментом, що дозволяє визначити основні групи споживачів з подібними характеристиками та створити персоналізовані стратегії асортиментного управління для кожної з них.

Задача формування оптимального асортименту виникає не лише у великих гравців на ринку ритейлу.

Сьогодні в багатьох місцях населених пунктів можна зустріти невеликі магазини популярного формату convenience store (C-store), відомі також як магазини «біля будинку» або магазини швидкого доступу. Обмежена площа таких торгових точок потребує надзвичайно раціонального підходу до підбору асортименту та розумного розташування товарів на полицях [19].

Популярність магазинів формату «біля будинку» активно зростає. Територіальна близькість до покупця, економія часу клієнта, наявність основних продуктів та товарів щоденного використання, комфортні умови для здійснення покупок роблять цей формат все більш затребуваним і зручним для споживача.

Проте, навіть якщо торгові точки знаходяться у безпосередній близькості одна від одної, будь то у кварталі, на кілометрі або на протилежних краях одного району, кожна з них може обслуговувати свою унікальну групу покупців.

Щоб надати своїм клієнтам саме те, що їм необхідно, оператори C-store повинні добре розуміти покупців, які відвідують кожен з магазинів.

Під час розуміння реального попиту велику роль відіграє кластеризація, яка допомагає відібрати правильний асортимент у категорії товарів. Можливість впровадження практики кластеризації є важливим показником успіху. Шляхом виявлення попиту на певну категорію товарів і вибору улюблених продуктів можна максимально задовольняти потреби своїх клієнтів. Такі відвідувачі з великою ймовірністю будуть повертатися за покупками знову і стануть вірними клієнтами мережі.

Забезпечити високий рівень розуміння споживчого попиту можна завдяки збору та аналізу даних у режимі реального часу. Крім того, аналітичні дані про

продажі, отримані в ході кластеризації магазинів, можуть бути використані для більш ефективного спілкування з постачальниками, що призведе до розподілу площ з пріоритетом на правильний асортимент.

Таким чином, кластеризація дозволить виявляти товари, які доцільно ввести до асортименту певних магазинів, прогнозувати показники продажу з урахуванням рекомендованого оновлення асортименту, що дозволить збільшити прибуток мережі, обравши найефективнішу асортиментну політику.

Кластеризація дозволить здійснювати аналіз за такими напрямками [18, 19, 20]:

1) Концепція магазину, включаючи його тип та формат. Штучний інтелект збирає дані, щоб визначити, яким чином магазин буде розвиватися та обслуговувати клієнтів. Це визначає, які товарні категорії і продукти будуть присутні на полицях.

2) Характеристики цільової аудиторії та її вподобання. Штучний інтелект аналізує велику кількість даних про клієнтів, щоб краще зрозуміти їх. Ці дані використовуються для надання рекомендацій з оптимізації асортименту, наприклад, вибір продуктів для різних груп покупців.

3) Внутрішні параметри магазину, такі як його розмір, наявність складу, обладнання та умови співробітництва з постачальниками. Ці фактори впливають на формування оптимального асортименту.

4) Конкурентне середовище. Штучний інтелект аналізує дані про конкурентів, включаючи їх асортимент, ціни та стратегії, щоб надати рекомендації для покращення асортименту магазину.

У проведенні кластеризації варто приділяти увагу всім властивостям товарів, оскільки навіть такий аспект, як кількість упаковки, може бути вирішальним при прийнятті рішення про покупку. Наприклад, у магазинах на околицях міста часто краще продаються товари великими упаковками, ніж малими. Для кожної категорії товарів необхідно застосовувати різні підходи до кластеризації торгових точок.

Вхідними даними для здійснення кластеризації є чеки та дані по програмам лояльності, на основі яких формується масив для його подальшої обробки [тези].

Таким чином, кластерний аналіз у ритейлі допомагає вирішувати завдання асортиментного управління, персоналізації маркетингових стратегій та вироблення індивідуалізованих підходів до обслуговування клієнтів, що сприяє збільшенню продажів, підвищенню задоволення споживачів та підвищення конкурентоспроможності торгового підприємства [17].

## **2.2 Структура і характеристика інтелектуальної системи аналізу ринку роздрібною торгівлі та її компонентів.**

**Бізнес-ціль** преектування інтелектуальної системи аналізу ринку роздрібною торгівлі: надання підтримки бізнесу у швидкому прийнятті рішень за допомогою інструмента для автоматичної групування магазинів за ключовими метриками, важливими для роздрібною торгівлі. Однією з типових задач, у яких цей інструмент може бути корисним, є адаптація CVP (Customer Value Proposition) з урахуванням кластерів.

CVP (Customer Value Proposition) – ціннісна пропозиція для споживача – в маркетингу це причина для покупки певного продукту чи послуги у зв'язку з цінністю, яку він пропонує [21]. Тобто це бізнес або маркетингова пропозиція, яка описує, чому клієнт повинен купити продукт або скористатися послугою.

### **Переваги від використання інтелектуальної системи:**

- автоматизація ручного вибору кластерів,
- зменшення трудовитрат і часу, поява можливості відтворення розрахунків,
- підвищення якості кластеризації (використання математичних алгоритмів).

Зазвичай бізнес використовує ручну кластеризацію, виходячи, переважно, з невеликої кількості факторів, таких як: географія розташування магазинів, їх

формат, обсяги продажів. При цьому не враховуються багато параметрів, які можуть суттєво вплинути на більш точний розподіл магазинів за кластерами.

**Передумови створення інтелектуальної системи:** багато часу витрачається на ручну класифікацію магазинів у кластери, і для зменшення трудовитрат бізнесу потрібний зручний і простий у використанні сервіс. Наприклад, бізнес-замовникам потрібно оперативно проводити кластеризацію магазинів, щоб розуміти, чи слід вносити зміни в поточну модель CVP (Customer Value Proposition).

**Мета інтелектуальної системи:** визначення різниці між магазинами на прийнятному рівні управління (тобто на рівні кластера) та можливість розробляти оптимальний асортимент продукції для кожного магазину роздрібною торгівлі, тобто можливість налаштовувати CVP (Customer Value Proposition) для кожного з кластерів.

**Бізнес-обмеження:**

- гнучке налаштування кластеризації - можуть використовуватися як усі параметри, так і вибір найбільш значущих для бізнесу показників (наприклад, виключно аналітичні/комерційні показники);
- існують певні стандартні параметри, які завжди повинні бути присутні під час кластеризації та за якими повинні виводитися метрики в звіті;
- кластеризація проводиться для магазинів різного формату;
- можливість вибору кількості кластерів вручну або отримання автоматичного розподілу на оптимальну кількість кластерів;
- швидкість отримання готового звіту (необхідно проводити розрахунки та вивантажувати звіт в режимі реального часу).

Робота інтелектуальної системи показана на UML-діаграмі послідовності на рисунку 2.1. Послідовність роботи наступна:

- 1) Користувач надає запит інтелектуальній системі на отримання кластеризації обраних торговельних точок.

2) Користувач звертається до таблиць з попередньо розрахованими даними по продажам, а також до таблиць з геоданими та даними про магазини (дата відкриття тощо), витягує дані, агрегує їх та запускає алгоритм кластеризації.

3) Вітрина даних оновлюється щокварталу.

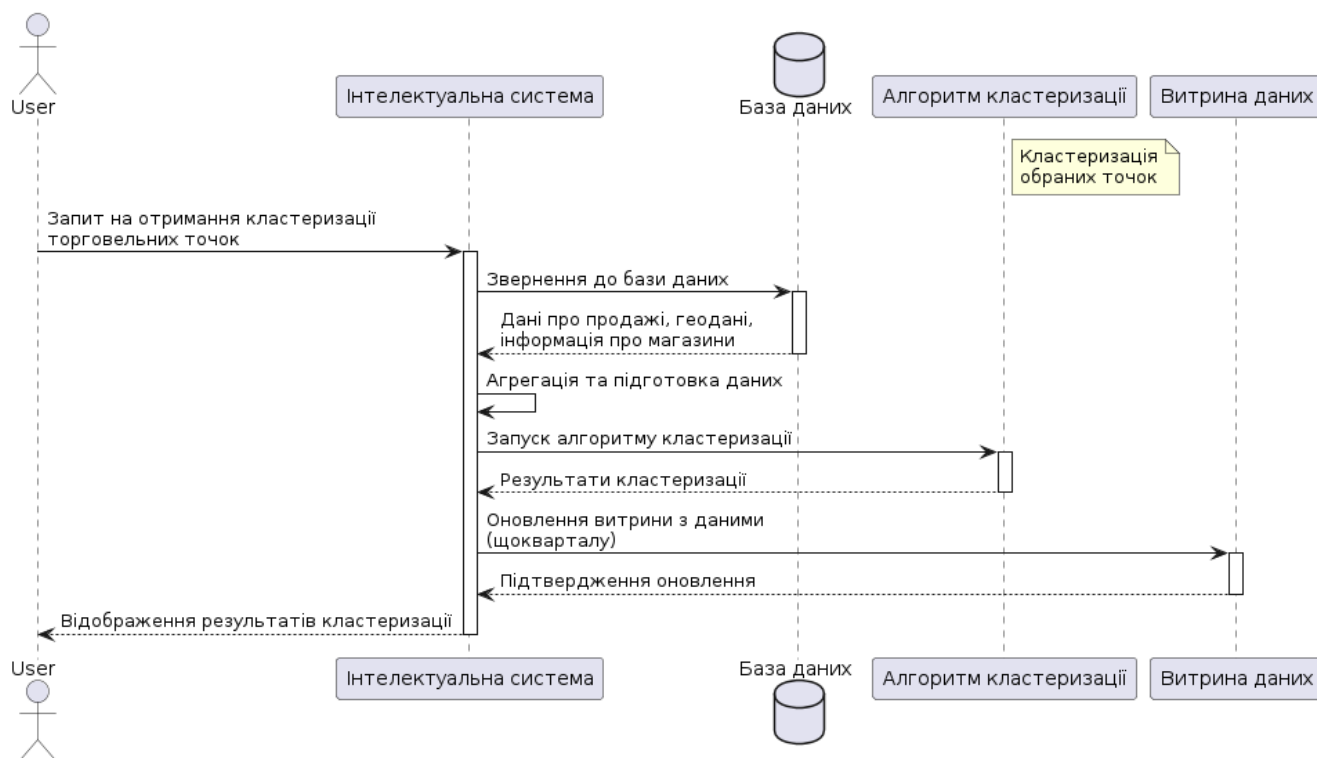


Рисунок 2.1 – Порядок роботи інтелектуальної системи.

*Джерело: розроблено автором самостійно*

Для підготовки даних до кластеризації потрібно виконати наступні дії:

1) Підготовка даних. Передбачається, що дані рядків будуть використовуватися для аналізу, а характеристики представлені у вигляді стовпців.

2) Наступним кроком є заповнення пропущених значень у даних медіанними значеннями.

3) Приведення даних до однакового стандартного вигляду.

4) Розмірність даних зменшується до  $n$  компонент.

5) Кількість компонент обирається динамічно згідно з критерієм Кайзера.

Суть критерію Кайзера полягає в тому, що він використовується для визначення кількості факторів або компонент у методах факторного та аналізу головних

компонент [22, 23 с. 258]. Цей критерій базується на власних значеннях коваріаційної (або кореляційної) матриці даних. Ідея полягає в тому, що фактори з власними значеннями, більшими за одиницю, відображають більш значущу частину дисперсії вихідних даних. Таким чином, вони є вагомими та можуть бути включені до моделі. Критерій Кайзера допомагає зрозуміти, яка кількість факторів є достатньою для відтворення основної частини дисперсії вихідних даних. Зазначаємо, що компонента важлива і її варто залишити, якщо власне значення  $\geq 1$ .

#### б) Застосування алгоритму кластеризації.

Процес підготовки даних для здійснення кластеризації показано на рисунку 2.2.

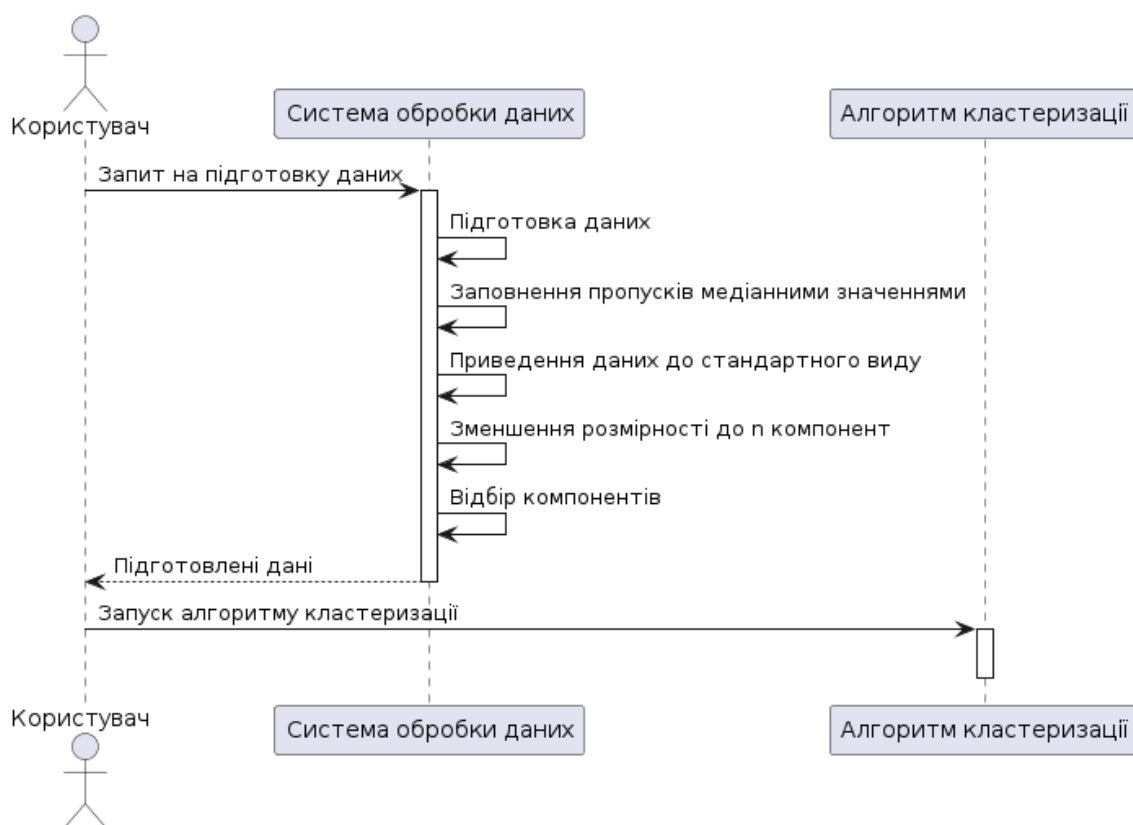


Рисунок 2.2 – Порядок підготовки даних до кластеризації

Джерело: розроблено автором самостійно

Структура системи показана на рисунку 2.3. До основних компонентів системи відносяться база даних та модуль інтелектуальної підсистеми. Дана інтелектуальна система пов'язана з підсистемою маркетингу (Marketing Sysytem).

