

Міністерство освіти і науки України
Луцький національний технічний університет
Факультет комп'ютерних та інформаційних технологій
Кафедра комп'ютерних наук

КВАЛІФІКАЦІЙНА РОБОТА
ЗА СТУПЕНЕМ ВИЩОЇ ОСВІТИ «МАГІСТР»

РОЗРОБКА ТА ДОСЛІДЖЕННЯ WEB-ДОДАТКУ
ДЛЯ ДАТА-МАЙНІНГУ В ОНЛАЙН-ТОРГІВЛІ

DEVELOPMENT AND RESEARCH OF A WEB APPLICATION
FOR DATA MINING IN ONLINE TRADING

спеціальність 122 Комп'ютерні науки

освітня програма «Комп'ютерні науки»

Виконав: здобувач вищої освіти
групи КНм-21
Гуменюк Назарій Олександрович

(підпис)

Керівник: к.т.н., доцент
Ліщина Валерій Олександрович

(підпис)

Кваліфікаційну роботу
допущено до захисту
«__» _____ 2025 р.
Гарант освітньої програми:
к.т.н., доцент
Ліщина Валерій Олександрович

(підпис)

Луцьк – 2025 року

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис	
		завдання видав	завдання прийняв
<i>Аналіз проблематики за темою роботи та постановка завдань дослідження</i>	<i>Ліщина В. О.</i>		
<i>Теоретичне дослідження та практична реалізація предмету дослідження</i>	<i>Ліщина В. О.</i>		
<i>Експериментальне дослідження результативності предмету дослідження</i>	<i>Ліщина В. О.</i>		
<i>Показник запозичень тексту</i>	%		
<i>Інструментальна перевірка</i>	<i>Кошелюк В. А.</i>		
<i>Нормоконтроль</i>	<i>Сачук В. О.</i>		
<i>Гарант ОПП</i>	<i>Ліщина В. О.</i>		

7. Дата видачі завдання *«14» травня 2025 р.*

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи бакалавра	Строк виконання етапів роботи	Примітка
1	<i>Провести огляд літературних джерел по темі кваліфікаційної роботи</i>	<i>до 30.06.2025 р</i>	
2	<i>Провести аналіз загальної проблеми і вибір напрямків дослідження</i>	<i>до 01.09.2025 р.</i>	
3	<i>Розробити функціональну схему роботи програмного продукту</i>	<i>до 01.10.2025 р</i>	
4	<i>Описати засоби розробки об'єкта проектування</i>	<i>до 15.10.2025 р.</i>	
5	<i>Практична реалізація об'єкта проектування</i>	<i>до 10.11.2025 р.</i>	
6	<i>Провести експериментальне дослідження результативності предмету дослідження</i>	<i>до 25.11.2025 р.</i>	
7	<i>Здача чистового варіанту кваліфікаційної роботи бакалавра на кафедрі</i>	<i>до 05.12.2025 р.</i>	

Здобувач вищої освіти _____ Назарій ГУМЕНЮК

Керівник роботи _____ Валерій ЛІЩИНА

АНОТАЦІЯ

Гуменюк Н. О. Розробка та дослідження web-додатку для дата-майнінгу в онлайн-торгівлі. Рукопис.

Кваліфікаційна робота магістра ОП «Комп'ютерні науки». Луцький національний технічний університет. Луцьк, 2025.

Кваліфікаційна робота магістра складається з вступу, 3 розділів, висновків, списку використаних джерел та додатків.

У роботі представлено розробку та дослідження web-додатку для дата-майнінгу в онлайн-торгівлі. Система забезпечує автоматизований збір, підготовку, аналіз і візуалізацію даних електронної комерції з використанням сучасних методів data-mining. Реалізовано модулі кластеризації, пошуку асоціативних правил, прогнозування продажів та виявлення аномалій. Запропонована архітектура поєднує транзакційні дані з аналітичним ядром, що дозволяє подолати ефект «аналітичного паралічу» та підвищити ефективність управлінських рішень у сфері e-commerce. Результатом є функціональний прототип web-додатку, здатний використовуватись для підтримки бізнес-аналітики та планування в онлайн-торгівлі.

Ключові слова: web-додаток, дата-майнінг, комерція, модуль, бізнес-аналітика.

ABSTRACT

Nazarii Humeniuk. Development and research of a web application for data mining in online trading. Manuscript.

Master's qualification work in specialty 122 «Computer Science», educational program «Computer Science». Lutsk National Technical University. Lutsk, 2025.

The master's thesis consists of an introduction, 3 chapters, conclusions, a list of references, and appendices.

The thesis presents the development and research of a web application for data mining in online commerce. The system provides automated collection, preparation, analysis, and visualization of e-commerce data using modern data mining methods. Modules for clustering, associative rule discovery, sales forecasting, and anomaly detection have been implemented. The proposed architecture combines transactional data with an analytical core, which allows overcoming the effect of analytical paralysis and increasing the effectiveness of management decisions in the field of e-commerce. The result is a functional prototype of a web application that can be used to support business analytics and planning in online commerce.

Keywords: web application, data mining, commerce, module, business analytics.

ЗМІСТ

ВСТУП	7
РОЗДІЛ 1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА МЕТОДІВ DATA-MINING В ОНЛАЙН-ТОРГІВЛІ	10
1.1 Аналітика даних у бізнесі, поняття, методичні підходи.....	10
1.2 Проблема «аналітичного паралічу» в онлайн-торгівлі	13
1.3 Огляд методів, алгоритмів та інструментів data-mining для онлайн-торгівлі	15
1.4 Постановка завдання дослідження.....	17
РОЗДІЛ 2 ПРОЄКТУВАННЯ ТА РОЗРОБКА WEB-ДОДАТКУ ДЛЯ ДАТА- МАЙНІНГУ В ОНЛАЙН-ТОРГІВЛІ	19
2.1 Загальна концепція системи (розширено)	19
2.2 Функціональна модель системи.....	21
2.3 Архітектура програмної системи	25
2.4 Структура даних та модель бази даних	29
2.5 Вибір технологій та засобів реалізації web-додатку	32
2.6 Практична реалізація ключових модулів та структура проєкту	35
2.7 Інтерфейс користувача та приклади ключових UI-компонентів	38
РОЗДІЛ 3 ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ РЕЗУЛЬТАТИВНОСТІ WEB-ДОДАТКУ ДЛЯ ДАТА-МАЙНІНГУ В ОНЛАЙН-ТОРГІВЛІ.....	44
3.1 Методика проведення дослідження	44
3.2 Обробка та аналіз отриманих результатів	46
ВИСНОВКИ.....	52
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	54
ДОДАТКИ.....	56

ВСТУП

Сучасний ринок онлайн-торгівлі характеризується швидким ростом обсягів даних, що формуються під час взаємодії користувачів із торговими платформами. Виникає потреба у інструментах, здатних здійснювати автоматизований data-mining, виявляти закономірності, прогнозувати попит та підтримувати аналітичні рішення у бізнесі. У багатьох компаніях спостерігається явище «аналітичного паралічу», коли даних багато, але інсайтів – недостатньо, що ускладнює прийняття управлінських рішень.

Стрімкий розвиток цифрових технологій та глобалізація економічних процесів докорінно змінили ландшафт сучасної торгівлі. Електронна комерція (e-commerce) трансформувалася з альтернативного каналу збуту в домінуючу модель ведення бізнесу, що генерує колосальні обсяги даних щосекунди. Сучасний бізнес-світ характеризується безпрецедентним накопиченням інформації: від транзакційних записів та логів поведінки користувачів до деталізованої історії ланцюжків постачання. У цьому контексті дані перестають бути просто побічним продуктом діяльності, а перетворюються на стратегічний актив, який, за умови правильного використання, здатний забезпечити суттєві конкурентні переваги.

Проте зростання обсягів доступної інформації не завжди корелює з підвищенням ефективності прийняття рішень. Навпаки, значна кількість організацій стикається з феноменом, відомим як «аналітичний параліч» (analysis paralysis). Цей стан характеризується нездатністю приймати своєчасні та обґрунтовані управлінські рішення не через брак інформації, а через її надмірність, неструктурованість та складність для когнітивної обробки людиною. Числа без контексту, розрізнені масиви даних та відсутність єдиної аналітичної платформи створюють інформаційний шум, що блокує стратегічне мислення менеджменту.

В умовах висококонкурентного ринку онлайн-торгівлі здатність компанії швидко трансформувати «сирі» дані в цінні інсайти (data-driven insights) стає

критичним фактором виживання. Проблема полягає у тому, що традиційні методи обробки даних часто виявляються неефективними при роботі з великими обсягами (Big Data), а готові BI-рішення (Business Intelligence) можуть бути занадто дорогими або недостатньо гнучкими для специфічних потреб конкретного бізнесу. Тому розробка спеціалізованих веб-додатків, що інтегрують потужні інструменти дата-майнінгу (такі як Apache Spark) та надають зручний інтерфейс для візуалізації результатів, є надзвичайно актуальним завданням. Такі системи дозволяють автоматизувати рутинні процеси аналізу, виявляти приховані закономірності у поведінці клієнтів, прогнозувати попит та, як наслідок, долати бар'єр «аналітичного паралічу».