Користувач ініціює процес, надаючи запити на кластеризацію торговельних точок та обробку даних. Інтелектуальна система приймає запити користувача і використовує їх для запуску алгоритмів кластеризації. Підготовка даних включає отримання даних з бази даних та їх обробку, що передається до алгоритму кластеризації. База даних використовується як джерело інформації для підготовки даних і кластеризації. Результати кластеризації використовуються для формування асортименту продукції торговельних точок та для персоналізованої розсилки.

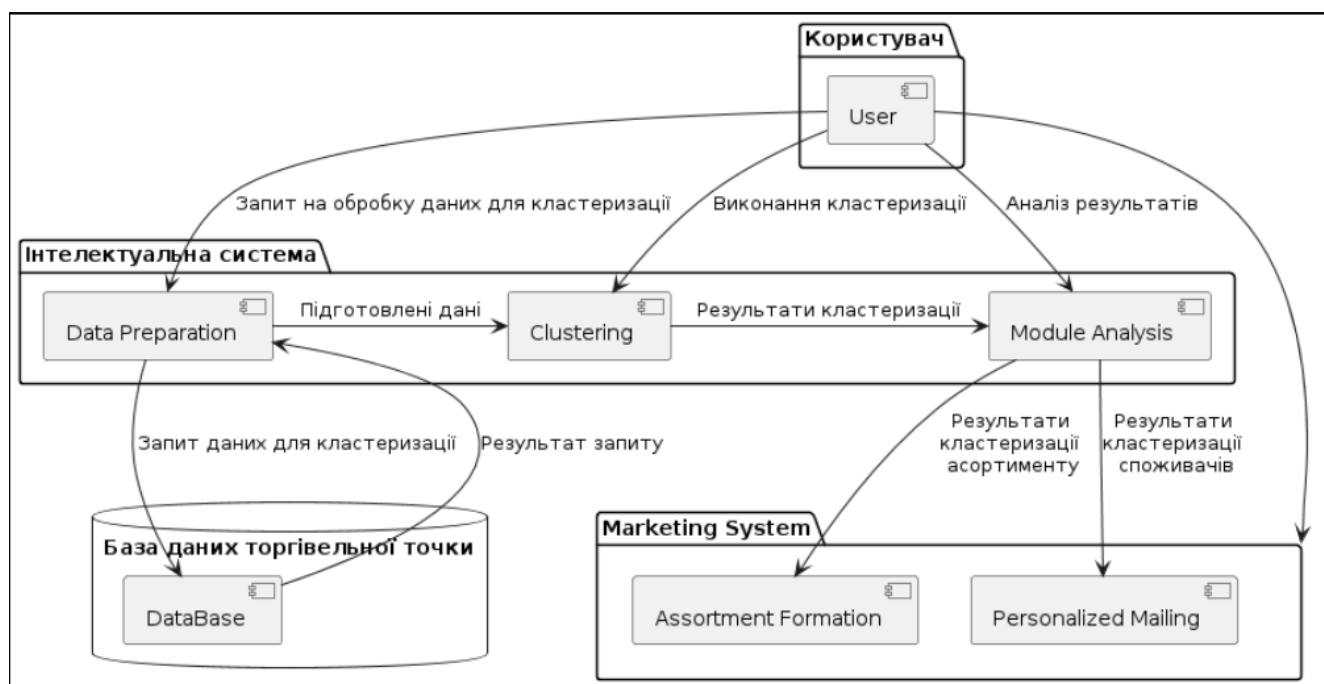


Рисунок 2.3 – Порядок підготовки даних до кластеризації

Джерело: розроблено автором самостійно

## 2.3 Методи, моделі і моделювання процесів і елементів складних систем

### 2.3.1 Методи дослідження й синтезу компонент систем штучного інтелекту

Для кластерного аналізу даних пропонується використовувати алгоритм К-середніх, перевагами якого є простота у використанні та масштабування, що робить його ефективним на великих обсягах даних. Даний алгоритм підтримується багатьма програмними засобами для аналізу даних та машинного навчання він також легко оптимізується під платформи типу Apache Hadoop, які використовуються для ефективного зберігання та обробки великих наборів даних [24].

Розглянемо етапи процесу кластеризації для аналізу даних в ритейлі [25].

1. **Підготовка та очищення даних.** Необхідно видалити інформацію, яка не впливає на поведінку покупців. Наприклад, несуттєвою інформацією є придбання пакетів на касі.

2. **Формування матриці із вхідними даними.** Важливо зазначити, що результати кластеризації сильно залежать від періоду часу, за який вона проводиться. Наприклад, короткий період може дозволити побачити поточні тенденції продажів. Кластеризація за тривалий період дозволить отримати узагальнену картину, виявивши стійкі групи клієнтів з різними стилістичними особливостями. Зокрема, можна виділити групи, що характеризуються споживчими вподобаннями чи поведінкою, наприклад, «Студенти», «Домогосподарки», «Пенсіонери» і т. д. Крім того, роздрібним мережам важливо враховувати особливості програм лояльності. Наприклад, учасники програми можуть відрізнятися за своїми покупками та частотою їх здійснення. Отже, матрицю необхідно формувати, враховуючи основні цілі кластеризації.

3. **Вибір оптимальної кількості кластерів.** Кількість кластерів обирається на основі власного досвіду шляхом експериментів. Недостатньо мале число

кластерів буде неефективним і непоказовим, оскільки при цьому виникає лише один-два «мегакластера», що охоплюють 98% клієнтів, і кілька малих кластерів, які не несуть корисної інформації. З іншого боку, при великій кількості кластерів утворюється занадто багато малих груп. Крім того, ніхто не бажає аналізувати 5000 окремих малих кластерів. В кожному випадку потрібен індивідуальний підхід. Наприклад, для аналізу тривалих періодів та великої кількості кластерів застосовується метод К-середніх.

4. *Здійснення кластеризації.* Для здійснення кластеризації пропонуємо використання алгоритму K-means на мові програмування Python з використанням бібліотеки scikit-learn, призначеної для кластерного аналізу [26].

5. *Аналіз результатів кластеризації.* Після застосування алгоритмів кластеризації всі об'єкти (наприклад, клієнти) розділяються на групи в залежності від їхніх характеристик. Клієнти з схожими характеристиками об'єднуються в один кластер. Якщо аналіз проводиться за певний період (наприклад, рік або місяць), то в кластеризації беруть участь лише ті клієнти, які здійснили покупки протягом вказаного періоду. Під час кластеризації алгоритм може виявити нетипову поведінку клієнта. Виявлення такої поведінки сприяє аналізу окремих характеристик кожного окремого кластера.

6. *Дослідження характеристик кожного кластера.* Кожному кластеру можна присвоїти назву, що відображає його особливості, наприклад, кластери «Домогосподарки», «Одружені», «Студенти», «Любителі тварин». Це допомагає ідентифікувати кластери залежно від їхнього поведінкового профілю.

Оборот кожного кластера дозволяє визначити, які кластери приносять найбільший дохід.

Доля кожного кластера в обороті відображає його вклад у загальний оборот, виражений у відсотках від загального обороту за вибраний період кластеризації.

Кількість клієнтів у кожному кластері вказує на його розмір.

Кількість нових клієнтів у кожному кластері визначається кількістю осіб, які вперше скористалися дисконтною картою протягом обраного періоду кластеризації.

Розподіл між чоловіками та жінками у кожному кластері у відсотковому співвідношенні дозволяє виявити типові покупки для кожної статі.

Загальна кількість чеків у кожному кластері та кількість чеків на одного клієнта дозволяють відстежити частоту покупок кожного клієнта протягом обраного періоду кластеризації.

Середня кількість товарів у чеку та середня вартість чеку дозволяють визначити, в якому кластері продаються найбільш дорогі товари.

Аналізуючи характеристики кожного кластера і використовуючи їх для визначення типу покупця, можна перейти до наступного етапу – персоналізованої розсилки.

**7. Формування асортименту і здійснення персоналізованих розсилок для кожного кластера.** Використовуючи кластеризацію клієнтів, можна розробити чітку систему рекомендацій - який товар, якому клієнту і в який час пропонувати. Знання того, що і якій групі людей пропонувати, дозволяє компаніям уникнути непотрібних розсилок SMS або електронною поштою. Пропонуючи клієнтам лише ті товари, які вони потребують (з урахуванням супутніх), можна досягти значного збільшення відгуку та конверсії у покупку.

### **2.3.2 Моделі та методи оптимізації в системах штучного інтелекту**

Для групування магазинів за певними ознаками пропонуємо використовувати методи кластеризації.

Кластеризація - це метод машинного навчання, що використовується для групування точок даних. Подаючи набір точок даних на вхід, ми можемо використати алгоритм кластеризації для призначення кожній точці даних певної групи. Точки даних у одній групі мають подібні характеристики, тоді як характеристики точок даних у різних групах відрізняються. Кластеризація

відноситься до методів навчання без учителя і є статистичною методикою обробки даних, яка застосовується в різних галузях [29, 30].

Для кластеризації даних застосовуються різні методи, але ціль цих методів однакова – об'єднати схожі за характеристиками об'єкти у групи. Для кожного методу кластеризації можна визначити модель кластеру та алгоритм кластеризації [31,32].

Алгоритм К-середніх є одним з найвідоміших методів кластеризації, де побудовані кластери розташовані якнайдалі один від одного.

Метод k-середніх є широко використовуваним методом кластеризації, який прагне мінімізувати середню квадратичну відстань між точками в одному кластері. Незважаючи на те, що він не гарантує абсолютної точності, його простота і швидкість це компенсують. Доповнюючи метод k-середніх простою рандомізованою технікою висіву, отримуємо алгоритм, який є  $O(\log k)$ -конкурентним з оптимальною кластеризацією. Експерименти показали, що таке доповнення досить сильно збільшує швидкість і точність k-середніх [33]

Основна ідея методу k-середніх полягає у визначенні кластерів таким чином, щоб мінімальна сумарна варіація між кластерами (тобто загальна варіація в межах кластера) була мінімізована. Існує декілька алгоритмів для цього методу, однак стандартний алгоритм визначає загальну варіацію в межах кластера як суму квадратичних евклідових відстаней між елементами і відповідним центроїдом (2.1):

$$W(C_k) = \sum_{x_i \in C_k} (x_i - \mu_k)^2, \quad (2.1)$$

де  $x_i$  – точка даних, що належить до кластера  $C_k$ ,  $\mu_k$  – середнє значення точок, що присвоєні кластеру  $C_k$ .

Кожне спостереження  $x_i$  присвоюється заданому кластеру таким чином, щоб сума квадратів відстані спостереження до їх призначених кластерних центрів  $\mu_k$  мінімізувалася. Алгоритм виглядає наступним чином:

1. Визначити  $k$  — кількість кластерів, які будуть створені.
2. Вибрати випадкові  $k$  об'єктів з набору даних у ролі початкових центрів кластерів.

3. Призначити кожне спостереження найближчому центроїду на основі евклідової відстані між об'єктом і центроїдом.

4. Для кожного з  $k$  кластерів перерахувати центроїд кластера шляхом обчислення нового середнього значення усіх точок даних у кластері.

5. Ітеративно мінімізувати загальну суму в межах площі. Повторити кроки 3 і 4, поки центроїди не зміняться або не буде досягнуто максимальної кількості ітерацій.

Загальна сума, або загальна варіація в межах кластера визначається таким чином (2.2):

$$\sum_{k=1}^k W(C_k) = \sum_{k=1}^k \sum_{x_i \in C_k} (x_i - \mu_k)^2, \quad (2.2)$$

Однією з переваг методу К-середніх є його швидкодія. Він має лінійну складність  $O(n)$ .

Однак, методу К-середніх має ряд недоліків. По-перше, потрібно наперед задавати кількість кластерів для розбиття, що не завжди є очевидним. По-друге, результат роботи алгоритму залежить від того, яким чином ми визначимо початкові центри кластерів. Тому при кількоразовому виконанні алгоритму можуть отримуватися різні результуючі кластери. Інші методи кластеризації можуть бути більш однозначними у цьому відношенні.

### 2.3.3 Методи та моделі управління в системах штучного інтелекту

При використанні кластерного аналізу важливим є визначення оптимальної кількості кластерів у методі k-середніх (k-means clustering). З цією метою пропонуємо використання методу Elbow (метод "ліктя"), який базується на аналізі сум квадратів помилок (WCSS), які виникають при внутрішньокластерній варіації [27, 28].

Основна ідея методу полягає в тому, що при збільшенні кількості кластерів SSE буде зменшуватися, оскільки кожен зразок буде ближче до свого власного

центру кластера. Проте, з певного моменту, додаткове ділення на кластери може призвести до невеликої зміни в WCSS, оскільки нові кластери вже не так сильно зменшують внутрішньокластерну дисперсію.

На графіку методу "ліктя" (рис. 2.4) кількість кластерів (по горизонтальній вісі) проти WCSS (по вертикальній вісі). Точка на графіку, де спостерігається раптова зміна (або "ліктьова точка"), вказує на оптимальну кількість кластерів для аналізу.

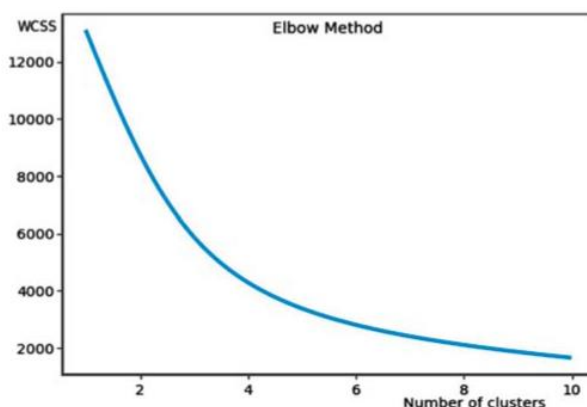


Рисунок 2.4 – Графік методу «ліктя»

*Джерело: [28]*

Варто зауважити, що вибір оптимальної кількості кластерів є неоднозначним і може залежати від конкретної задачі, даних та методу аналізу. Тому рекомендується використовувати додаткові методи для підтвердження результатів.

## РОЗДІЛ 3 РОЗРОБЛЕННЯ ПРОЕКТНИХ РІШЕНЬ

### 3.1 Моделювання та проектування бази знань прийняття інтелектуальних рішень і управління

Змоделюємо процес підготовки даних та здійснення кластеризації.

Для проведення кластеризації необхідно підготувати набір даних, отриманих з бази даних торговельних точок і сформованих у файлі набору даних. Підготовлені дані будемо зберігати у файлі типу .csv. Для отримання даних у такому вигляді потрібно зібрати інформацію з бази даних кожної торговельної точки. На рисунку 3.1 схематично показаний процес отримання інформації та підготовки даних.

Модель потоку даних, показаний на рис.3.1 полягає у наступних діях.

#### 1. Отримання даних від торгових точок (ТТ):

- кожна торгова точка надсилає дані про продані товари за певний період, виконуючи запити SQL;
- отримана інформація містить такі дані, як код торгової точки, дата продажу, кількість транзакцій, кількість одиниць товару, сума продаж.

2. Обробка отриманих даних. Отримані дані обробляються для формування файлу Stores.csv, який містить необхідні атрибути:

- код торгової точки,
- дата продажу,
- кількість транзакцій,
- кількість одиниць товару,
- сума продаж.

3. Розрахунок показників ефективності продажів. На основі даних з файлу Stores.csv розраховуються показники ефективності продажів, такі як:

- ефективність продажів на площі торгового залу,
- коефіцієнт продаж на одиницю товару,
- коефіцієнт продаж на одного клієнта.

4. Отримання набору даних для кластеризації. Отримані дані підготовляються для подальшої кластеризації, утворюючи набір даних *df*.

Таким чином, модель потоку даних включає отримання, обробку та аналіз даних з торгових точок, включаючи розрахунок показників ефективності продажів і підготовку даних для кластеризації.

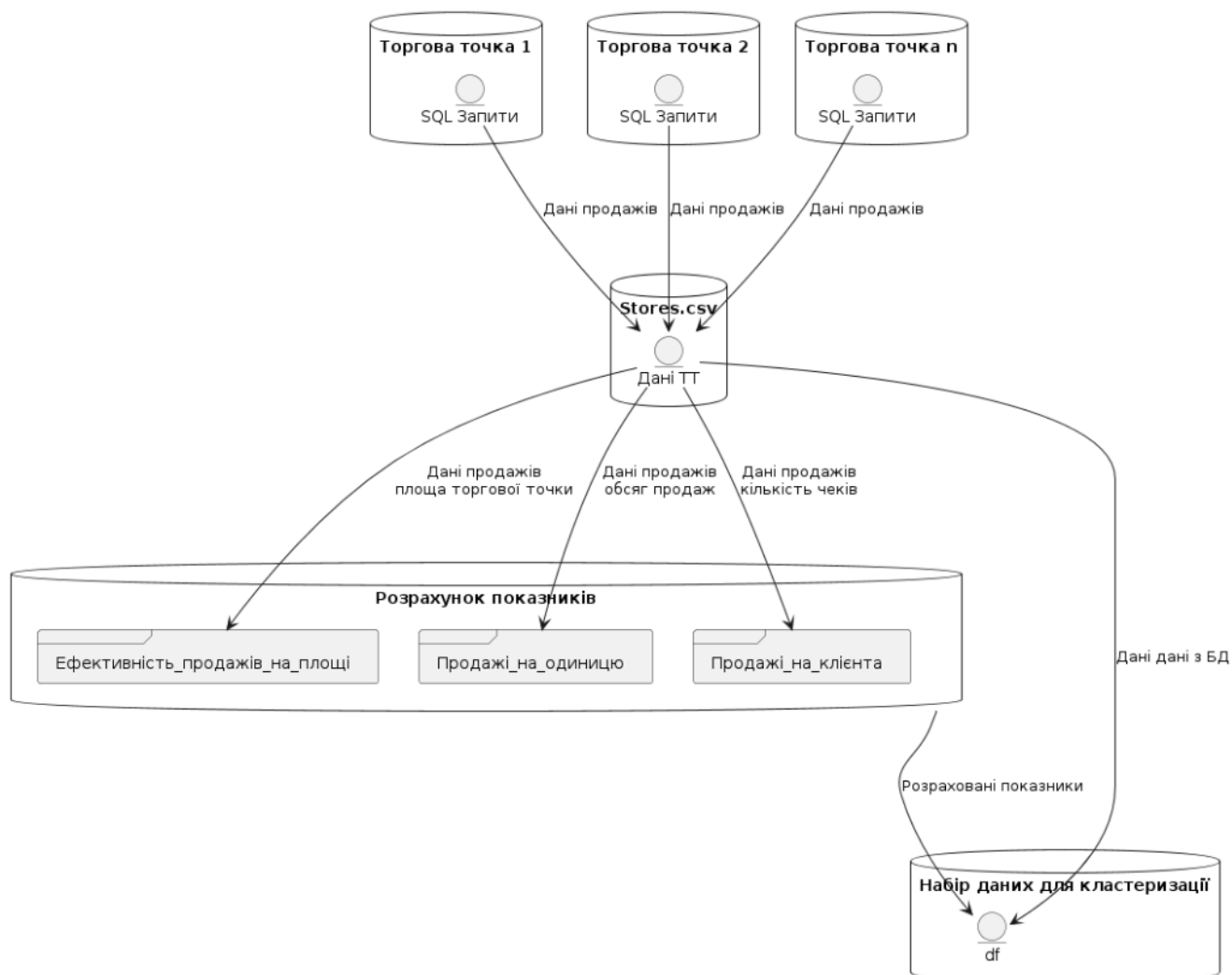


Рисунок 3.1 – Модель потоку даних для формування набору даних для кластеризації

*Джерело: розроблено автором самостійно*

З метою формування *Stores.csv* кожна торгівельна точка надає інформацію з бази даних, яку можна реалізувати шляхом групувального запиту, зміст якого і його результат показаний на рис. 3.2.

```

SELECT Trade_check.id_point,
       Month(Trade_check.date_check) as 'Місяць',
       Year(Trade_check.date_check) as 'Рік',
       Trade_check.id_check as 'Номер чеку',
       count(Tovar_check.vol_tovar) AS 'Кількість артиклів',
       sum(Tovar_zalish.retail_price*Tovar_check.vol_tovar) as Вартість
FROM   Trade_check INNER JOIN
       Tovar_check ON Trade_check.id_check = Tovar_check.id_check INNER JOIN
       Tovar_zalish ON Tovar_check.articul = Tovar_zalish.articul INNER JOIN
       Tovar ON Tovar_zalish.id_tovar = Tovar.id_tovar INNER JOIN
       Categoria ON Tovar.id_categ = Categoria.id_categ
group by Trade_check.id_point, Trade_check.id_check,
         Month(Trade_check.date_check), Year(Trade_check.date_check)

```

	id_point	Місяць	Рік	Номер чеку	Кількість артиклів	Вартість
1	1	5	2024	1	5	430.63
2	2	5	2024	2	6	466.96
3	1	5	2024	3	8	661.90
4	2	5	2024	4	6	496.93
5	2	5	2024	5	4	668.23
6	1	5	2024	6	6	449.86
7	4	5	2024	7	7	468.44
8	3	5	2024	8	7	554.98
9	3	5	2024	9	6	455.90
10	3	5	2024	10	9	649.16

Рисунок 3.2 – Запит і його результат для формування файлу *Stores.csv*

*Джерело: розроблено автором самостійно*

Наступним кроком потрібно згрупувати інформацію по кожній точці продажу. На основі продемонстрованого на рис.3.2 запиту створили представлення, і з даного представлення сформуємо групувальний запит, який надає інформацію про кількість товарних чеків в певному магазині (торгівельній точці), кількість проданих товарів, вартість продаж (рис. 3.3).

```

SELECT id_point, Місяць, Рік, count([Номер чеку]) as 'Кількість покупок' ,
       sum([Кількість артиклів]) as 'Кількість товарних одиниць',
       sum (Вартість) as 'Вартість'
FROM   Prodaji
group by id_point, Місяць, Рік

```

id_point	Місяць	Рік	Кількість покупок	Кількість товарних одиниць	Вартість
1	5	2024	3	19	1542.39
2	5	2024	3	16	1632.12
3	5	2024	3	22	1660.04
4	5	2024	1	7	468.44

Рисунок 3.3 – Підготовлені дані для завантаження у файл *Stores.csv*

*Джерело: розроблено автором самостійно*

На рисунку 3.4 показаний вміст файлу *Stores.csv*, який отримуємо на основі даних з баз даних торгівельних точок. Структура файлу показана у таблиці 3.1

	Store ID	Площа	Кількість_Артикулів	Кількість_покупців	Обсяг_продаж
0	1	1659	1961	530	66490
1	2	1461	1752	210	39820
2	3	1340	1609	720	54010
3	4	1451	1748	620	53730
4	5	1770	2111	450	46620
5	6	1442	1733	760	45260
6	7	1542	1858	1030	72240
7	8	1261	1507	1020	37720
8	9	1090	1321	680	46310
9	10	1030	1235	1130	44150

Рисунок 3.4 – Набір даних файлу *Stores.csv*

*Джерело: розроблено автором самостійно*

Таблиця 3.1 – Структура файлу *Stores.csv*

Номер ЗП	Назва колонки	Позначення у файлі	Одиниці вимірювання	Тип
1.	Код магазину (торгової точки)	Store ID		integer
2.	Площа торгової точки	Площа	кв.м	decimal(8,2)
3.	Кількість найменувань товарів у магазині	Кількість_Артикулів	штуки	integer
4.	Кількість покупців за місяць	Кількість_покупців	чоловік	integer
5.	Обсяг продажів за місяць	Обсяг_продаж	тис.грн.	integer

*Джерело: розроблено автором самостійно*

Для отримання додаткових показників аналізу пропонуємо розрахувати показники, ефективність продажів на площі торгового залу (3.1), коефіцієнт продаж на одиницю товару (3.2), коефіцієнт продаж на клієнта (3.3).

Показники розрахуємо наступним чином.

Ефективність продажів на площі торгового залу:

$$ET_{mi} = \frac{S_{mi}}{PL_i}, \quad (3.1)$$

де  $ET_{mi}$  – ефективність продажів у торговій точці  $i$  за місяць- $m$  на площі торгового залу;

$S_{mi}$  – сума продажів у торговій точці  $i$  за місяць- $m$ ;

$PL_i$  – торгівельна площа у торговій точці  $i$ .

Коефіцієнт продаж на одиницю товару:

$$KP_{mi} = \frac{S_{mi}}{T_{mi}}, \quad (3.2)$$

де  $KP_{mi}$  – коефіцієнт продажів товару у торговій точці  $i$  за місяць- $m$ ;

$T_{mi}$  – кількість одиниць продукції (артикулів), проданих у торговій точці  $i$  за місяць- $m$ ;

Коефіцієнт продаж на клієнта:

$$KK_{mi} = \frac{S_{mi}}{C_{mi}},$$

де  $KK_{mi}$  – коефіцієнт продажів на клієнта у торговій точці  $i$  за місяць- $m$ ;

$C_{mi}$  – кількість покупок, які здійснив клієнт (кількість товарних чеків) у торговій точці  $i$  за місяць- $m$ .

Після розрахунку показників (3.1)-(3.3) отримали набір даних, показаних на рис. 3.5. На основі отриманого набору даних будемо здійснювати кластеризацію. В якості тестових будемо використовувати набір даних з ресурсу Kaggle [33].

Масив даних файлу *Stores.csv* складається з 896 рядків, тому пропонуємо застосувати 5 кластерів.