Актуальність теми полягає у необхідності створення програмного рішення, яке забезпечить автоматизований збір, обробку, аналіз та візуалізацію даних онлайн-торгівлі, тим самим підвищуючи якість аналітичних процесів та ефективність бізнес-рішень.

Мета роботи – створити та дослідити веб-додаток для дата-майнінгу в онлайн-торгівлі з використанням сучасних інструментів аналізу великих даних.

Об'єкт дослідження. Процеси аналітичної обробки та візуалізації даних у сфері онлайн-торгівлі.

Предмет дослідження – методи дата-майнінгу, алгоритми обробки даних та архітектурні рішення веб-систем для онлайн-торгівлі.

Завдання роботи:

- провести огляд проблем аналітики в e-commerce;
- проаналізувати існуючі підходи та інструменти data-mining; розробити архітектуру веб-додатку;
- реалізувати модуль аналізу й агрегації даних;
- провести експериментальне дослідження роботи системи;
- оцінити ефективність застосованих алгоритмів.

Наукова новизна. Запропоновано комплексний підхід до інтегрованого дата-майнінгу у веб-системі онлайн-торгівлі з використанням автоматизованої генерації, агрегації та візуалізації даних.

Практична цінність. Результатом є прототип web-додатку, що може застосовуватись у реальних торговельних системах для прогнозування попиту, оцінки динаміки продажів та підвищення якості управлінських рішень.

РОЗДІЛ 1

АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА МЕТОДІВ DATA-MINING В ОНЛАЙН-ТОРГІВЛІ

1.1 Аналітика даних у бізнесі, поняття, методичні підходи.

Аналітика даних у системах електронної комерції є однією з ключових складових сучасного цифрового бізнесу. Під час здійснення онлайн-продажів формується значний обсяг інформації: історія транзакцій, поведінкові характеристики користувачів, дані про товари, маркетингові метрики, цінові зміни, кількість повернень, швидкість обробки замовлення, джерела трафіку тощо. Такий масив інформації формує основу для методів аналітичної обробки, прогнозування та автоматизації стратегічних рішень [1].

Аналітика даних визначається як процес послідовного збору, очищення, перетворення, аналізу та візуального представлення інформації з метою отримання нових знань, закономірностей та висновків, що дозволяють оптимізувати бізнес-процеси. Вона охоплює як статистичні методи, так і алгоритмічні підходи машинного навчання, технології великих даних, методи кластеризації, пошук асоціативних правил, аналіз часових рядів, сегментацію клієнтів та рекомендаційні моделі.

У сфері онлайн-торгівлі роль аналітики є визначальною. На відміну від офлайн-ритейлу, де поведінка покупця часто залишається «чорною скринькою», онлайн-торгівля дозволяє фіксувати кожен крок клієнта: від джерела переходу на сайт до часу, проведеного на сторінці конкретного товару. Це створює передумови для глибокого розуміння споживацьких потреб та оптимізації бізнес-процесів.

Основними підходами до бізнес-аналітики, які сьогодні виділяють експерти, є:

– дескриптивна аналітика (Descriptive Analytics): це фундамент будь-якої аналітичної системи. Вона відповідає на питання «Що відбулося?». Методи дескриптивної аналітики фокусуються на агрегації історичних даних та їх

візуалізації. В контексті нашого дослідження, це звіти про загальний обсяг продажів, кількість нових клієнтів, середній чек тощо. Хоча цей підхід констатує факти минулого, він є необхідним для розуміння поточного контексту [2-3];

– діагностична аналітика (Diagnostic Analytics): спрямована на розуміння причин – «Чому це відбулося?». Вона передбачає більш глибоке занурення в дані (drill-down), пошук кореляцій та аномалій. Наприклад, якщо продажі впали, діагностична аналітика допомагає з'ясувати, чи це пов'язано з сезонністю, технічними проблемами на сайті чи діями конкурентів;

– предиктивна аналітика (Predictive Analytics): це рівень, на якому дані починають працювати на випередження. Вона відповідає на питання «Що може статися у майбутньому?». Використовуючи статистичні моделі, алгоритми машинного навчання та аналіз часових рядів, предиктивна аналітика дозволяє прогнозувати майбутні тренди, попит на товари та ймовірну поведінку клієнтів. Саме цей аспект є ключовим у даній кваліфікаційній роботі, оскільки дозволяє бізнесу переходити від реактивного до проактивного управління;

– прескриптивна аналітика (Prescriptive Analytics): найвищий рівень зрілості, що відповідає на питання «Що потрібно зробити?». Вона не просто прогнозує майбутнє, а й рекомендує конкретні дії для досягнення бажаних результатів (наприклад, автоматична зміна цін або замовлення певної партії товару).

Інтеграція цих підходів у єдину систему дозволяє створити цілісну картину бізнесу. Проте, як показує практика, саме перехід від накопичення даних до їх ефективного використання є найскладнішим етапом для більшості компаній.

У більшості наукових праць автори сходяться на тому, що ефективність управлінських рішень у бізнесі прямо залежить від якості та швидкості обробки даних. Наявність інформації без інструментів для її інтерпретації не приносить користі – компанія отримує «сірі дані», але не знання. Саме аналіз, моделювання та автоматизація дозволяють перетворювати дані на цінність [4].

Аналіз наукових публікацій та досліджень демонструє кілька еволюційних напрямів розвитку аналітики в онлайн-торгівлі (табл. 1.1).

Таблиця 1.1 – Напрями розвитку аналітики в онлайн-торгівлі

Підхід	Характерні риси	Типові результати
Статистичний аналіз	Регресія, кореляція, тренди	прогнозування продажів, сезонність
Data Mining	Кластеризація, асоціативні правила, виявлення патернів	сегментація клієнтів, аналіз попиту
Big Data Analytics	Масштабовані обчислення, обробка потоків даних	робота з великими e-commerce масивами
Predictive Modelling	ML-алгоритми, часові ряди, нейромережі	прогноз ринку, рекомендаційні системи
Visualization Analytics	Дашборди, інтерактивні графіки	швидке ухвалення рішень, скорочення аналітичних циклів

Дані практичних досліджень показують, що впровадження аналітики у сферу електронної комерції забезпечує [5]:

- зростання точності прогнозу попиту й товарного обігу;
- своєчасне виявлення падіння або пікового росту продажів;
- покращення управління складськими залишками;
- збільшення конверсії за рахунок персоналізації рекомендацій;
- скорочення ризику дефіциту або надлишку товару;
- підвищення швидкості прийняття рішень менеджерами.

Таким чином, сучасні дослідження підкреслюють необхідність переходу від базового збору інформації до автоматизованого аналізу, прогнозування та візуального подання результатів. Саме ці підходи й формують методичне підґрунтя для розвитку систем дата-майнінгу в e-commerce.

На момент виконання роботи стан розвитку цифрової торгівлі характеризується двома паралельними тенденціями: обсяг даних зростає швидше, ніж здатність бізнесу їх аналізувати; існуючі системи збору інформації не забезпечують достатньо швидкого та структурованого перетворення даних у рішення.

Це створює явище, яке науковці називають аналітичним паралічем – ситуацію, коли даних багато, але висновків немає. Підприємство бачить показники, але не розуміє інсайти, не може прогнозувати попит, визначити провальні чи успішні товарні позиції, оцінити ефективність маркетингових кампаній та своєчасно оптимізувати запаси [6].

Отже, узагальнення проведеного огляду дозволяє сформулювати ключове завдання дослідження:

- побудувати web-систему, яка автоматизує процеси аналізу та візуалізації даних онлайн-торгівлі;
- розробити алгоритмічний модуль для аналітичної обробки великих масивів транзакцій;
- забезпечити швидке отримання результатів у форматі, що доступний для менеджерів.

Саме це визначає актуальність і доцільність розробки, реалізованої в даній кваліфікаційній роботі.

1.2 Проблема «аналітичного паралічу» в онлайн-торгівлі

Швидкий розвиток онлайн-торгівлі призвів до ситуації, коли здатність бізнесу накопичувати дані значно перевищує його здатність їх осмислювати та використовувати у вигляді стратегічних рішень. У більшості компаній формується парадоксальна картина: інформації стає більше – але якість управління не зростає. Цей стан визначають як аналітичний параліч – ситуацію, у якій надмір даних і недостатність інструментів їх аналітичної інтерпретації блокують процес прийняття рішень [7].

Аналітичний параліч виникає не тоді, коли даних бракує, а навпаки – тоді, коли аналізованої інформації настільки багато, що вона перестає перетворюватися на знання. Руководитель, маючи перед собою десятки таблиць, звітів, графіків, метрик, може бачити зміни, але не розуміти їх причин і не здатний оперативно реагувати. Дані стають складнішими, глибшими та фрагментованішими, втрачається здатність розрізняти тенденції, а пошук причинно-наслідкових зв'язків перетворюється на трудомісткий процес. Аналітичний параліч – це момент, коли дані є, але відповіді немає; коли звіти створені, але рішення відкладаються; коли бізнес володіє інформацією, але позбавлений можливості використати її вчасно.