Store ID	Площа	Кількість_Артикулів	Кількість_покупців	Обсяг_продаж	Ефективність_продажів_на_площі	Продажі_на_одиницю	Продажі_на_клієнта
1	1659	1961	530	66490	40.078360	33.906170	125.452830
2	1461	1752	210	39820	27.255305	22.728311	189.619048
3	1340	1609	720	54010	40.305970	33.567433	75.013889
4	1451	1748	620	53730	37.029635	30.737986	86.661290
5	1770	2111	450	46620	26.338983	22.084320	103.600000
6	1442	1733	760	45260	31.386963	26.116561	59.552632
7	1542	1858	1030	72240	46.848249	38.880517	70.135922
8	1261	1507	1020	37720	29.912768	25.029861	36.980392
9	1090	1321	680	46310	42.486239	35.056775	68.102941
10	1030	1235	1130	44150	42.864078	35.748988	39.070796

Рисунок 3.5 – Підготовлений для кластеризації набір даних

*Джерело: розроблено автором самостійно*

Для підтвердження гіпотези про оптимальну кількість кластерів, скористаємося методом Elbow method, описаним в п.2.3.3.

За допомогою програмного коду на Python отримали графік, який свідчить, що кількість кластерів – 5 – є оптимальною величиною (рис.3.6).

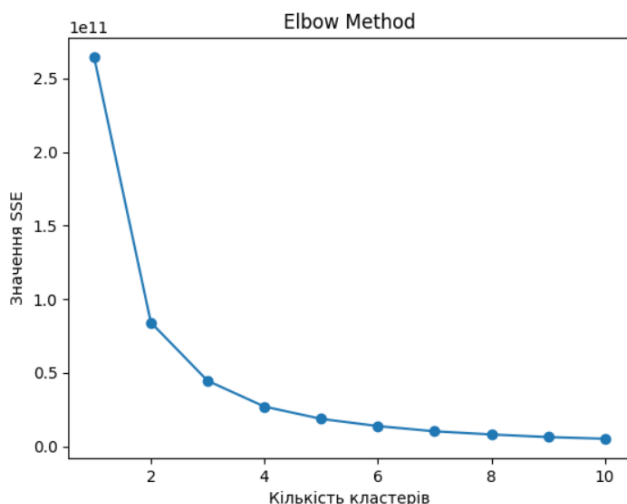


Рисунок 3.6 – Графік методу «ліктя» для визначення оптимальної кількості кластерів

*Джерело: розроблено автором самостійно на основі [28]*

Наступним кроком є здійснення кластеризації, яке виконане на мові Python з використанням бібліотеки scikit-learn.

### 3.2 Розроблення користувацького інтерфейсу. Елементи та структура.

Структуру користувацького інтерфейсу опишемо у вигляді прототипу екранної форми інтелектуальної системи аналізу ринку роздрібною торгівлі.

Користувачами системи є фахівці відділу маркетингу торгівельної мережі роздрібною торгівлі.

Основними складовими інтерфейсу є:

- модуль підготовки даних, при виборі якого користувач системи отримує необхідну інформацію, здійснює перегляд показників, вибір необхідних для кластеризації, їх налаштування;
- модуль кластеризації – виконання алгоритмів кластеризації;
- модуль аналізу результатів і формування звітів. Користувач має змогу переглянути результати у графічному вигляді та сформувати відповідні звіти.

Опишемо результати реалізації моделі.

Після розрахунку моделі кожен елемент даних отримав мітку, яка класифікує торговельну точку у окремій кластер. На рис.3.7 показаний фрагмент оброблених даних, остання колонка **Cluster\_Label** показує приналежність до кластеру.

```
[ ] df.head(10)ТЬ текстовую ячейку
```

	Store ID	Площа	Кількість_Артикулів	Кількість_покупців	Обсяг_продаж	Ефективність_продажів_на_площі	Продажі_на_одиницю	Продажі_на_клієнта	Cluster_Label
0	1	1659	1961	530	66490	40.078360	33.906170	125.452830	2
1	2	1461	1752	210	39820	27.255305	22.728311	189.619048	3
2	3	1340	1609	720	54010	40.305970	33.567433	75.013889	1
3	4	1451	1748	620	53730	37.029635	30.737986	86.661290	1
4	5	1770	2111	450	46690	26.388883	22.081420	103.600000	1

Рисунок 3.7 – Результати кластеризації даних.

*Джерело: розроблено автором самостійно*

За допомогою програмного коду, показаного на рис. 3.8 отримали інформацію про належність кожної торговельної точки до кластеру:

```

# Перебір кожного магазину
for store_id in df['Store ID'].unique():
    # Пошук мітки кластера для поточного магазину
    cluster_label = cluster_labels[df['Store ID'] == store_id]
    # Збереження мітки кластера для поточного магазину в словнику
    cluster_labels_stores[store_id] = cluster_label

# Виведення мітки кластерів для всіх магазинів
for store_id, cluster_label in cluster_labels_stores.items():
    print("Магазин Store ID =", store_id, "належить до кластера з міткою:", cluster_label)

```

```

Магазин Store ID = 1 належить до кластера з міткою: [4]
Магазин Store ID = 2 належить до кластера з міткою: [3]
Магазин Store ID = 3 належить до кластера з міткою: [1]
Магазин Store ID = 4 належить до кластера з міткою: [1]
Магазин Store ID = 5 належить до кластера з міткою: [6]
Магазин Store ID = 6 належить до кластера з міткою: [6]
Магазин Store ID = 7 належить до кластера з міткою: [7]
Магазин Store ID = 8 належить до кластера з міткою: [3]
Магазин Store ID = 9 належить до кластера з міткою: [6]

```

Рисунок 3.8 – Виведення мітки кластера кожному магазину

*Джерело: розроблено автором самостійно*

На рис.3.9 показано графік розсіювання по двом ознакам, зокрема, показано розсіювання між ознаками «Об'єм продаж» і «Площа», інша пара – «Ефективність продаж на площі» і «Продажі на клієнта».

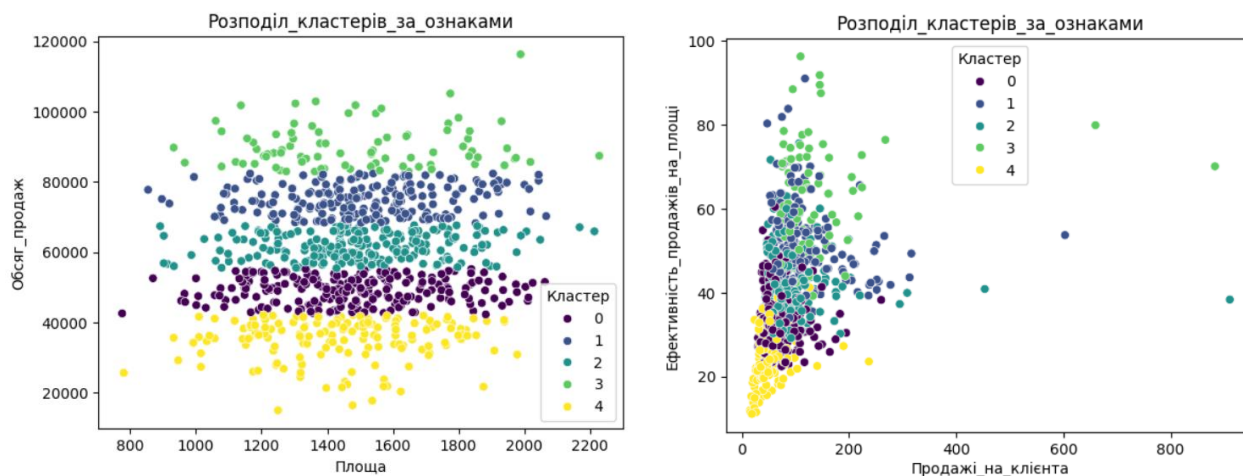


Рисунок 3.9 – Графік розсіювання для двох ознак

*Джерело: розроблено автором самостійно*

На наступному рисунку 3.10 показані матриці гістограм для ознак «Кількість покупців» і «Обсяг продаж». Гістограми розфарбовані за кластерами, що дозволяє візуально порівняти розподіл кожної ознаки між кластерами.

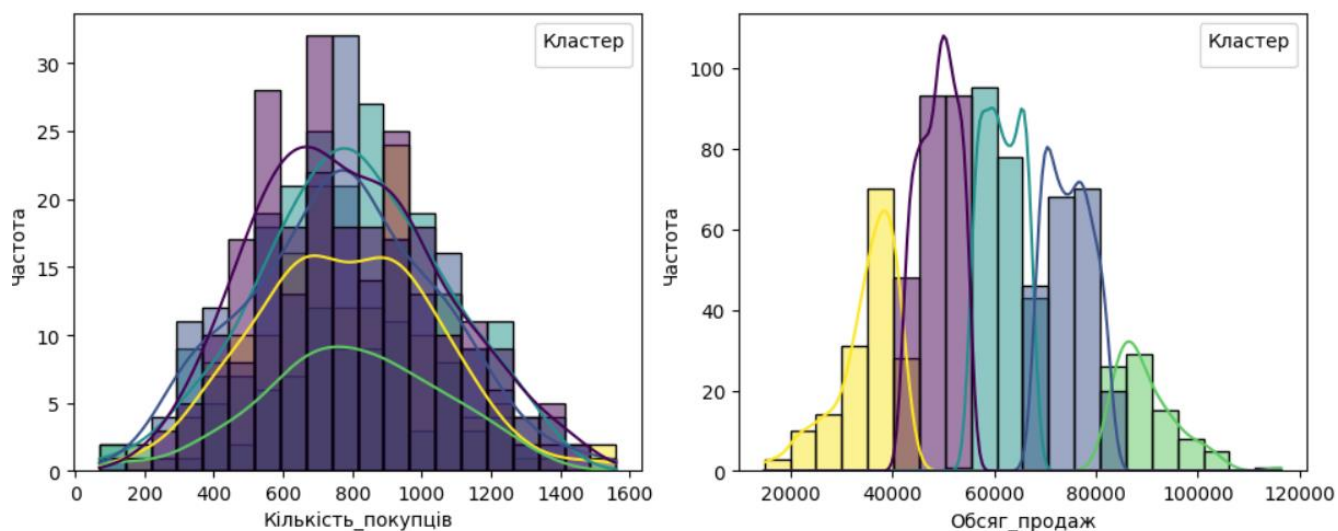


Рисунок 3.10 – Графік гістограм для ознак «Кількість покупців» і «Обсяг продаж»

*Джерело: розроблено автором самостійно*

На рисунку 3.11 показано перші 50 із набору даних магазинів із розподілом їх на кластери. Для кращої візуалізації біля кожної точки показаний id відповідного магазину.

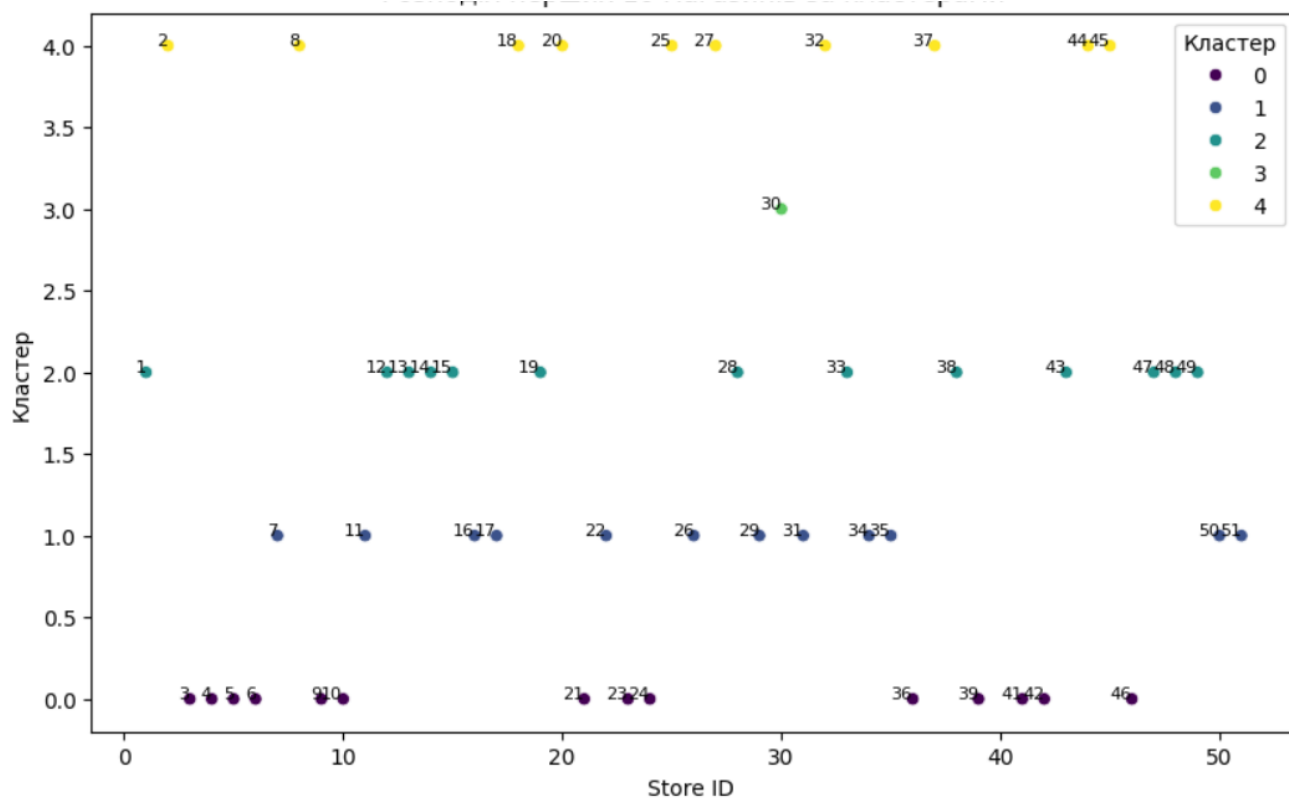


Рисунок 3.11 – Графік розподілу магазинів за кластерами

*Джерело: розроблено автором самостійно*

### 3.3 Проектування забезпечувальних підсистем СШ. Реалізація системи.

#### 3.3.1 Інформаційне забезпечення.

До складу інформаційного забезпечення інтелектуальної системи аналізу ринку роздрібною торгівлі входять нормативно-методичні матеріали, система класифікації і кодування, баз даних, в якій міститься інформацію про продаж товарів у точках продажу та набір даних, необхідний для інтелектуального аналізу.

До складу нормативного забезпечення входять закони і постанови Уряду, пов'язані з регулюванням роздрібною торгівлі. Визначимо деякі з них:

- Закон України “Про захист прав споживачів” [34]
- Наказ Міністерства економіки № 19.04.2007 № 104 «Про затвердження Правил роздрібною торгівлі непродовольчими товарами» [35];
- Порядок провадження торговельної діяльності та правила торговельного обслуговування на ринку споживчих товарів, затверджено постановою Кабінету Міністрів України від 15.06.2006 № 833 [36].

На рис. 3.12 показана схема інформаційного забезпечення.

Для роботи бази даних потрібні наступні класифікатори:

1. Довідник категорій товарів (Categoria) – застосовується порядковий метод кодування, довжина коду – 5 символів.

2. Довідник торговельних марок і брендів (brand). Для кодування застосовується послідовний метод кодування.

Структура коду наступна:

XX – категорія, до якої відноситься бренд,

XXX – порядковий номер бренду відповідно категорії.

Довжина коду – 5 символів.

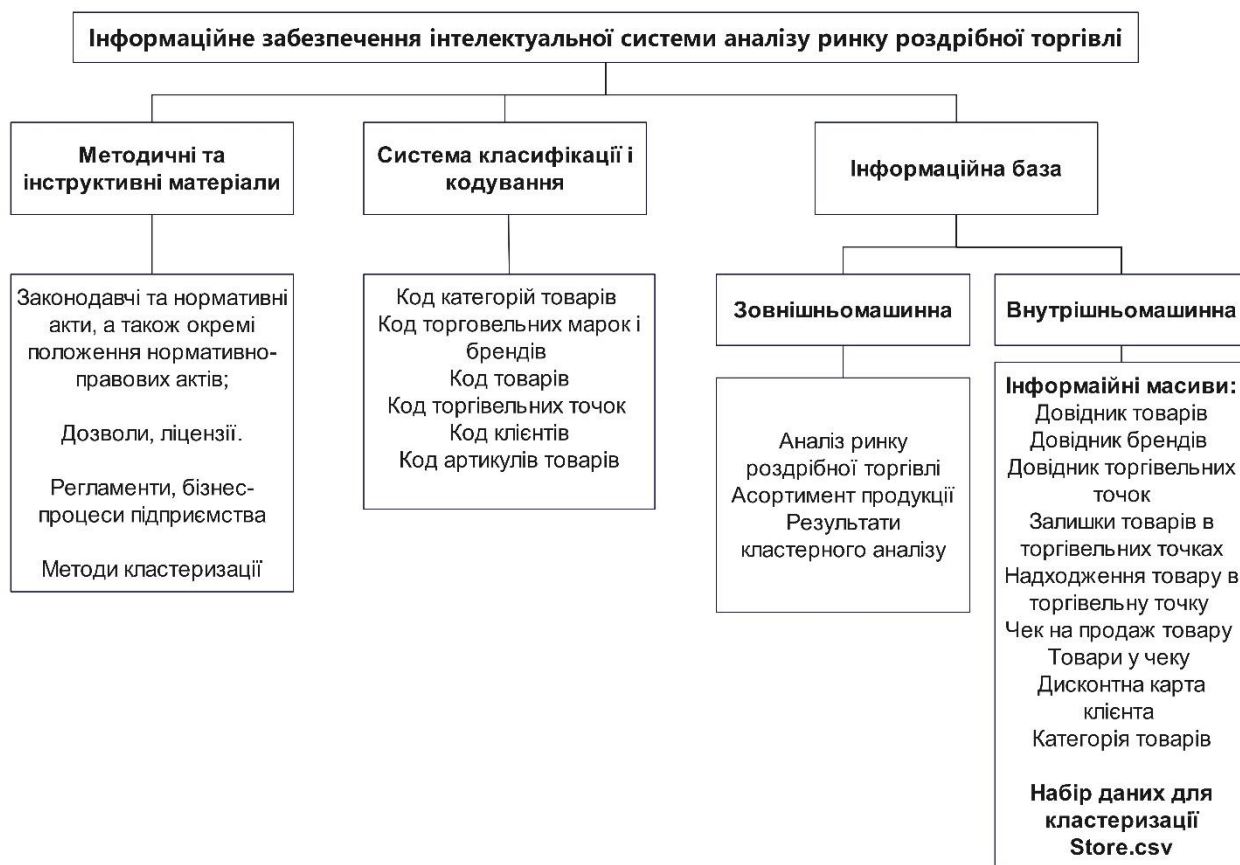


Рисунок 3.12 – Інформаційне забезпечення

*Джерело: розроблено автором самостійно*

3. Довідник товарів (Tovar) – послідовний метод кодування. Структура коду наступна:

XX – група товарів;

XX – підгрупа товарів;

XXX – порядковий номер товару у відповідній підгрупі.

Довжина коду – 6 символів.

4. Довідник торговельних точок (Trade\_point) – пропонується послідовний метод кодування, структура коду:

XX – код регіону, де розташована торговельна точка;

XXX – порядковий номер точки продажу.

Довжина коду складає 5 символів.

5. Довідник клієнтів. Для цього довідника ключовим є номер дисконтної карти, який використовують клієнти.

Структура коду карти наступна:

XXXXXX – код торговельної точки, де первинно оформлювалась карта;

XX – рік отримання карти;

XXX – порядковий номер клієнта

X – контрольне число.

Довжина номеру карти – 11 символів.

6. Класифікатор артикулів товарів. Для обліку товарів в кожній торговельній точці пропонується вести облік в розрізі артикулів товарів.

Артикул – це код, який продавці використовують для ідентифікації та обліку своїх товарів та їх варіантів. Він може складатися з будь-яких комбінацій букв, цифр та символів. Покупці можуть побачити артикули в інвойсах або опис товару на сайті магазину. Якогось універсального формату, відповідно до якого потрібно створювати артикули, не існує (хоча деякі формати і можна віднести до «найкращих практик»). Можна використовувати літери алфавіту або, навпаки, обійтися лише цифрами [37].

Пропонуємо наступну структуру артикулу товару:

XX – бренд;

XXX – категорія товару;

X – цільова аудиторія споживачів ;

XX – розмір товару (за потребою, інакше 00);

XX – вид товару;

XX – підвид товару.

Довжина коду артикулу – 12 символів

Схема бази даних, в якій зберігається інформація про продажі товарів у різних торговельних точках, показана на рисунку 3.13.

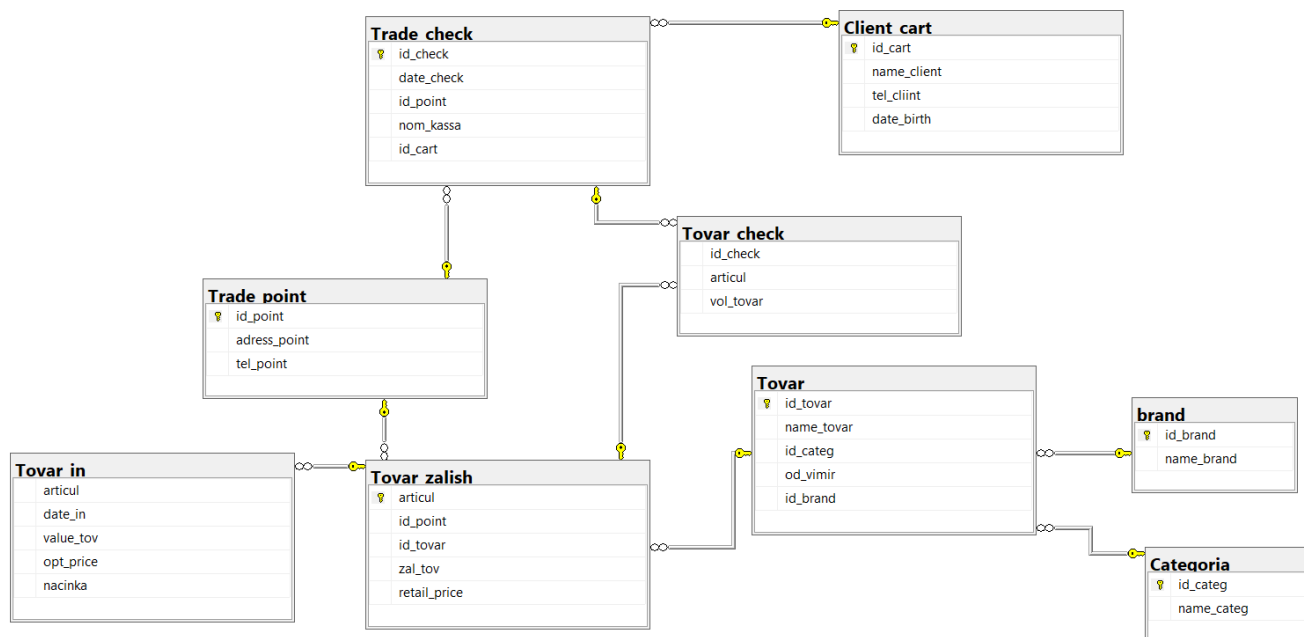


Рисунок 3.13 – Схема бази даних

Джерело: розроблено автором самостійно

Структура інформаційних масивів показана в таблицях 3.2 – 3.10.

Таблиця 3.2 – Опис масиву “Довідник товарів”

Найменування масиву — Довідник товарів

Ідентифікатор масиву — Товар

Найменування носія інформації — Жорсткий диск

Максимальний об’єм масиву — 100000 записів.

Довжина запису — 86 символів.

Метод організації — послідовний.

Ключі упорядкування — id\_ tovar

Ідентифікатор індексного масиву — ind\_ tovar

Найменування	Ідентифікатор у программі	Умовне позначення у формулах	Формат	Бізнес-правила				Логічні чи семантичні зв'язки
				Первинний /вторинний ключ	Умова на значення	Обов'язкове поле	Індексне поле	
Код товару	id_tovar	t	9(6)	PK	-	так	ІН Д	-
Назва товару	name_tovar	-	X (50)	-	-	так	-	-
Код категорії	id_categ	k	9(5)	FK	-	так	ІН Д	Category
Одиниця виміру	od_vimir	-	X (20)	-	-	так	-	-
Код бренду	id_brand	b	9(5)	FK	-	так	ІН Д	brand

Джерело: розроблено автором самостійно

Таблиця 3.3 – Опис масиву “Довідник брендів”

Найменування масиву — Довідник брендів

Ідентифікатор масиву — brand

Найменування носія інформації — Жорсткий диск

Максимальний об'єм масиву — 1000 записів.

Довжина запису — 55 символів.

Метод організації — послідовний.

Ключі упорядкування — id\_brand

Ідентифікатор індексного масиву — ind\_brand

Найменування	Ідентифікатор у программі	Умовне позначення у формулах	Формат	Бізнес-правила				Логічні чи семантичні зв'язки
				Первинний /вторинний ключ	Умова на значення	Обов'язкове поле	Індексне поле	
Код бренду	id_brand	b	9(5)	PK	-	так	ІН Д	Tovar
Назва бренду	name_brand	-	X(50)	-	-	так	-	

Джерело: розроблено автором самостійно

Таблиця 3.4 – Опис масиву “Довідник торгівельних точок”

Найменування масиву — Довідник торгівельних точок

Ідентифікатор масиву — Trade\_point

Найменування носія інформації — Жорсткий диск

Максимальний об’єм масиву — 10000 записів.

Довжина запису — 128 символів.

Метод організації — послідовний.

Ключі упорядкування — id\_point

Ідентифікатор індексного масиву — ind\_point

Найменування	Ідентифікатор у програмі	Умовне позначення у формулах	Формат	Бізнес-правила				Логічні чи семантичні зв’язки
				Первинний /вторинний ключ	Умова на значення	Обов’язкове поле	Індексне поле	
Код торгової точки	id_point	p	9(8)	PK	-	так	ІН Д	Tovar_zalish, Trade_check
Адреса торгової точки	adress_point	-	X(100)	-	-	так	-	-
Телефон торгової точки	tel_point	-	X(20)	-	-	ні	-	-

*Джерело: розроблено автором самостійно*

Таблиця 3.5 – Опис масиву “Залишки товарів в торгівельних точках”

Найменування масиву — Залишки товарів в торгівельних точках

Ідентифікатор масиву — Tovar\_zalish

Найменування носія інформації — Жорсткий диск

Максимальний об’єм масиву — 1000000 записів.

Довжина запису — 28 символів.

Метод організації — послідовний.

Ключі упорядкування — articul

Ідентифікатор індексного масиву — ind\_zalish

Найменування	Ідентифікатор у программі	Умовне позначення у формулах	Формат	Бізнес-правила				Логічні чи семантичні зв'язки
				Первинний /вторинний ключ	Умова на значення	Обов'язкове поле	Індексне поле	
Артикул товару	articul	a	9(12)	PK	-	так	ІН Д	Tovar_in
Код торгової точки	id_point	p	9(8)	-	-	так	ІД Д	Trade_point
Код товару	id_tovar	t	9(8)	-	-	так	ІД Д	Tovar
Залишок товару	zal_tov	Z	9(10,2)	-	-	так	-	-
Роздрібна ціна товару	retail_price	RP	9(10,2)	-	-	так	-	-

*Джерело: розроблено автором самостійно*

Таблиця 3.6 – Опис масиву “Надходження товару в торгівельну точку”

Найменування масиву — Надходження товару в торгівельну точку

Ідентифікатор масиву — Tovar\_in

Найменування носія інформації — Жорсткий диск

Максимальний об'єм масиву — 1000000 записів.

Довжина запису — 46 символів.