Цьому сприяє кілька взаємопов'язаних явищ. По-перше, гетерогенність джерел даних: CRM містить історію взаємодій з клієнтами, база транзакцій – фінансові операції, аналітичні сервіси фіксують поведінку користувачів, складські системи відстежують залишки, а маркетингові платформи – покази реклами та кліки. Кожен із цих фрагментів має цінність, але сукупно вони утворюють інформаційний лабіринт. По-друге, значні затримки між збором даних і їх аналізом призводять до того, що управлінські рішення приймають у момент, коли ситуація вже змінилася. Якщо формування звіту займає години або дні, а поведінка ринку змінюється протягом хвилин – аналітика втрачає сенс. По-третє, навіть правильні цифри не завжди перетворюються на висновки – надлишок параметрів, показників та графіків призводить до перевантаження мислення, коли замість вибору стратегії підприємство відкладає рішення «на потім». Це – когнітивна форма аналітичного паралічу, коли страх помилки підсилюється достатком варіантів [8].

Наслідки такої ситуації відчутні на всіх рівнях управління. Бізнес реагує на ринкові зміни із запізненням, не встигає перехоплювати тренди, неефективно розподіляє маркетингові бюджети, стикається з дефіцитом чи надлишком товарів на складах. Коли відсутній інструмент, що узагальнює патерни, прогнозує попит і формує оперативну картину стану продажів, компанія переважно працює реактивно – відповідає на події після їх настання, а не до. У такий спосіб аналітичний параліч стає каталізатором уповільнення бізнес-процесів, скорочення прибутковості та втрати конкурентних можливостей.

Сутність проблеми можна виразити просто: швидкість генерації даних росте нелінійно, тоді як швидкість їх аналітичної обробки залишається лінійною або навіть ручною. Підприємство збільшує обсяг даних, інтегрує нові канали продажів, маркетингові інструменти, але пропускна здатність аналітичної системи – людської чи технічної – не встигає адаптуватися. Таким чином утворюється інформаційна асиметрія: дані існують, але не працюють, рішення потрібні, але не приймаються, можливості є, але не реалізуються [9].

Подолати аналітичний параліч можливо лише переходом від пасивного збору до активного осмислення даних. Це потребує інтегрованого середовища, яке не просто зберігає інформацію, а перетворює її на готові аналітичні сценарії: сегментацію покупців, правила асоціацій, рекомендаційні моделі, аналіз сезонності, пошук кореляцій, прогнозування. Візуалізація має відбуватися не у вигляді багатосторінкових таблиць, а в компактних динамічних графіках, heat-map-відображенні, інтерактивних дашбордах. Web-інтерфейс повинен забезпечувати доступ до результатів у реальному часі, а система – виконувати аналіз без участі людини, автоматично агрегуючи, узагальнюючи та прогножуючи.

Таким чином, аналітичний параліч – це не просто технічна, а системна бізнес-проблема, яка виникає через розрив між потенціалом даних і здатністю підприємства їх осмислити. Його подолання можливе лише через побудову комплексної архітектури – єдиної точки входу до інформації, автоматизованого циклу data-mining, алгоритмічних прогнозних моделей і інтерактивної візуалізації, що дозволяє перетворювати факти на рішення. Подальші розділи роботи спрямовані на створення саме такої системи [10].

1.3 Огляд методів, алгоритмів та інструментів data-mining для онлайн-торгівлі

Data-mining у сфері онлайн-торгівлі охоплює комплекс алгоритмічних підходів, спрямованих на виявлення закономірностей у транзакційних масивах, побудову прогнозних моделей та автоматизоване формування управлінських висновків. У сучасному e-commerce щодня генеруються десятки тисяч операцій, переходів, клієнтських взаємодій, що робить аналітику нерозривним елементом бізнес-моделювання. Основою інтелектуальної обробки є вміння виділити структуру в хаотичних даних, знизити інформаційний шум і перетворити числові ряди на практичні інсайти [11].

Методи data-mining умовно можна об'єднати за завданнями, які вони вирішують: групування, класифікація, виявлення правил, прогнозування поведінки та аномалій. Алгоритми кластеризації, такі як k-means, DBSCAN та ієрархічні методи, дозволяють ділити споживачів на сегменти за параметрами активності, середнього чеку, частоти покупок та чутливості до знижок. На практиці це дає змогу формувати індивідуальні пропозиції, впливати на поведінку цільових груп та автоматизувати маркетингові розсилки. Класифікаційні моделі – логістична регресія, дерева рішень, SVM, ансамблеві Random Forest – застосовуються для прогнозу ймовірності купівлі та визначення ризикових транзакцій. Система отримує нові дані про замовлення й визначає їхню приналежність до категорій «вірогідно завершиться покупкою» або «потребує верифікації», що знижує втрати від шахрайства та зменшує потребу в ручній модерації [12].

Окрему роль відіграє аналіз асоціативних правил, що визначає товари, які часто купують разом. Алгоритми Apriori та FP-Growth дозволяють побудувати правила типу «якщо обрано ноутбук, імовірність покупки миші зростає удвічі» – такі залежності формують основу рекомендаційних механізмів, up-sale та cross-sale-стратегій. У більш розвинених системах ці моделі працюють у режимі online-learning, реагуючи на трендові зміни в реальному часі. Для прогнозування обсягів продажів, сезонності та пікових навантажень застосовують моделі часових рядів – ARIMA, Prophet, LSTM-мережі. Їхня цінність у можливості передбачати поведінку попиту, забезпечувати оптимізацію складу, визначати періоди завантаження кур'єрів і логістичних ресурсів [13].

Алгоритми лише тоді дають результат, коли вони підтримані технічною інфраструктурою. У великих системах кількість даних перевищує можливості одномашинної обробки, тому використовуються розподілені обчислювальні платформи. Hadoop забезпечує масштабоване сховище й пакетну обробку великих наборів інформації, а Spark працює безпосередньо в оперативній пам'яті, що прискорює агрегацію, SQL-запити та машинне навчання у десятки разів. У потоковій аналітиці застосовують Kafka, яка дозволяє обробляти події з

мілісекундними затримками – це критично для аналізу ціноутворення, моніторингу фроду та відстеження поведінкових патернів. Мови програмування є інструментальним рівнем аналізу: Python домінує завдяки Pandas, NumPy, SciKit-Learn, TensorFlow та PyTorch, R використовується для статистичних моделей, а Scala та Java інтегруються зі Spark у високопродуктивних системах.

Утім, жодна модель не має цінності без можливості подати її результат людині зрозуміло та швидко. Саме тому системи візуалізації – Tableau, Power BI, Grafana, а також web-бібліотеки Chart.js, D3.js – стали ключовим компонентом аналітики. Графічне представлення даних скорочує час сприйняття, дає змогу порівнювати тренди за секунди та ухвалювати рішення не постфактум, а в режимі поточного часу [14].

Таким чином, ефективний аналітичний процес в e-commerce поєднує кілька шарів: збирання транзакцій, їх структурування, машинне навчання, оцінку закономірностей, прогнозування майбутньої поведінки та представлення результатів у придатному для рішень вигляді. Web-додаток аналітики має інтегрувати ці рівні в єдину систему, де дані перетворюються на знання, а знання – на управлінські дії. У межах цієї роботи саме такий підхід обрано як основний напрям розроблення інтелектуального програмного рішення для дата-майнінгу в онлайн-торгівлі.

1.4 Постановка завдання дослідження

Аналіз предметної області, розгляд методів data-mining для онлайн-торгівлі та виявлення проблеми аналітичного паралічу дозволяють сформулювати чітке наукове завдання цієї роботи. З огляду на високі обсяги транзакційних даних, нерівномірність їх надходження, фрагментацію джерел та обмеженість традиційних засобів аналітики, виникає потреба у створенні інтегрованої web-системи, здатної автоматично збирати, обробляти, агрегувати й інтерпретувати інформацію у вигляді управлінських висновків.

Основна проблема, що потребує вирішення, полягає у розриві між кількістю доступних даних і здатністю бізнесу використовувати їх у стратегічних рішеннях. Це визначає необхідність побудови засобу, який забезпечить перехід від реактивної моделі управління (аналіз подій) до проактивної (виявлення тенденцій критичних змін). Така трансформація можлива за умови впровадження автоматизованого data-mining з інструментами прогнозування динаміки продажів, виявлення асоціативних залежностей, кластеризації клієнтів і процесів та інтерактивної візуалізації результатів.

У межах дослідження вирішується комплекс завдань, що охоплює увесь цикл інформаційної аналітики – від збору сирих даних до формування готових рекомендацій для бізнесу:

Завдання дослідження:

- здійснити теоретичний аналіз методів data-mining та їх застосування в електронній торгівлі;
- дослідити фактори виникнення аналітичного паралічу та визначити критерії його подолання;
- розробити архітектуру web-системи для інтегрованого аналізу даних онлайн-продажів;
- сформуванати модель обробки та зберігання транзакційних наборів із можливістю масштабування;
- реалізувати програмний модуль дата-майнінгу з підтримкою прогнозування та виявлення закономірностей;
- розробити інтерфейс візуалізації, що забезпечує швидке прийняття управлінських рішень;
- провести експериментальне дослідження розробленої системи та оцінити її ефективність.

Визначення мети і завдань дослідження створює методологічну основу для переходу до проєктування та реалізації програмної системи.