Метод організації — послідовний.

Ключі упорядкування — articul

Ідентифікатор індексного масиву — ind\_Tovar\_in

Найменування	Ідентифікатор у программі	Умовне позначення у формулах	Формат	Бізнес-правила				Логічні чи семантичні зв'язки
				Первинний /вторинний ключ	Умова на значення	Обов'язкове поле	Індексне поле	
Артикул товару	articul	a	9(12)	FK	-	так	ІД Д	Tovar_zalish
Дата введення	date_in	d	D(8)	-	-	так	-	-
Кількість товару, що надійшло	value_tov	KN	9(10,2)	-	-	так	-	-
Оптова ціна	opt_price	OP	9(10,2)	-	-	так	-	-
Націнка	nasinka	N	9(6,2)	-	-	-	-	-

*Джерело: розроблено автором самостійно*

Таблиця 3.7 – Опис масиву “Чек на продаж товару”

Найменування масиву — Чек на продаж товару

Ідентифікатор масиву — Trade\_check

Найменування носія інформації — Жорсткий диск

Максимальний об'єм масиву — 1000000 записів.

Довжина запису — 44 символи.

Метод організації — послідовний.

Ключі упорядкування — id\_check

Ідентифікатор індексного масиву — ind\_Trade\_check

Найменування	Ідентифікатор у программі	Умовне позначення у формулах	Формат	Бізнес-правила				Логічні чи семантичні зв'язки
				Первинний /вторинний ключ	Умова на значення	Обов'язкове поле	Індексне поле	
Номер чеку	id_check	c	9(15)	PK	-	так	ІН Д	Tovar_check
Дата чеку	date_check	d	D(8)	-	-	так	-	-
Код торгової точки	id_point	p	9(8)	FK	-	так	ІД Д	Trade_point
Номер каси	nom_kassa	s	9(2)	-	-	ні	-	-
Код картки	id_cart	i	9(11)	FK	-	ні	ІД Д	Client_cart

*Джерело: розроблено автором самостійно*

Таблиця 3.8 – Опис масиву “Товари у чеку”

Найменування масиву — Товари у чеку

Ідентифікатор масиву — Tovar\_check

Найменування носія інформації — Жорсткий диск

Максимальний об'єм масиву — 1000000 записів.

Довжина запису — 44 символи.

Метод організації — послідовний.

Ключі упорядкування — id\_check

Ідентифікатор індексного масиву — ind\_ Tovar\_check

Найменування	Ідентифікатор у программі	Умовне позначення у формулах	Формат	Бізнес-правила				Логічні чи семантичні зв'язки
				Первинний /вторинний ключ	Умова на значення	Обов'язкове поле	Індексне поле	
Номер чеку	id_check	c	int	FK	-	так	ІД Д	Trade_ check
Артикул товару	articul	a	int	-	-	так	ІД Д	Tovar_ zalish
Кількість товару	vol_tovar	KS	int	-	-	так	-	-

Джерело: розроблено автором самостійно

Таблиця 3.9 – Опис масиву “Дисконтна карта клієнта”

Найменування масиву — Дисконтна карта клієнта

Ідентифікатор масиву — Client\_cart

Найменування носія інформації — Жорсткий диск

Максимальний об'єм масиву — 1000000 записів.

Довжина запису — 64 символи.

Метод організації — послідовний.

Ключі упорядкування — id\_check

Ідентифікатор індексного масиву — ind\_Client\_cart

Найменування	Ідентифікатор у программі	Умовне позначення у формулах	Формат	Бізнес-правила				Логічні чи семантичні зв'язки
				Первинний /вторинний ключ	Умова на значення	Обов'язкове поле	Індексне поле	
Номер картки клієнта	id_cart	-	9(11)	PK	-	так	ІН Д	Trade_ check
ПІБ клієнта	name_client	-	A(30)	-	-	так	-	-
Телефон клієнта	tel_cliint	-	X(15)	-	-	так	-	-
Дата народження	date_birth	-	D(8)	-	-	ні	-	-

Джерело: розроблено автором самостійно

Таблиця 3.10 – Опис масиву “Категорія товарів”

Найменування масиву — Категорія товарів

Ідентифікатор масиву — Categoria

Найменування носія інформації — Жорсткий диск

Максимальний об’єм масиву — 9999 записів.

Довжина запису — 55 символів.

Метод організації — послідовний.

Ключі упорядкування — id\_categ

Ідентифікатор індексного масиву — ind\_categ

Найменування	Ідентифікатор у програмі	Умовне позначення у формулах	Формат	Бізнес-правила				Логічні чи семантичні зв’язки
				Первинний /вторинний ключ	Умова на значення	Обов’язкове поле	Індексне поле	
Код категорії	id_categ	k	9(5)	ПК	-	так	ІН Д	Tovar
Назва категорії	name_categ	-	X(50)	-	-	так	ІН Д	-

*Джерело: розроблено автором самостійно*

### 3.3.2 Програмне забезпечення

Для розроблення клієнтської частини інтелектуальної системи аналізу ринку роздрібною торгівлі пропонується використовувати мову програмування Java, а для алгоритмів машинного навчання, зокрема, для кластеризації – мову програмування Python.

Python – вважається одним з найпопулярніших мов програмування у сфері машинного навчання та аналізу даних. Він широкий спектр бібліотек для машинного навчання та аналізу даних, таких як Scikit-learn, TensorFlow, PyTorch тощо. Ці бібліотеки надають різноманітні алгоритми кластеризації, такі як k-means,

DBSCAN, hierarchical clustering тощо, які можна легко використовувати для аналізу даних та отримання результатів. Python дозволяє легко і швидко розробляти власні алгоритми кластеризації або розширювати функціональність існуючих бібліотек за допомогою власних модулів та плагінів.

Python має чистий і простий синтаксис, що робить його досить доступним для новачків у сфері аналізу даних. Більшість бібліотек машинного навчання мають документацію та приклади коду, які полегшують їх використання.

Python також має багато бібліотек для візуалізації даних, таких як Matplotlib, Seaborn, Plotly, що дозволяє відобразити результати кластеризації у вигляді графіків та діаграм, що сприяє кращому розумінню отриманих результатів.

Також має велику та активну спільноту розробників, що робить його досить надійним і стабільним для використання. Завдяки цьому, можна легко знайти відповіді на питання або допомогу вирішити проблеми, які можуть виникнути під час роботи з алгоритмами кластеризації.

Для реалізації алгоритму кластеризації була застосована бібліотека Scikit-learn [38]. Це безкоштовна бібліотека машинного навчання, реалізована на Python, яка пропонує широкий вибір алгоритмів як для навчання з учителем, так і без нього. Однією з основних переваг цієї бібліотеки є те, що вона базується на кількох поширених математичних бібліотеках і легко інтегрує їх одну з одною. Ще однією перевагою є широка спільнота користувачів та докладна документація. Scikit-learn спеціалізується на алгоритмах машинного навчання для розв'язання завдань як навчання з учителем - класифікації та регресії, так і без учителя - кластеризації, зменшення розмірності та виявлення аномалій [39]. Для реалізації кластерного аналізу був застосований метод KMeans [40].

Візуалізацію результатів було виконано за допомогою методів:

scatterplot [41] з бібліотеки Seaborn [42] – для побудови графіка розсіювання для двох ознак з розфарбуванням за кластерами;

subplots [43] з бібліотеки Matplotlib [44] – для створення множини графіків розподілу даних для кожної ознаки.

### 3.3.3 Технічне забезпечення

Для роботи інтелектуальної системи аналізу ринку роздрібною торгівлі потрібен сучасний комп'ютер з можливостями візуалізації. Для отримання друкованих звітів необхідний принтер.

В [45] описані рекомендації при виборі комп'ютера для роботи з графікою. На основі даних рекомендацій пропонуємо наступну конфігурацію комп'ютера, достатнього для роботи інтелектуальної системи аналізу ринку роздрібною торгівлі.

**Процесор:** для функціонування інтелектуальної системи буде достатньо Intel Core i3-10100. Це 4 ядерний процесор з тактовою частотою 3.6 ГГц. Якщо є потреба у великій кількості візуалізації, можна скористатися більш потужним ПК з комплектацією Intel Core i3-10320 і Intel Core i5-10500. Це процесори з гарною тактовою частотою і високими показниками продуктивності.

**Відеокарта:** NVIDIA GeForce 1660. Для більш потужної візуалізації можна скористатися відеокартами NVIDIA GeForce RTX 2060 або NVIDIA GeForce RTX 2070.

**Оперативна пам'ять:** для роботи інтелектуальної системи достатньо 4GB оперативної пам'яті. Проте, якщо необхідно обчислювати великі обсяги даних, потрібно буде 8GB.

**SSD:** рекомендується використовувати 3 накопичувача: 2xSSD і HDD. HDD потрібен для зберігання об'ємних файлів, а SSD для завантаження системи і візуалізація результатів програм. Samsung PRO і Kingston A1000 – найкращі SSD M.2 на сьогоднішній день.

**Монітор** – обов'язкова складова будь-якого професійного ПК. Монітори на основі IPS матриці найкраще підходять для даної інтелектуальної системи.

Монітор рекомендується з FullHD дозволом.

### 3.3.4 Організаційно- економічне забезпечення

Критерієм успіху при впровадженні інтелектуальної системи аналізу ринку роздрібною торгівлі є позитивний зворотний зв'язок від користувачів стосовно розробленого інструмента кластеризації та зменшення трудовитрат на ручний розрахунок.

За дослідженнями фахівців, наразі бізнес витрачає на кластеризацію від 3 до 5 годин робочого часу, який включає в себе також очікування від інших команд та залучення їх ресурсів для вивантаження даних з різних сервісів та аналітичних звітів. Передбачається, що використання інтелектуальної системи зменшить час на формування даних та кластеризацію до 15 хвилин.

Оскільки рекомендується аналіз асортименту продукції проводити один раз на тиждень, це значно скоротить час на даний аналіз, що в свою чергу, значно скоротить витрати.

## ВИСНОВКИ

В результаті написання кваліфікаційної магістерської роботи було досліджено предметну область щодо створення інтелектуальної системи аналізу роздрібною торгівлі та запропоновано проєктні рішення щодо її створення.

Зокрема, було досліджено основні поняття роздрібною торгівлі та ритейлу, визначено, що дані поняття мають деякі відмінності. Визначені основні особливості ритейлу в порівнянні із звичайним магазином роздрібною торгівлі: розташування магазину на вигідній території, яка ретельно підбирається; легкість знаходження необхідного товару; використання відповідного торгового або холодильного обладнання, яке виконує кілька функцій; застосування маркетингових прийомів, таких як акції та знижки, що допомагають привернути увагу покупців; можливість самообслуговування клієнтами, що сприяє швидкому здійсненню покупок та не вимагає великої кількості працівників для обслуговування клієнтів, автоматизація зберігання товарів, закупок та обліку за допомогою комп'ютерних програм, що значно спрощує логістичні та бізнес-процеси.

Проаналізовано сучасні інформаційні технології, які застосовуються в роздрібній торгівлі. Зокрема, виявлено, що інформаційні системи і технології використовуються в наступних напрямках: **оптова торгівля, дистрибуція, виробничий збут, роздрібна торгівля, Cash&Carry, онлайн-торгівлі**. Також активно розвиваються цифрові технології та системи штучного інтелекту.

Визначено, що у зв'язку з різноманітністю покупців та їхніх потреб, виникає потреба в ефективних стратегіях асортиментного управління, спрямованих на максимізацію задоволення клієнтів та підвищення рентабельності підприємств роздрібною торгівлі, якої можна досягти шляхом використання інтелектуальних систем на основі сучасних методів штучного інтелекту і кластерного аналізу зокрема.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Ритейл [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: <https://www.carrotquest.io/glossary/retail/>.
2. Ритейл – що це таке, як працює та хто такий Ритейлер. [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: <https://termin.in.ua/ryteyl-ryteyler/>.
3. Що таке ритейл простими словами [Електронний ресурс]. – 2023. – Режим доступу до ресурсу: <https://art-trade.com.ua/chto-takoe-ritejl-prostymi-slovami/>.
4. Що таке рітейл? Види, бізнес-процеси. [Електронний ресурс]. – 2019. – Режим доступу до ресурсу: [https://olteba.com/article/chto\\_takoe\\_riteil\\_vidi\\_biznesprocessi](https://olteba.com/article/chto_takoe_riteil_vidi_biznesprocessi).
5. Ритейл Дізнайтеся, в чому суть ритейлу та ознайомтеся з його видами [Електронний ресурс]. – 2023. – Режим доступу до ресурсу: <https://sendpulse.ua/ru/support/glossary/retail>.
6. Концепція використання товарооблікових систем в галузі торгівлі. Практичне застосування УТП, УВП, CRM, TMS, WMS і ін. рішень для управління. [Електронний ресурс]. – 2019. – Режим доступу до ресурсу: <https://tqm.com.ua/ua/likbez/ua-articles/zastosuvannja-1c-utp-crm-na-praktyci-it-dlja-torgivli>.
7. Микита Попов. ШІ-технології в ритейлі. Як утримати покупців та зменшити витрати [Електронний ресурс] / Микита Попов. – 2022. – Режим доступу до ресурсу: <https://speka.media/ai/ai-tehnologiyi-v-riteili-yak-utrimati-pokupciv-ta-zmensiti-vitrati-p1w069>.
8. «Делойт» в новинах Як українські ритейлери використовують штучний інтелект [Електронний ресурс]. – 2023. – Режим доступу до ресурсу: <https://www2.deloitte.com/ua/uk/pages/press-room/deloitte-press/2023/11-16.html>
9. Вікторія Саравас. Як українські ритейлери використовують штучний інтелект [Електронний ресурс] / Вікторія Саравас. – 2023. – Режим доступу до

ресурсу: <https://biz.liga.net/ua/all/it/article/yak-ukrainski-ryteilery-vykorystovuiut-shtuchnyi-intelekt>.