РОЗДІЛ 2

ПРОЄКТУВАННЯ ТА РОЗРОБКА WEB-ДОДАТКУ ДЛЯ ДАТА-МАЙНІНГУ В ОНЛАЙН-ТОРГІВЛІ

2.1 Загальна концепція системи (розширено)

Розроблювана система є web-додатком аналітичного типу, призначеним для автоматизованого дата-майнінгу в онлайн-торгівлі. Її концепція базується на ідеї створення єдиного інтегрованого середовища, у межах якого дані електронної комерції перетворюються на практично значущі управлінські інсайти. Система орієнтована на подолання ключової проблеми сучасного e-commerce – розриву між накопиченням великих масивів інформації та здатністю бізнесу оперативно її аналізувати і використовувати для прийняття рішень.

У більшості реальних торговельних платформ інформація про користувачів, товари, замовлення, маркетингові активності та поведінкові події існує у різних джерелах і з різною структурою. Це зумовлює фрагментацію аналітики та збільшує час отримання результатів. Запропонований web-додаток концептуально спрямований на те, щоб звести ці дані до єдиної логіки обробки, забезпечити уніфікований аналітичний контур і реалізувати повний цикл: збір → підготовка → аналіз → моделювання → візуалізація → рекомендації.

Головною особливістю системи є те, що вона розглядає аналітику не як пасивний інструмент статистичної звітності, а як активну інтелектуальну функцію. Кінцевою метою не є створення набору таблиць або історичних агрегатів. Натомість, система повинна генерувати результати, які можуть бути безпосередньо використані бізнесом: оцінка динаміки продажів, прогноз попиту, виявлення сезонних ефектів, визначення груп покупців з подібною поведінкою, аналіз співпокупок товарів та раннє виявлення ризикових аномалій у транзакціях. Таким чином, web-додаток виконує роль прикладної аналітичної платформи для підтримки рішень на рівні менеджменту, маркетингу, логістики та комерційної стратегії.

Концептуальна модель роботи системи передбачає наявність двох рівнів функціонування. Перший рівень – операційний, який забезпечує стабільний збір і оновлення базових даних: інформації про замовлення, товарні позиції, клієнтів, часові параметри продажів, дані про склад і ціни. Другий рівень – аналітичний, що реалізує алгоритмічне перетворення цих даних у знання. Саме на цьому рівні застосовуються інструменти data-mining, прогнозування та аналітичної агрегації, а результати подаються у форматі інтерактивних графіків, дашбордів і рекомендаційних блоків [15].

Цільовими користувачами системи можуть виступати бізнес-аналітики, менеджери з продажів, маркетологи, керівники інтернет-магазинів та адміністратори торговельних платформ. Для кожної з цих груп важливо забезпечити різні сценарії використання. Аналітик потребує гнучких інструментів фільтрації та порівняння наборів даних, менеджер – швидких КРІ-індикаторів та прогнозів, маркетолог – сегментації та рекомендацій щодо поведінки клієнтів, тоді як адміністратор відповідає за коректність імпорту та підтримку стабільності сервісу. Така різноманітність сценаріїв визначає необхідність побудови системи з модульною логікою бізнес-функцій та універсальним web-інтерфейсом.

Окремо слід виділити вимоги до якості та експлуатаційних характеристик системи. Вона має бути масштабованою з погляду збільшення кількості транзакцій та користувацьких запитів, стабільною при роботі з великими наборами даних, а також здатною забезпечувати прийнятний час відповіді для аналітичних операцій. Важливими є вимоги до безпеки та контролю доступу, оскільки об'єктом обробки є комерційно чутлива інформація. Тому концепція системи передбачає застосування ролей користувачів, журналювання звернень та обмеження доступу до критичних наборів даних.

З позиції інженерії програмного забезпечення web-додаток доцільно реалізовувати із чітким розділенням відповідальності між компонентами: серверний модуль виконує завдання збору, підготовки та аналітичних обчислень, тоді як клієнтський модуль фокусується на представленні результатів у

зрозумілому інтерфейсному вигляді. Такий підхід забезпечує підтримуваність та дозволяє розширювати функціонал системи шляхом додавання нових алгоритмічних сценаріїв без істотної перебудови базової архітектури.

Таким чином, загальна концепція розроблюваної системи полягає у створенні web-платформи, яка інтегрує методи дата-майнінгу та інструменти візуалізації з транзакційними даними онлайн-торгівлі, забезпечуючи оперативний перехід від даних до управлінських рішень. Це формує основу для подальшого проектування архітектури, моделі даних, вибору технологічного стеку та реалізації функціональних модулів.

2.2 Функціональна модель системи

Функціональна модель розроблюваного web-додатку визначає набір сервісів і сценаріїв, які забезпечують повний цикл інтелектуальної аналітики даних онлайн-торгівлі. Система проектується як інтегрована платформа, що трансформує різноманітні набори транзакційної та поведінкової інформації у прикладні аналітичні результати. Її функціонування організовано так, щоб користувач отримував не тільки статистичні зведення, а й інтерпретовані висновки: сегменти клієнтів, тренди продажів, індикатори попиту, правила спільних покупок, результати прогнозування та автоматизовані рекомендації.

На концептуальному рівні система реалізує такі групи функцій: інтеграція даних, підготовка та валідація наборів, аналітична агрегація, запуск алгоритмів data-mining, візуальне представлення результатів та сервісне керування аналітичними сценаріями. Важливою вимогою є підтримка повторюваності аналітичних експериментів, що означає можливість зберігати налаштування моделей, параметри навчання, структуру вхідних датасетів та результати обчислень для подальшого порівняння.

Система передбачає кілька ролей користувачів. Бізнес-аналітик здійснює поглиблену роботу з даними, запускає моделі, порівнює результати та формує звіти. Менеджер з продажів отримує агреговані показники, прогнози і короткі

інсайти для оперативних управлінських рішень. Маркетолог використовує сегментацію, аналіз поведінки та рекомендаційні блоки для налаштування кампаній. Адміністратор відповідає за коректність імпорту даних, налаштування доступів, контроль стабільності та журналювання дій.

З погляду сценаріїв використання, система повинна забезпечувати такі типові дії:

- завантаження або підключення джерел даних;
- автоматизоване очищення, нормалізацію та формування структурованих таблиць;
- побудову агрегатів за часовими зрізами;
- запуск алгоритмів сегментації й аналізу попиту;
- генерацію асоціативних правил та рекомендацій;
- прогнозування продажів на основі історичних рядів;
- виявлення аномалій у транзакціях;
- перегляд результатів у вигляді інтерактивних графіків і дашбордів;
- експорт звітів для подальшого використання в управлінських процесах.

Функціональна модель також передбачає відокремлення ядра аналітики від інтерфейсної частини. Ядро виконує підготовку та обчислення, а клієнтський рівень забезпечує інтерпретацію результатів у зрозумілому форматі. Такий підхід дозволяє масштабувати аналітичні модулі та незалежно розвивати візуалізаційні компоненти.

Нижче наведено ключові рисунки, що відображають функціональну структуру та логіку взаємодії користувачів із системою.

Зобразимо діаграма прецедентів web-додатку для дата-майнінгу (рис. 2.1).

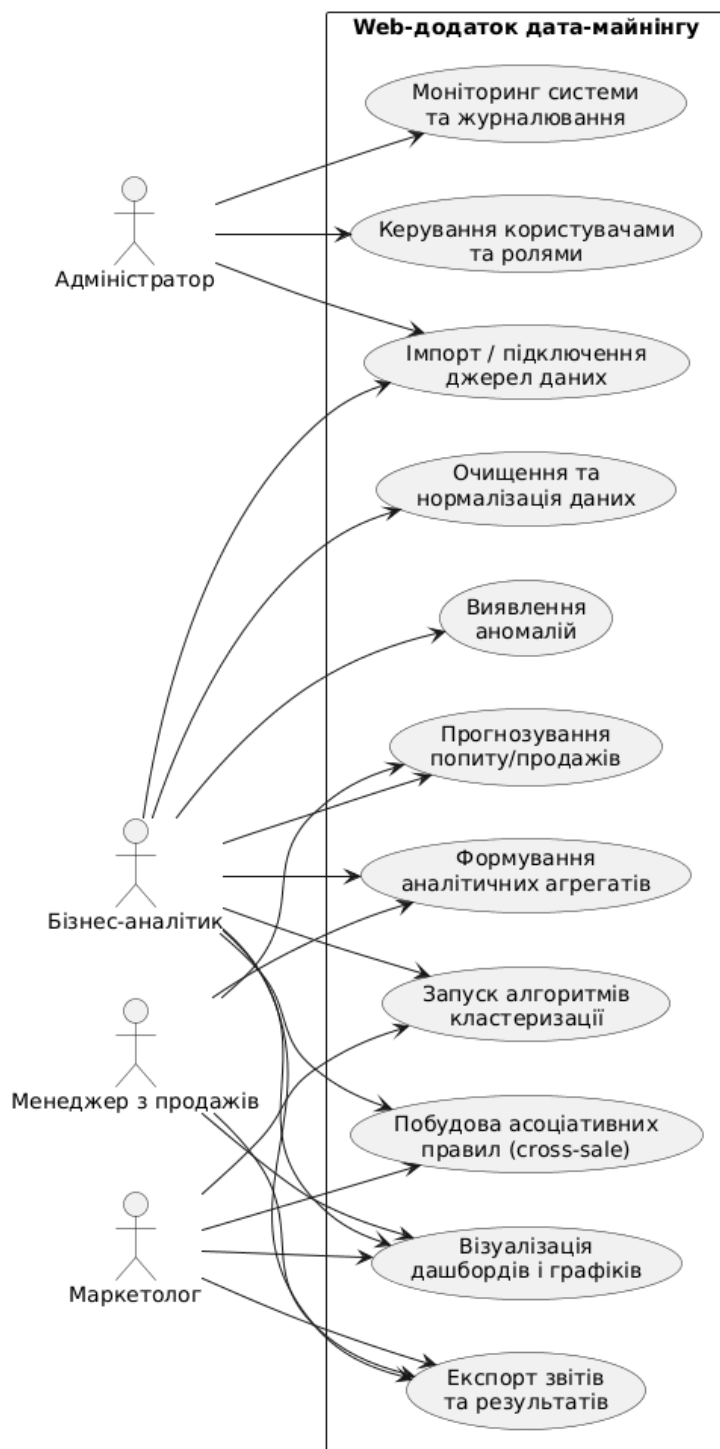


Рисунок 2.1 – Діаграма прецедентів web-додатку для дата-майнінгу

Зобразимо контекстну діаграму потоків даних (DFD Level 0) (рис. 2.2).

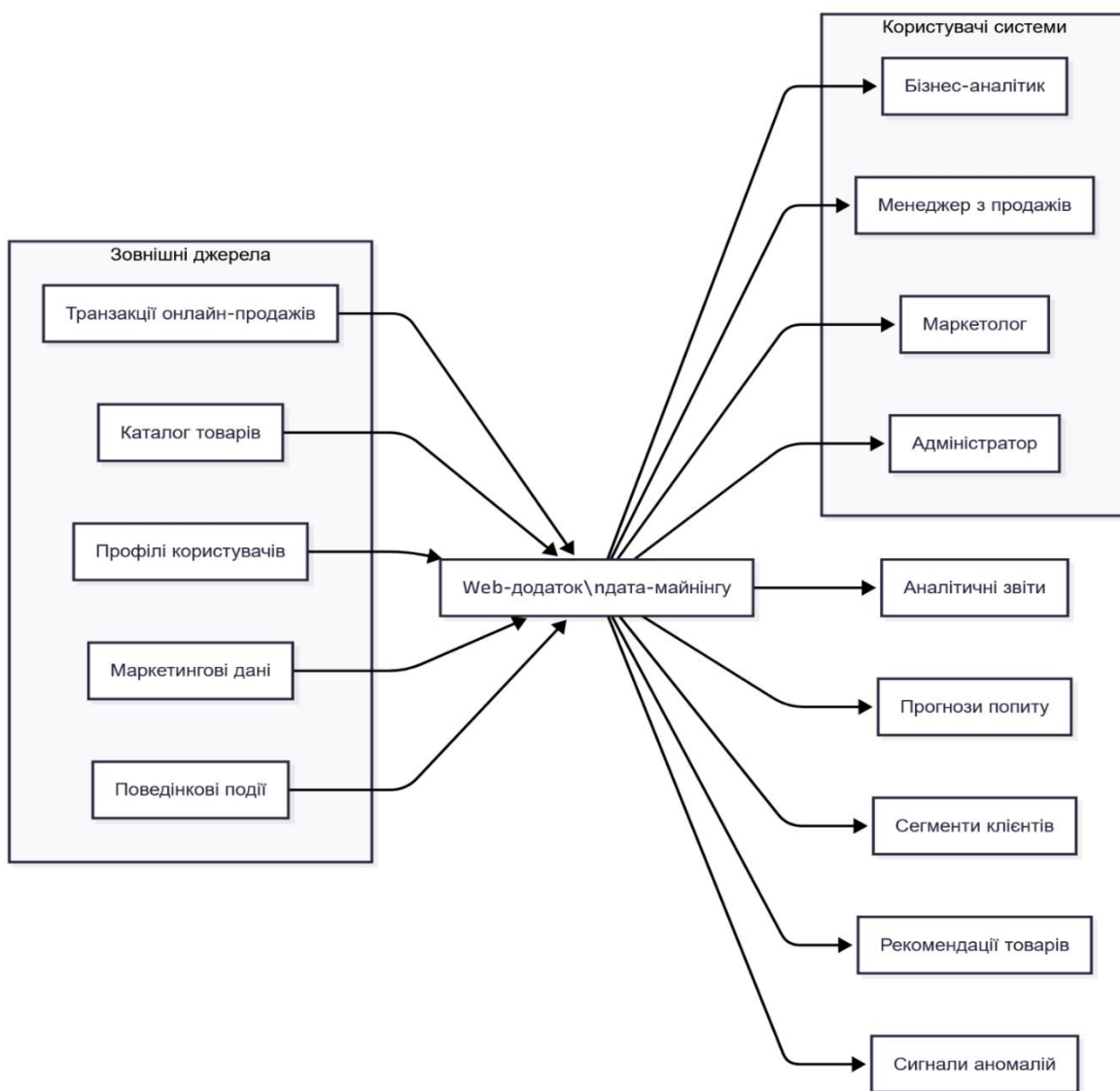


Рисунок 2.2 – Контекстна діаграма потоків даних (DFD Level 0)

Функціональна декомпозиція системи показана на рисунку 2.3.

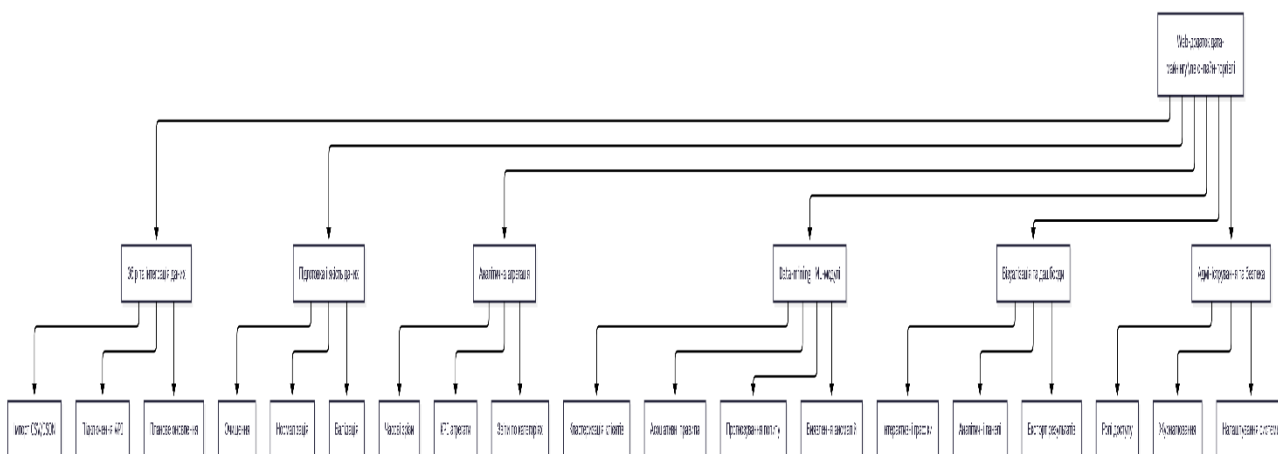


Рисунок 2.3 – Функціональна декомпозиція системи

Узагальнений процес аналітичного циклу показано рисунку 2.4.



Рисунок 2.4 – Узагальнений процес аналітичного циклу

Представлена функціональна модель демонструє, що система орієнтована на повний життєвий цикл аналітики даних e-commerce. Ключем до практичної цінності є поєднання автоматизованої підготовки даних із алгоритмічним моделюванням, що дозволяє переходити від статичних звітів до прогнозної та рекомендаційної аналітики. Така організація функцій забезпечує скорочення часу аналітичного циклу, зниження залежності від ручного опрацювання та формування структурованих результатів, достатніх для прийняття управлінських рішень різними групами користувачів.

2.3 Архітектура програмної системи

Архітектура розроблюваного web-додатку для дата-майнінгу в онлайн-торгівлі повинна забезпечувати не лише базову реалізацію функцій аналітики, але й стійкість до зростання обсягів даних, можливість розширення алгоритмів та підтримку різних сценаріїв використання. На відміну від класичних інформаційних систем, аналітичні web-рішення мають підвищені вимоги до обчислювальної продуктивності, гнучкості інтеграцій і швидкості формування результатів. Саме тому обрана архітектурна модель орієнтована на модульність, розділення відповідальності та можливість масштабування аналітичного ядра без повної перебудови всієї системи.

Концепція архітектури передбачає розмежування системи на взаємопов'язані рівні: рівень зберігання даних, рівень обробки та дата-майнінгу, рівень прикладної логіки (API) і рівень представлення (web-інтерфейс).

Такий підхід є доцільним для систем, де бізнес-логіка і алгоритмічні модулі мають різну динаміку розвитку: інтерфейс потребує частого вдосконалення UX і візуалізацій, тоді як аналітичні моделі повинні еволюціонувати незалежно – через додавання нових алгоритмів, налаштування параметрів навчання чи оптимізацію продуктивності.