10. More than half of retailers will deploy AI, machine learning tech in next year [Електронний ресурс]. – 2023. – Режим доступу до ресурсу: <https://capcom-ncr.com/insights/2023/more-than-half-of-retailers-will-deploy-ai-machine-learning-tech-in-next-year>.

11. Коцовський В. М. Методи та системи штучного інтелекту / В. М. Коцовський. – Ужгород: Конспект лекцій, 2016. – 76 с.

12. Вахнюк, С.В. Технологія створення програмних та інтелектуальних систем [Текст] : навчальний посібник / С. В. Вахнюк. – Суми : ДВНЗ “УАБС НБУ”, 2011. – 254 с.

13. Winston W. Microsoft Excel Data Analysis and Business Modeling. 7th ed. Microsoft Press, 2022. 1168 p.

14. The latest in machine learning. Papers With Code [Електронний ресурс]. Електрон. дан. Режим доступу: World Wide Web. URL: <https://paperswithcode.com/>.

15. Гужва В.М. Системи бізнес-аналітики [Електронний ресурс]: навч. посіб. / Київ: КНЕУ, 2016. 265 с.

16. Ситник В. Ф., Краснюк М.Т. Інтелектуальний аналіз даних (дейтамайнінг): Навч. посібник. К: КНЕУ, 2007. 376 с.

17. Як кластеризація допомагає підібрати правильний асортимент у категорії? [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: <https://www.c4r.eu/ukr/blog/biznes-konsalt/klasterizatsiya/>

18. Як грамотно оптимізувати асортименти в умовах кризи? [Електронний ресурс]. – 2022. – Режим доступу до ресурсу: <https://datawiz.io/ru/blog/optimization-of-the-assortment>.

19. Як кластеризація магазинів допомагає малоформатному бренду мислити масштабно [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: <https://www.c4r.eu/rus/blog/biznes-konsalt/klasterizatsiya-magazinov/>.

20. Безпарточний, Максим Григорович. "Використання кластерного аналізу при оцінці ефективності діяльності торговельних підприємств." Торгівля, комерція, підприємництво 17 (2014): 24-27.

21. Cambridge Dictionary. Customer value proposition [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: <https://dictionary.cambridge.org/uk/dictionary/english/customer-value-proposition>.

22. Тоцька О. Л. Визначення прихованих факторів впливу на виробництво продуктів харчування у Волинському регіоні / О. Л. Тоцька // Регіональна економіка. - 2008. - № 3. - С. 182-188. - Режим доступу: [http://nbuv.gov.ua/UJRN/regek\\_2008\\_3\\_25](http://nbuv.gov.ua/UJRN/regek_2008_3_25)

23. Бізнес-аналітика багатовимірних процесів : навчальний посібник [Електронний ресурс] / Т. С. Клебанова, Л. С. Гур'янова, Л. О. Чаговець та ін. – Харків : ХНЕУ ім. С. Кузнеця, 2018. – 272 с.

24. What is Hadoop? [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: [https://aws.amazon.com/what-is/hadoop/?nc1=h\\_ls](https://aws.amazon.com/what-is/hadoop/?nc1=h_ls).

25. Кластеризація: розкажи мені, що ти купуєш, і я скажу, хто ти [Електронний ресурс]. – 2015. – Режим доступу до ресурсу: <https://habr.com/ru/companies/datawiz/articles/248863/>.

26. Clustering [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: <https://scikit-learn.org/stable/modules/clustering.html#clustering>.

27. Cui M. Introduction to the K-Means Clustering Algorithm Based on the Elbow Method // Accounting, Auditing and Finance. 2020. Vol. 1, № 1.

28. Олійник, С. В., & Бойко, Б. М. (2023). Порівняння методів кластеризації для створення цільових груп клієнтів у наборі даних. Науковий вісник НЛТУ України, 33(5), 77-83. <https://doi.org/10.36930/40330510>

29. Ситник В. Ф., Краснюк М.Т. Інтелектуальний аналіз даних (дейтамайнінг): Навч. посібник. К: КНЕУ, 2007. 376 с.

30. Стандартні методи кластеризації даних [Електронний ресурс]. – 2020. – Режим доступу до ресурсу:

[https://csc.knu.ua/media/study/asp/mod\\_prob1\\_inf\\_tech\\_sys\\_analysis\\_ivohin/lecture/lec2.pdf](https://csc.knu.ua/media/study/asp/mod_prob1_inf_tech_sys_analysis_ivohin/lecture/lec2.pdf).

31. Jain A., Murty M., Flynn P. (1999) Data Clustering: A Review. // ACM Computing Surveys

32. Основні поняття кластеризації та постановка задачі [Електронний ресурс]. – 2020. – Режим доступу до ресурсу: [https://csc.knu.ua/media/study/asp/mod\\_prob1\\_inf\\_tech\\_sys\\_analysis\\_ivohin/lecture/lec11.pdf](https://csc.knu.ua/media/study/asp/mod_prob1_inf_tech_sys_analysis_ivohin/lecture/lec11.pdf).

33. Store Efficiency Segmentations and Strategies [Електронний ресурс]. – 2022. – Режим доступу до ресурсу: <https://www.kaggle.com/code/sasakitetsuya/store-efficiency-segmentations-and-strategies/notebook>.

34. Закон України “Про захист прав споживачів” – Режим доступу до ресурсу: <https://zakon.rada.gov.ua/laws/show/1023-12#Text>

35. Про затвердження Правил роздрібної торгівлі непродовольчими товарами. Наказ Міністерства економіки №. 104 від 19.04.2007 – 2007. – Режим доступу до ресурсу: <https://zakon.rada.gov.ua/laws/show/z1257-07#Text>.

36. Порядок провадження торговельної діяльності та правила торговельного обслуговування на ринку споживчих товарів, затверджено постановою Кабінету Міністрів України від 15.06.2006 № 833 – Режим доступу до ресурсу: <https://zakon.rada.gov.ua/laws/show/833-2006-%D0%BF#Text>

37. Артикули товарів та як ними користуватися [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: <https://support.ecwid.com/hc/ru/articles/360011125640-%D0%90%D1%80%D1%82%D0%B8%D0%BA%D1%83%D0%BB%D1%8B-%D1%82%D0%BE%D0%B2%D0%B0%D1%80%D0%BE%D0%B2-%D0%B8-%D0%BA%D0%B0%D0%BA-%D0%B8%D0%BC%D0%B8-%D0%BF%D0%BE%D0%BB%D1%8C%D0%B7%D0%BE%D0%B2%D0%B0%D1%82%D1%8C%D1%81%D1%8F>.

38. scikit-learn Machine Learning in Python [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: <https://scikit-learn.org/stable/>.

39. Scikit-Learn [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: [http://cloud-5.bitp.kiev.ua/?page\\_id=640](http://cloud-5.bitp.kiev.ua/?page_id=640).
40. sklearn.cluster.KMeans [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: <https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html#sklearn.cluster.KMeans>.
41. seaborn.scatterplot [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: <https://seaborn.pydata.org/generated/seaborn.scatterplot.html#seaborn.scatterplot>.
42. seaborn: statistical data visualization [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: <https://seaborn.pydata.org/>.
43. matplotlib.figure [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: [https://matplotlib.org/stable/api/figure\\_api.html#matplotlib.figure.Figure.subplots](https://matplotlib.org/stable/api/figure_api.html#matplotlib.figure.Figure.subplots).
44. Matplotlib: Visualization with Python [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: <https://matplotlib.org/>.
45. Комп'ютер для графіки 2021 [Електронний ресурс]. – 2021. – Режим доступу до ресурсу: <https://artline.ua/uk/news/kompyuter-dlya-grafiki-2021>.

## ДОДАТКИ

### Додаток А. SQL-код для створення бази даних

```
use Retail

create table Categoria
(id_categ int primary key,
name_categ char(50) not null)

create table Tovar
(id_tovar int primary key,
name_tovar char(50) not null,
id_categ int not null,
constraint tov_cat foreign key (id_categ) references Categoria)

alter table Tovar
drop constraint tov_cat

alter table Tovar
add constraint tov_cat foreign key (id_categ)
references Categoria (id_categ)
on update cascade
on delete cascade

alter table Tovar
add od_vimir char(20) not null

create table brand
(id_brand int primary key,
name_brand char (50) not null)

alter table Tovar
add id_brand int not null
foreign key references brand (id_brand)

create table Trade_point
(id_point int primary key,
adress_point char(100) not null,
tel_point char(20) )

create table Tovar_zalish
(articul int not null primary key,
id_point int not null,
id_tovar int not null,
zal_tov numeric (10,2) not null,
retail_price numeric (10,2) not null
)

alter table Tovar_zalish
add constraint fk_point foreign key (id_point)
references Trade_point (id_point)
on update cascade
on delete cascade

alter table Tovar_zalish
add constraint fk_tovar foreign key (id_tovar)
references Tovar (id_tovar)
on update cascade
on delete cascade
```

```

create table Tovar_in
(articul int not null,
date_in date not null,
value_tov numeric (10,2) not null,
opt_price numeric (10,2)not null,
nacinka numeric (6,2),
constraint fk_tovar_in foreign key (articul)
references Tovar_zalish (articul)
on update cascade
on delete cascade
)

create table Trade_check
(id_check int primary key,
date_check date not null,
id_point int not null,
nom_kassa smallint null,
constraint fk_check foreign key (id_point)
references Trade_point (id_point)
on update cascade
on delete cascade)

create table Tovar_check
(id_check int not null,
articul int not null,
vol_tovar int not null,
constraint fk_articul foreign key (articul)
references Tovar_zalish (articul) on update cascade on delete cascade)

alter table Tovar_check
add constraint fk_tov_check foreign key (id_check)
references Trade_check (id_check) on update cascade on delete cascade

create table Client_cart
(id_cart int not null primary key,
name_client char(30) not null,
tel_cliint char (15) null,
date_birth date null)

alter table Trade_check
add id_cart int null
foreign key references Client_cart (id_cart) null

```

## Додаток Б. Код програми кластеризації

```
# -*- coding: utf-8 -*-
"""Ритейл3.ipynb

Automatically generated by Colab.

Original file is located at
    https://colab.research.google.com/drive/1PEn9sdSilFAzUoFfbjNl4_8YtpiMeDI9
"""

import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)

import os
for dirname, _, filenames in os.walk('/content/drive/MyDrive/Colab Notebooks/Дані і
датасети/Retail'):
    for filename in filenames:
        print(os.path.join(dirname, filename))

# Commented out IPython magic to ensure Python compatibility.
# %matplotlib inline
import matplotlib.pyplot as plt
import seaborn as sns

import warnings
warnings.filterwarnings('ignore')

df=pd.read_csv('/content/drive/MyDrive/Colab Notebooks/Дані і
датасети/Retail/Stores3.csv')

df.head(10)

df.info()

sns.pairplot(df)

plt.figure(figsize = (10,10))
sns.heatmap(df.corr(),annot=True, cbar=True, cmap='Greens', fmt='.1f')

df['Ефективність_продажів_на_площі']=df['Обсяг_продаж']/df['Площа']
df['Продажі_на_одиночку']=df['Обсяг_продаж']/df['Кількість_Артикулів']
df['Продажі_на_клієнта']=df['Обсяг_продаж']/df['Кількість_покупців']

df.info()

df.head(20)

num_list=['Площа', 'Кількість_Артикулів', 'Кількість_покупців',
```

```

        'Обсяг_продаж', 'Ефективність_продажів_на_площі', 'Продажі_на_одиницю',
        'Продажі_на_клієнта']
fig = plt.figure(figsize=(15,15))

for i in range(len(num_list)):
    plt.subplot(3,3,i+1)
    plt.title(num_list[i])
    sns.boxplot(data=df,y=df[num_list[i]])

plt.tight_layout()

df.sort_values(by="Продажі_на_клієнта",ascending=False)

df=df.drop(df.index[39])

df = df.reset_index(drop=True)

!pip install scikit-learn

from sklearn.cluster import KMeans

kmeans = KMeans(n_clusters=5)

kmeans.fit(df)

centers = kmeans.cluster_centers_

labels = kmeans.labels_

import seaborn as sns

from sklearn.cluster import KMeans
import matplotlib.pyplot as plt

# Згенеруйте значення SSE для різної кількості кластерів
sse = []
for k in range(1, 11):
    kmeans = KMeans(n_clusters=k, random_state=42)
    kmeans.fit(df) # df - ваші дані
    sse.append(kmeans.inertia_)

# Побудуйте графік "ліктя" (elbow plot)
plt.plot(range(1, 11), sse, marker='o')
plt.xlabel('Кількість кластерів')
plt.ylabel('Значення SSE')
plt.title('Elbow Method')
plt.show()

# Побудова графіка розсіювання для двох ознак з розфарбуванням за кластерами
sns.scatterplot(data=df, x='Площа', y='Обсяг_продаж', hue=labels, palette='viridis')