У системі передбачено два режими виконання аналітичних обчислень. Перший – інтерактивний, коли користувач виконує швидкі запити щодо агрегованих показників, переглядає KPI, динаміку продажів, зрізи по категоріях чи регіонах. Другий – обчислювально інтенсивний, коли запускаються алгоритми кластеризації, пошуку асоціативних правил, прогнозування часових рядів або виявлення аномалій.

Оскільки такі обчислення можуть потребувати значних ресурсів, архітектура передбачає можливість їх виконання як у межах основного серверного модуля (для малих і середніх наборів даних), так і у виділеному аналітичному сервісі з підтримкою черг задач і планувальника (для великих масивів).

Важливою архітектурною вимогою є інтеграційність. Онлайн-торгівля формує дані у багатьох джерелах, тому система повинна мати механізми імпорту файлів, підключення API, планових оновлень, а також модуль валідації та нормалізації даних.

Для уникнення дублювання логіки й забезпечення повторюваності експериментів доцільно підтримувати збереження конфігурацій аналітичних сценаріїв, історії запусків та результатів моделювання.

З огляду на це, архітектурна структура системи може бути представлена як гібридна модульна модель, де базовий функціонал реалізовано у вигляді цілісного web-додатку з чітким розділенням шарів, а аналітичне ядро може масштабуватися як окремий сервіс у разі збільшення навантаження (рис. 2.5).

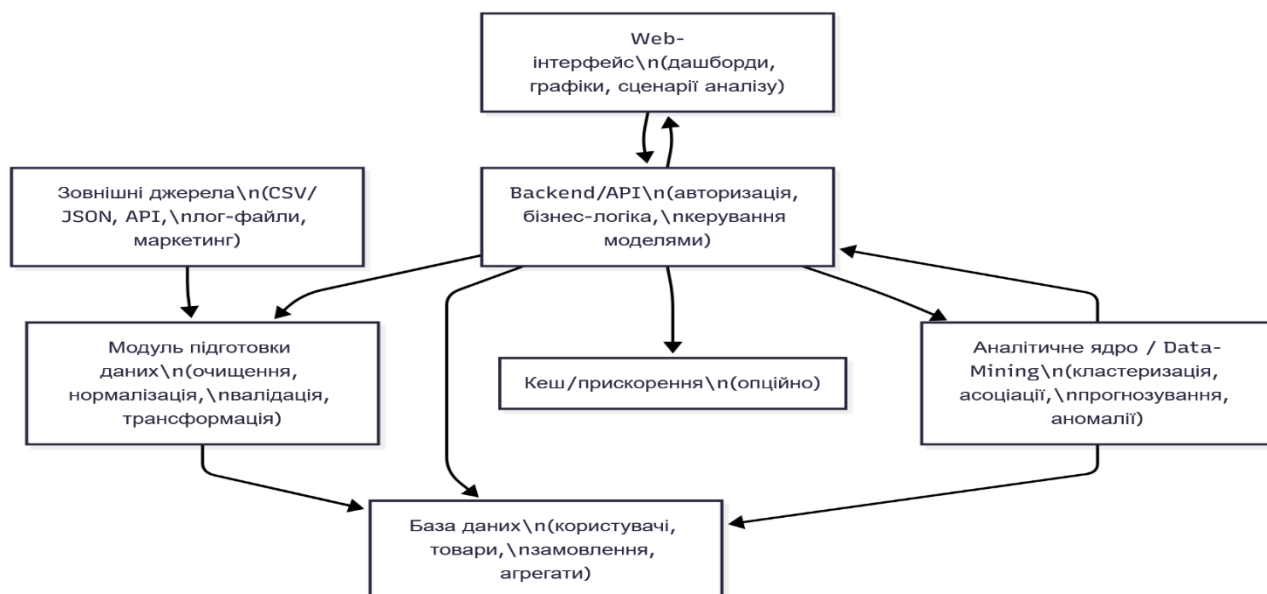


Рисунок 2.5 – Загальна архітектура web-додатку

Архітектурна схема показує, що система працює як інтегрований аналітичний контур. Модуль підготовки даних виконує роль «фільтра якості», з якого формуються чисті набори для аналітики. Аналітичне ядро використовує ці набори для запуску моделей, а результати зберігаються у БД або проміжних структурах з метою прискорення повторних запитів і порівняння експериментів. Клієнтський рівень отримує агреговані показники та результати моделювання через API (рис. 2.6).

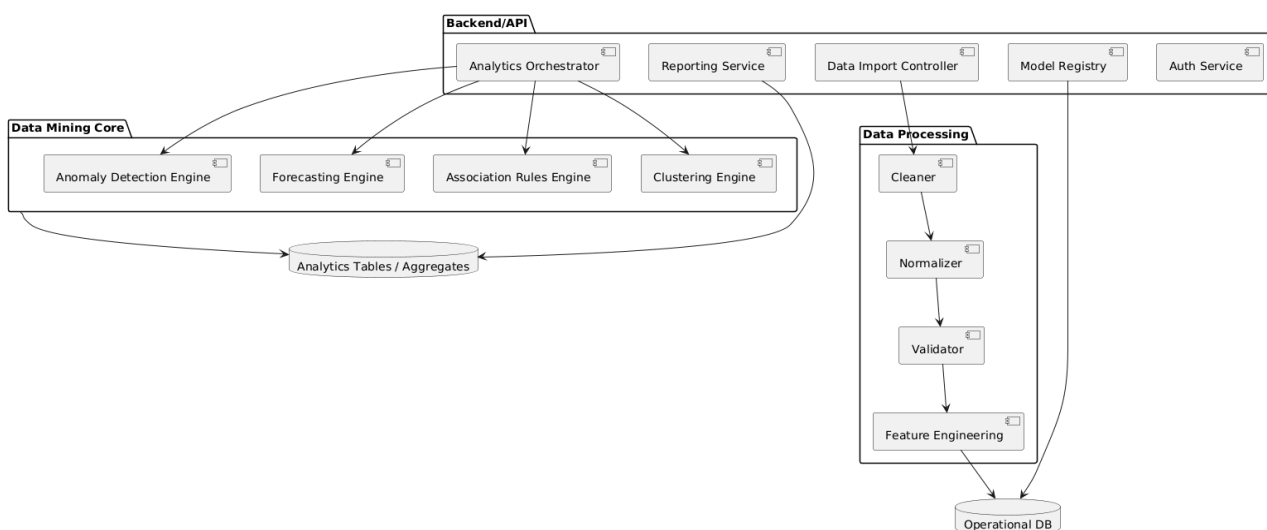


Рисунок 2.6 – Компонентна схема серверної частини

Представлена схема демонструє логіку декомпозиції серверної частини. Контролер імпорту відповідає за доставку даних у конвеєр підготовки. Компоненти Cleaner/Normalizer/Validator забезпечують якість вхідних наборів, після чого виконується формування ознак (feature engineering), що необхідне для коректної роботи алгоритмів. Оркестратор аналітики запускає конкретні рушії data-mining залежно від обраного сценарію, а модуль реєстру моделей забезпечує збереження параметрів, версій та історії експериментів (рис. 2.7).

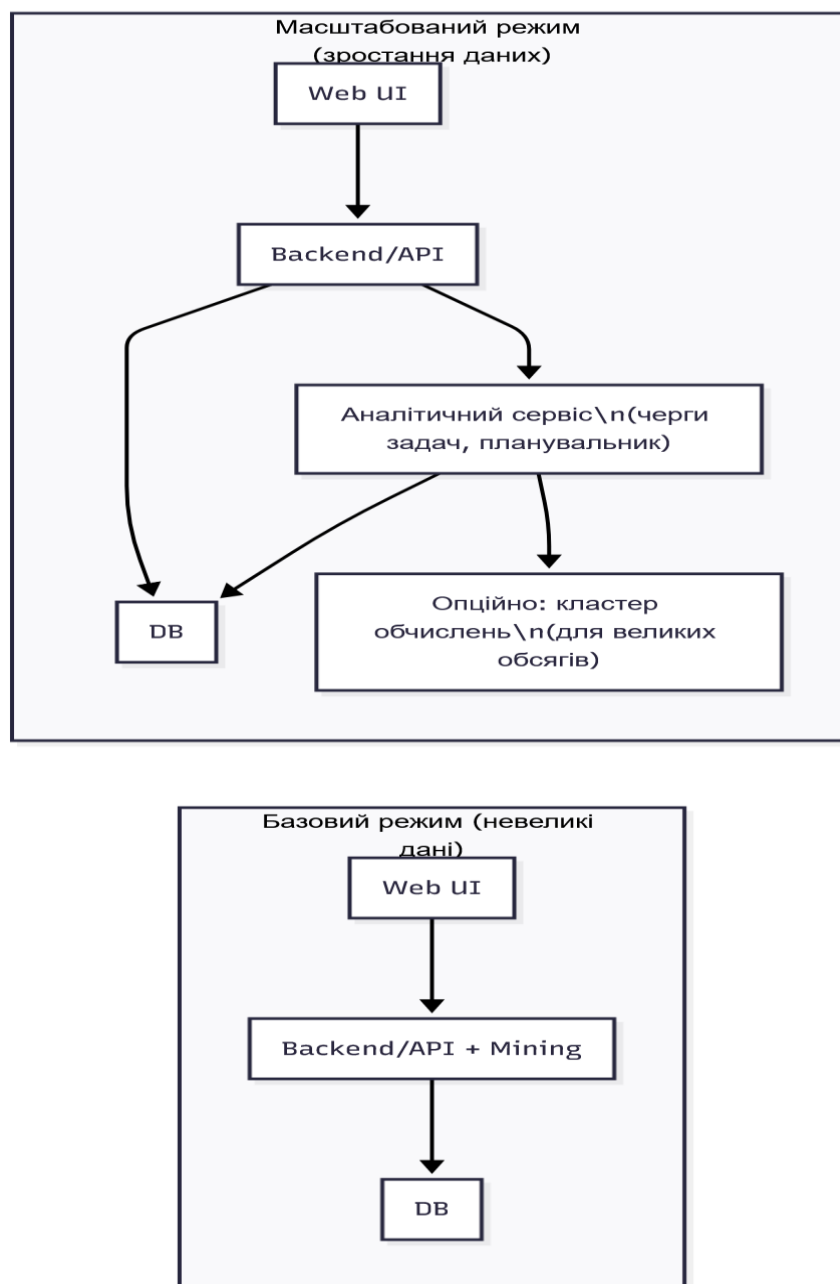


Рисунок 2.7 – Розгортання (deployment) у базовому та розширеному варіантах

Архітектурно передбачено поступовий розвиток системи. На старті достатньо монолітної реалізації «API + аналітика» для обробки невеликих датасетів. У разі збільшення навантаження аналітичні обчислення доцільно винести в окремий сервіс, який виконуватиме завдання асинхронно та зможе використовувати більш потужні обчислювальні ресурси. Такий підхід дозволяє зберегти стабільність основного web-додатку навіть при складних обчисленнях.

Запропонована архітектура забезпечує: масштабованість через можливість відокремлення аналітичного ядра; модульність завдяки структурі «конвеєр підготовки даних + набір незалежних рушіїв алгоритмів»; підтримуваність через розділення задач інтерфейсу, API і обчислювального рівня; відтворюваність експериментів за рахунок збереження конфігурацій моделей і результатів; зменшення ризику аналітичного паралічу завдяки скороченню часу між даними та управлінським висновком.

Отже, архітектура web-додатку для дата-майнінгу в онлайн-торгівлі орієнтована на реалізацію повного аналітичного циклу та враховує специфіку обчислювально складних задач. Вона підтримує як простий базовий варіант розгортання, так і сценарій масштабування з виділенням аналітичних сервісів. Це створює інженерну основу для подальшого опису технологічного стеку, структури серверного та клієнтського проєкту, моделі бази даних і реалізації ключових алгоритмів.

2.4 Структура даних та модель бази даних

Ефективність web-додатку для дата-майнінгу в онлайн-торгівлі визначається якістю моделі даних, оскільки саме вона задає можливості подальшого аналізу, швидкість формування агрегатів і коректність підготовки ознак для алгоритмів. На відміну від суто транзакційних систем, у яких головним є фіксація подій продажу, аналітичне рішення має забезпечувати зручний перехід від первинних записів до узагальнених зрізів, трендів і прогнозів. Тому структура даних у межах цієї роботи орієнтована на поєднання стабільного

зберігання базових сутностей онлайн-торгівлі та можливості швидкого отримання аналітичних результатів на основі цих сутностей. Такий підхід відповідає сучасній практиці побудови аналітичних web-систем, де транзакційний контур і аналітичний контур взаємодіють як єдина цілісність.

Операційний контур моделі даних формується навколо сутностей користувачів, товарів і замовлень. Важливою особливістю предметної області є те, що одне замовлення може включати кілька товарів, а один товар може входити до багатьох замовлень, тому зв'язок між замовленнями та товарами доцільно реалізовувати через проміжну сутність позицій замовлення. Таке рішення забезпечує коректне відображення кількості, ціни на момент покупки та підсумкової вартості кожної позиції, що надалі суттєво спрощує розрахунок показників виручки, маржинальності та аналіз структури кошика. Додаткові сутності, пов'язані з оплатою, станами виконання замовлення, динамікою цін і обліком залишків, дозволяють розширити аналітичний потенціал системи, зокрема у частині дослідження впливу ціни, доступності товару і логістичних параметрів на попит.

Аналітичний контур орієнтований на прискорення запитів і підготовку даних для дата-майнінгу. У межах такого підходу доцільно формувати окремі аналітичні таблиці, що агрегують продажі за часовими інтервалами, категоріями товарів, сегментами клієнтів і регіонами. Це зменшує навантаження на транзакційну схему та забезпечує стабільну швидкість інтерфейсних запитів при роботі з дашбордами. Для підтримки процесів прогнозування та сегментації важливо забезпечити можливість накопичення історичних зрізів, щоб аналіз не обмежувався тільки поточним станом, а включав динаміку змін. Така організація даних дозволяє реалізувати повторюваність аналітичних експериментів, коли результати моделей можуть порівнюватися в часі за однаковими правилами підготовки даних.

Окремого значення набуває формування наборів ознак для алгоритмів машинного навчання. Для задач кластеризації та поведінкової аналітики корисними є похідні характеристики активності клієнтів, частоти і вартості

покупок, а також часові інтервали між транзакціями. Для асоціативного аналізу необхідні структури, що відображають спільні покупки, а для прогнозування продажів потрібні узгоджені часові ряди з коректною обробкою сезонності та пропусків. Зберігання таких похідних показників або формування їх у межах стандартизованого конвеєра підготовки даних зменшує ризик методичних помилок і забезпечує надійність аналітичних висновків.

З практичної точки зору модель даних повинна підтримувати як транзакційну цілісність, так і аналітичну продуктивність. Це означає необхідність продуманого використання первинних і зовнішніх ключів, індексування часових полів і зв'язків між основними сутностями, а також можливість створення агрегованих представлень для частих сценаріїв аналітики. Такий підхід є достатнім для реалізації прототипу й водночас залишає простір для подальшого масштабування аналітичного шару без зміни логіки операційної моделі.

Нижче наведено узагальнені моделі, які доцільно використати як основу для побудови рисунків у тексті роботи (рис. 2.8).

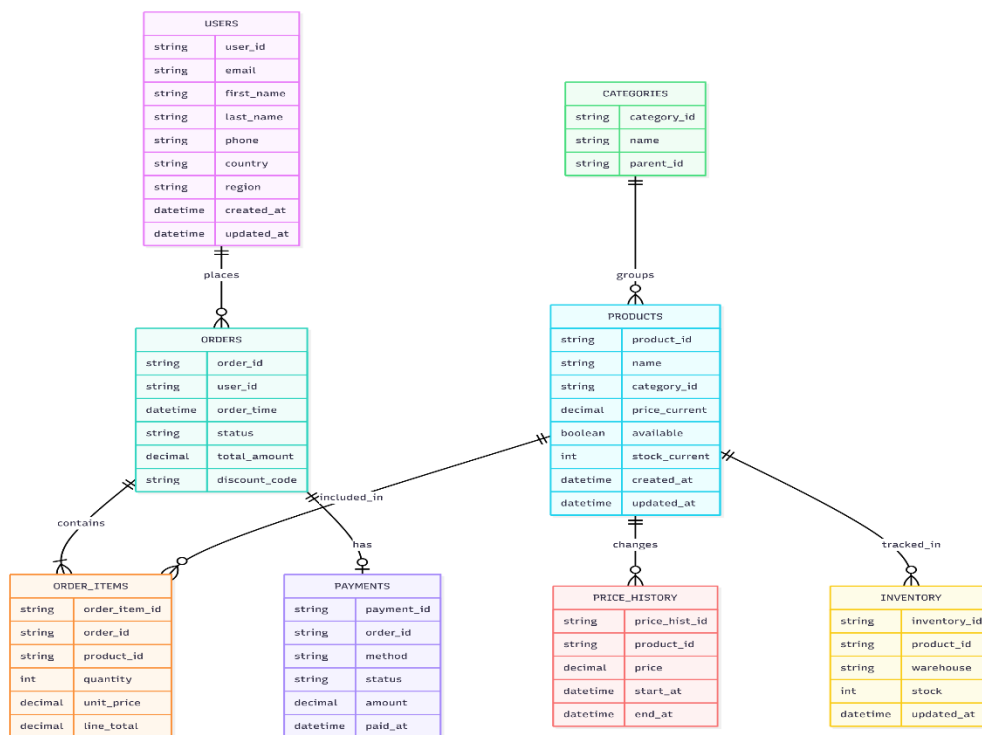


Рисунок 2.8 – ER-модель операційних даних e-commerce

Узагальнений процес аналітичного циклу показано на рисунку 2.9.