```

```

plt.xlabel('Площа')
plt.ylabel('Обсяг_продаж')
plt.title('Розподіл_кластерів_за_ознаками')
plt.legend(title='Кластер')
plt.show()

# Побудова графіка розсіювання для двох ознак з розфарбуванням за кластерами
sns.scatterplot(data=df, x='Продажі_на_клієнта', y='Ефективність_продажів_на_площі',
hue=labels, palette='viridis')
plt.xlabel('Продажі_на_клієнта')
plt.ylabel('Ефективність_продажів_на_площі')
plt.title('Розподіл_кластерів_за_ознаками')
plt.legend(title='Кластер')
plt.show()

fig, axes = plt.subplots(nrows=2, ncols=2, figsize=(10, 8))
for i, ax in enumerate(axes.flatten()):
    sns.histplot(data=df, x=df.columns[i+1], hue=labels, palette='viridis', ax=ax,
kde=True)
    ax.set_xlabel(df.columns[i+1])
    ax.set_ylabel('Частота')
    ax.legend(title='Кластер')
plt.tight_layout()
plt.show()

# Отримання міток кластерів для кожного зразка
cluster_labels = cluster_labels = kmeans.labels_

# Виведення перших 10 міток кластерів
print(cluster_labels[:10])

# Знайдіть мітку кластера для магазину з Store ID = 1
cluster_label_store_1 = cluster_labels[df['Store ID '] == 20]

# Візуалізуйте мітку кластера
print("Магазин Store ID = 1 належить до кластера з міткою:", cluster_label_store_1)

# Створіть порожній словник для збереження міток кластерів для кожного магазину
cluster_labels_stores = {}

# Перебір кожного магазину
for store_id in df['Store ID '].unique():
    # Пошук мітки кластера для поточного магазину
    cluster_label = cluster_labels[df['Store ID '] == store_id]
    # Збереження мітки кластера для поточного магазину в словнику
    cluster_labels_stores[store_id] = cluster_label

# Виведення мітки кластерів для всіх магазинів
for store_id, cluster_label in cluster_labels_stores.items():

```

```

    print("Магазин Store ID =", store_id, "належить до кластера з міткою:",
cluster_label)

import matplotlib.pyplot as plt

# Виберіть дані для конкретного магазину (наприклад, Store ID = 1)
store_data = df[df['Store ID '] == 56]

# Побудуйте графік розподілу для вибраних даних
plt.figure(figsize=(8, 6))
plt.hist(store_data['Обсяг_продаж'], bins=20, color='skyblue', edgecolor='black')
plt.title('Розподіл обсягу продаж для магазину Store ID = 1')
plt.xlabel('Обсяг продаж')
plt.ylabel('Частота')
plt.grid(True)
plt.show()

import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Створення DataFrame з масивом міток кластерів
df['Cluster_Label'] = labels

# Побудова парної діаграми для всіх ознак з розфарбуванням за кластерами
sns.pairplot(df, hue='Cluster_Label', palette='viridis', diag_kind='kde')
plt.suptitle('Парна діаграма ознак з розподілом кластерів', y=1.02)
plt.show()

import seaborn as sns
import matplotlib.pyplot as plt

# Відбір перших 10 магазинів та їх кластерів
top_10_stores = df.head(50)
top_10_stores.reset_index(drop=True, inplace=True)

# Створення графіка розсіювання для перших 10 магазинів та їх кластерів
plt.figure(figsize=(10, 6))
sns.scatterplot(data=top_10_stores, x='Store ID ', y='Cluster_Label',
hue='Cluster_Label', palette='viridis')
plt.title('Розподіл перших 50 магазинів за кластерами')
plt.xlabel('Store ID')
plt.ylabel('Кластер')
plt.legend(title='Кластер')
plt.show()

# Створення графіка розсіювання для перших 10 магазинів та їх кластерів
plt.figure(figsize=(10, 6))
sns.scatterplot(data=top_10_stores, x='Store ID ', y='Cluster_Label',
hue='Cluster_Label', palette='viridis')

```

```
plt.title('Розподіл перших 10 магазинів за кластерами')
plt.xlabel('Store ID')
plt.ylabel('Кластер')
plt.legend(title='Кластер')

# Додавання текстових міток для кожної точки
for index, row in top_10_stores.iterrows():
    plt.text(row['Store ID'], row['Cluster_Label'], str(row['Store ID']),
             ha='right', fontsize=8, color='black')

plt.show()

# Створення графіка розсіювання для перших 50 магазинів та їх кластерів
plt.figure(figsize=(10, 6))
scatter = sns.scatterplot(data=top_10_stores, x='Store ID', y='Cluster_Label',
                           hue='Cluster_Label', palette='viridis')
plt.title('Розподіл перших 10 магазинів за кластерами')
plt.xlabel('Store ID')
plt.ylabel('Кластер')
plt.legend(title='Кластер')

# Додавання текстових міток для кожної точки
for index, row in top_10_stores.iterrows():
    plt.text(int(row['Store ID']), row['Cluster_Label'], str(int(row['Store ID'])),
             ha='right', fontsize=8, color='black')

plt.show()
```

Имя пользователя: Інформаційних систем в економіці Шкуратовська Те...	ID проверки: 1016260329
Дата проверки: 18.05.2024 02:23:14 EEST	Тип проверки: Doc vs Internet + Library
Дата отчета: 18.05.2024 02:28:31 EEST	ID пользователя: 100005745

Название файла: Сіденко В. ІШІ-602

Количество страниц: 71 Количество слов: 12781 Количество символов: 105381 Размер файла: 1.83 MB ID файла: 1016048762

**Обнаружены модификации текста (могут влиять на процент совпадений)**

## 15% Совпадения

Наибольшее совпадение: 2.87% с источником из Библиотеки (ID файла: 1014943693)

9.9% Источники из Интернета 289 ..... Страница 73

9.55% Источники из Библиотеки 574 ..... Страница 76

## 0% Цитат

Исключение цитат выключено

Исключение списка библиографических ссылок выключено

## 0% Исключений

Нет исключенных источников

## Модификации

Обнаружены модификации текста. Подробная информация доступна в онлайн-отчете.

Замененные символы 7

Подозрительное форматирование 19 страниц

1

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ  
ІМЕНІ ВАДИМА ГЕТЬМАНА**

**Навчально-науковий інститут  
«Інститут інформаційних технологій в економіці»**

**Кафедра інформаційних систем в економіці**

**ОСВІТНЬО-ПРОФЕСІЙНА ПРОГРАМА  
«Системи штучного інтелекту»**

галузь знань 12 Інформаційні технології  
спеціальність 122 Комп'ютерні науки

Форма навчання: очна (денна)

**КВАЛІФІКАЦІЙНА МАГІСТЕРСЬКА РОБОТА**

на тему: «Інтелектуальна система аналізу ринку роздрібної торгівлі»

здобувача Сіденко Владислава Вячеславовича

Науковий керівник: к.е.н., доцент Гордієнко І.В.

**Робота допущена до захисту перед екзаменаційною  
комісією з атестації здобувачів вищої освіти (ЕК)**

Завідувач кафедри: к.е.н., доцент Тішков Б.О.

**Київ 2024**

## ВИКОРИСТАННЯ КЛАСТЕРНОГО АНАЛІЗУ В РОЗДРІБНІЙ ТОРГІВЛІ

В сучасному ринковому середовищі роздрібна торгівля (рітейл) є однією з ключових галузей економіки, де вирішальне значення має ретельне вивчення та аналіз споживчого попиту. Різноманітність покупців та їхніх потреб обумовлюють необхідність використання ефективних стратегій асортиментного управління, спрямованих на максимізацію задоволення клієнтів та підвищення рентабельності підприємств роздрібною торгівлі. В цьому контексті кластерний аналіз виявляється важливим інструментом, що дозволяє визначити основні групи споживачів з подібними характеристиками та створити персоналізовані стратегії асортиментного управління для кожної з них.

Кластерний аналіз є статистичним методом аналізу даних, який використовується для групування об'єктів або спостережень у взаємоподібні групи, відомі як кластери. Метою кластерного аналізу є виявлення внутрішньої структури даних та визначення груп об'єктів, які подібні за певними характеристиками, але відрізняються від об'єктів інших груп [1].

У рітейлі кластерний аналіз дає змогу виявити закономірності у споживчому попиті та групувати магазини в кластери зі схожими характеристиками споживання. Це дозволяє ефективно адаптувати асортимент та маркетингові стратегії для кожного кластеру магазинів. Цей метод також використовується для сегментації клієнтів на групи з подібними характеристиками, як-от: споживчі звички, вік, доходи, географічне положення тощо [2].

У проведенні кластеризації варто приділяти увагу всім властивостям товарів, оскільки навіть такий аспект, як кількість упаковки, може бути вирішальним при прийнятті рішення про покупку. Наприклад, у магазинах на околицях міста часто краще продаються товари великими упаковками, ніж малими.

Вхідними даними для здійснення кластеризації є чеки та дані по програмам лояльності, на основі яких формується масив для його подальшої обробки.

Для кластерного аналізу даних пропонується використовувати алгоритм К-середніх, перевагами якого є простота у використанні та масштабування, що робить його ефективним на великих обсягах даних. Даний алгоритм підтримується багатьма програмними засобами для аналізу даних та машинного навчання він також легко оптимізується під платформи типу Apache Hadoop, які використовуються для ефективного зберігання та обробки великих наборів даних [3].

Розглянемо етапи процесу кластеризації для аналізу даних в рітейлі.

**1. Підготовка та очищення даних.** Необхідно видалити інформацію, яка не впливає на поведінку покупців. Наприклад, несуттєвою інформацією є придбання пакетів на касі.

**2. Формування матриці із вхідними даними.** Важливо зазначити, що результати кластеризації сильно залежать від періоду часу, за який вона проводиться. Наприклад, короткий період може дозволити побачити поточні тенденції продажів. Кластеризація за тривалий період дозволить отримати узагальнену картину, виявивши стійкі групи клієнтів з різними стилістичними особливостями. Зокрема, можна виділити групи, що характеризуються споживчими вподобаннями чи поведінкою, наприклад, «Студенти», «Домогосподарки», «Пенсіонери» і т. д. Крім того, роздрібним мережам важливо враховувати особливості програм лояльності. Наприклад, учасники програми можуть відрізнятися за своїми покупками та частотою їх здійснення. Отже, матрицю необхідно формувати, враховуючи основні цілі кластеризації.

**3. Вибір оптимальної кількості кластерів.** Кількість кластерів обирається на основі власного досвіду шляхом експериментів. Недостатньо мале число кластерів буде неефективним

і непоказовим, оскільки при цьому виникає лише один-два «мегакластера», що охоплюють 98% клієнтів, і кілька малих кластерів, які не несуть корисної інформації. З іншого боку, при великій кількості кластерів утворюється занадто багато малих груп.

**4. Здійснення кластеризації.** Для здійснення кластеризації пропонуємо використання алгоритму K-means на мові програмування Python з використанням бібліотеки scikit-learn, призначеної для кластерного аналізу [4].

**5. Аналіз результатів кластеризації.** Після застосування алгоритмів кластеризації всі об'єкти (наприклад, клієнти) розділяються на групи в залежності від їхніх характеристик. Клієнти з схожими характеристиками об'єднуються в один кластер. Якщо аналіз проводиться за певний період (наприклад, рік або місяць), то в кластеризації беруть участь лише ті клієнти, які здійснили покупки протягом вказаного періоду.

**6. Дослідження характеристик кожного кластера.** Кожному кластеру можна присвоїти назву, що відображає його особливості, наприклад, кластери «Домогосподарки», «Студенти», «Любителі тварин». Це допомагає ідентифікувати кластери залежно від їхнього поведінкового профілю. Обороти кожного кластера дозволяють визначити, які кластери приносять найбільший дохід. Доля кожного кластера в обороті відображає його вклад у загальний оборот, виражений у відсотках від загального обороту за вибраний період кластеризації. Кількість клієнтів у кожному кластері вказує на його розмір. Загальна кількість чеків у кожному кластері та кількість чеків на одного клієнта дозволяють відстежити частоту покупок кожного клієнта протягом обраного періоду кластеризації. Середня кількість товарів у чеку та середня вартість чеку дозволяють визначити, в якому кластері продаються найдорожчі товари.

**7. Здійснення персоналізованих розсилок для кожного кластера.** Використовуючи кластеризацію клієнтів, можна розробити чітку систему рекомендацій - який товар, якому клієнту і в який час пропонувати. Знання того, що і якій групі людей пропонувати, дозволяє компаніям уникнути непотрібних розсилок SMS або електронною поштою. Пропонуючи клієнтам лише ті товари, яких вони потребують (з урахуванням супутніх), можна досягти значного збільшення відгуку та конверсії у покупку.

Отже, кластерний аналіз у ритейлі допомагає вирішувати завдання асортиментного управління, персоналізації маркетингових стратегій та вироблення індивідуалізованих підходів до обслуговування клієнтів, що сприяє збільшенню продажів, підвищенню задоволення споживачів та підвищенню конкурентоспроможності торгового підприємства [2].

#### **Список використаних джерел**

1. Ситник В. Ф., Краснюк М.Т. Інтелектуальний аналіз даних (дейтамайнінг): Навч. посібник. К: КНЕУ, 2007. 376 с.

2. Як кластеризація допомагає підібрати правильний асортимент у категорії? [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: <https://www.c4r.eu/ukr/blog/biznes-konsalt/klasterizatsiya/>.

3. What is Hadoop? [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: [https://aws.amazon.com/what-is/hadoop/?nc1=h\\_ls](https://aws.amazon.com/what-is/hadoop/?nc1=h_ls).

4. Clustering [Електронний ресурс]. – 2024. – Режим доступу до ресурсу: <https://scikit-learn.org/stable/modules/clustering.html#clustering>.

**Науковий керівник:** Гордієнко І.В., к.е.н., доцент

## Рецензія

У цілому кваліфікаційна дипломна робота ЗВО Сіденка Владислава В'ячеславовича на тему «Інтелектуальна система аналізу ринку роздрібної торгівлі» виконана на належному рівні. По тексту роботи є несуттєві зауваження:

- окремі описки і помилки, які не зменшують підсумкової позитивної оцінки – це переважно незначні відхилення в термінології;
- наведені в додатках програмні коди значно б виграли, якщо б супроводжувались коментарями для ключових модулів.

Представлена на попередній захист кваліфікаційна дипломна робота відповідає у цілому вимогам вищої школи за спеціальністю «Комп'ютерні науки», може бути допущена до захисту і заслуговує на високу оцінку.

*Професор каф. ІСЕ, д.е.н.*

*Pinna С.П., 19.05.2024*