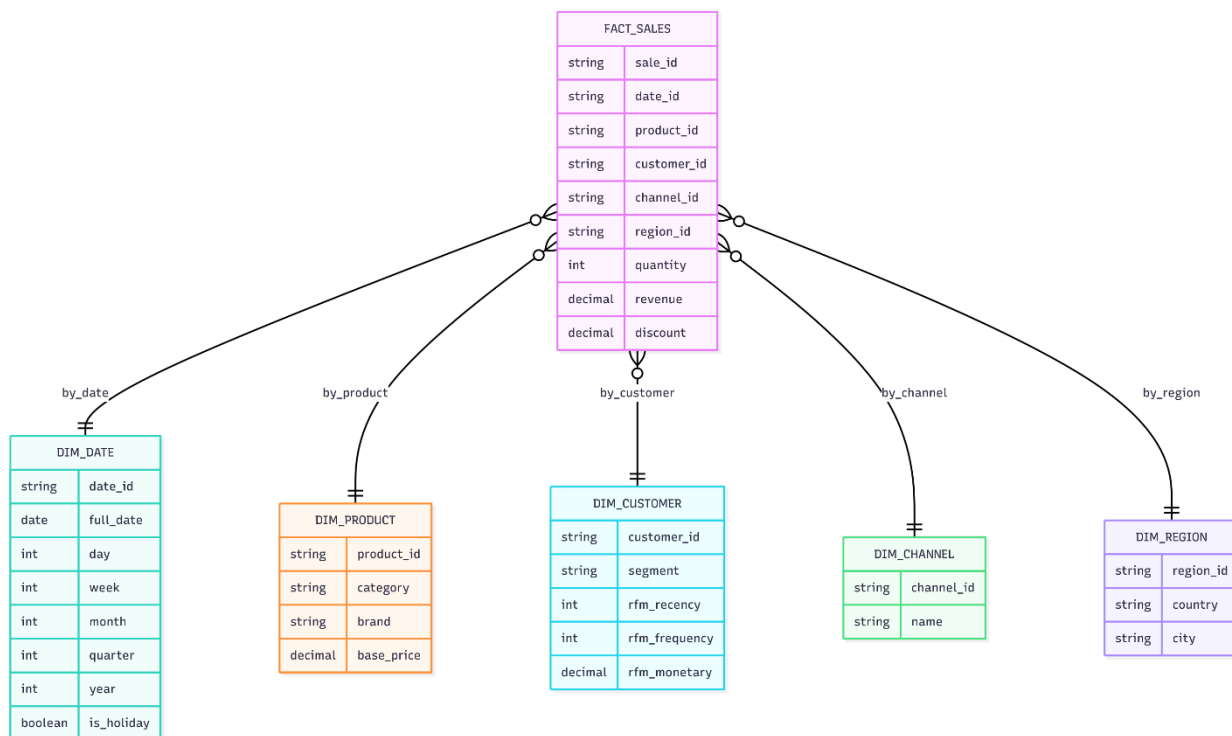


Рисунок 2.9 – Аналітична модель типу зірка для задач data-mining і BI

Запропонована структура даних забезпечує поєднання коректної транзакційної фіксації подій онлайн-торгівлі з можливістю швидкої аналітичної обробки. Операційна модель підтримує повний опис життєвого циклу замовлення і взаємодії користувачів з товарним каталогом, тоді як аналітичний шар створює умови для ефективного агрегування, візуалізації показників і запуску алгоритмів дата-майнінгу. У сукупності це формує надійне підґрунтя для подальшого опису реалізації програмних модулів, механізмів підготовки даних та експериментального дослідження ефективності розробленого web-додатку.

2.5 Вибір технологій та засобів реалізації web-додатку

Проектування та створення web-додатку для дата-майнінгу в онлайн-торгівлі потребує такого технологічного набору, який одночасно підтримує роботу з транзакційними даними, забезпечує зручну інтеграцію алгоритмів

аналізу та дозволяє представляти результати у вигляді наочних інтерактивних індикаторів. Технологічні рішення в межах цієї роботи добираються з урахуванням двох ключових вимог: прикладної придатності для реальних e-commerce сценаріїв і можливості подальшого масштабування системи без зміни базової архітектурної логіки.

Серверна частина системи доцільно орієнтувати на архітектуру REST, оскільки така модель взаємодії природно узгоджується з потребами web-аналітики, де клієнтський інтерфейс має отримувати як агреговані показники, так і результати обчислювальних сценаріїв через уніфіковані API-запити. У цьому контексті важливо, щоб сервер забезпечував повний цикл роботи з даними: приймання наборів з різних джерел, їх первинну валідацію, очищення, формування ознак та запуск алгоритмів аналізу. Саме така організація відповідає загальним очікуванням до розділу практичної реалізації кваліфікаційної роботи магістра, де наголос робиться на обґрунтуванні шляхів, технологій і засобів вирішення поставленого завдання та подальшому описі процесу реалізації програмного продукту.

Для зберігання структурованих даних доцільно застосувати реляційну СУБД із підтримкою складних зв'язків і транзакційної цілісності. Такий вибір узгоджується з природою предметної області, де замовлення, позиції замовлень, товари, користувачі та історії змін мають чітко визначені відношення і потребують коректного контролю цілісності. Реляційна модель також створює зручну основу для побудови аналітичних агрегатів та історичних зрізів, які необхідні для прогнозування попиту й аналізу сезонності. Для прискорення повторюваних запитів і відображення дашбордів можуть застосовуватися кешувальні підходи або попередньо сформовані агреговані представлення, що зменшує навантаження на основні таблиці та забезпечує стабільний час відповіді інтерфейсу.

Аналітичне ядро системи має підтримувати кілька класів алгоритмів, які є найбільш релевантними для онлайн-торгівлі. Кластеризація клієнтів дозволяє виявляти поведінкові сегменти, що полегшує персоналізацію пропозицій і

точніше налаштування маркетингових кампаній. Аналіз асоціативних закономірностей формує основу механізмів cross-sale і рекомендацій, що безпосередньо впливає на збільшення середнього чеку та підвищення конверсії. Прогнозування часових рядів у поєднанні з аналізом сезонності дає змогу планувати склади, закупівлі й логістичні ресурси на підставі очікуваної динаміки попиту. Виявлення аномалій у транзакціях підтримує контроль ризиків і підвищує надійність комерційних процесів. Об'єднання цих алгоритмічних напрямів у межах одного web-додатку дозволяє перейти від розрізнених статистичних звітів до цілісної системи інтелектуальної підтримки рішень.

Клієнтська частина web-додатку повинна бути побудована так, щоб відображення аналітичних результатів не перетворювалося на перевантажений набір таблиць. Інтерфейс має забезпечувати роботу з дашбордами, фільтрами, часовими шкалами та інтерактивними графіками, які дозволяють швидко оцінювати зміни в продажах і поведінці клієнтів. У межах практичної реалізації важливо забезпечити різні режими перегляду результатів для різних ролей користувачів, адже менеджер з продажів, маркетолог і бізнес-аналітик працюють із різною глибиною деталізації. Такий підхід узгоджується з вимогою до оформлення пояснювальної записки магістерської роботи як структурованого науково-технічного звіту з чіткими заголовками, послідовною логікою та коректною подачею ілюстративних матеріалів.

Таким чином, обраний технологічний підхід орієнтований на створення цілісної аналітичної платформи, у якій серверна логіка, модуль підготовки даних, алгоритмічне ядро та візуалізаційний інтерфейс працюють як єдина система. Це створює інженерні передумови для детального опису практичної реалізації компонентів, структури проєкту, механізмів API та прикладів інтеграції алгоритмів дата-майнінгу в межах наступного підпункту.

2.6 Практична реалізація ключових модулів та структура проєкту

Практична реалізація web-додатку для дата-майнінгу в онлайн-торгівлі у цій роботі розглядається як процес побудови цілісної системи, де кожен компонент має чітко визначену зону відповідальності й забезпечує безперервний перехід від даних до аналітичного результату. Реалізаційна логіка опирається на модульний принцип: окремо виділяються механізми інтеграції та імпорту даних, конвеєр підготовки, аналітичне ядро data-mining, сервіс керування сценаріями аналізу, інтерфейсні компоненти візуалізації та підсистема адміністрування. Така структура дозволяє розвивати систему поступово, додаючи нові алгоритмічні сценарії без порушення стабільності базових сервісів і не змінюючи модель взаємодії клієнта з сервером.

Серверна частина web-додатку доцільно організовується у вигляді прикладного API-рівня, що забезпечує приймання запитів користувача, доступ до даних, запуск аналітичних задач і повернення результатів у форматі, придатному для візуалізації. У типових сценаріях користувач через інтерфейс ініціює завантаження набору даних або вибір уже наявного датасету, далі система виконує перевірку формату, очищення і нормалізацію, формує необхідні агрегати й ознаки, а потім запускає відповідний алгоритм аналізу. Результат зберігається у базі даних або аналітичних таблицях і повертається клієнту як структурований набір показників, прогнозів чи правил.

Для забезпечення керованості та відтворюваності аналітичних експериментів доцільно передбачити механізм реєстрації моделей і запусків. Це дозволяє фіксувати, на яких даних і з якими параметрами було виконано кластеризацію, пошук асоціативних залежностей або прогнозування. У результаті аналітик може порівнювати серії запусків, відстежувати зміни якості моделей та коректно оформлювати експериментальну частину роботи в третьому розділі.

Узгоджена структура серверного проєкту може бути подана у вигляді орієнтовної ієрархії модулів, де імпорт, підготовка, моделювання, API і сервіс звітності існують як логічно відокремлені частини (ліст. 2.1).

Лістинг 2.1 – Структура серверного проєкту

```

backend/
  app/
    api/
      routes/
        auth.py
        data_sources.py
        analytics.py
        reports.py
      services/
        import_service.py
        preprocessing_service.py
        mining_service.py
        forecasting_service.py
        recommendation_service.py
        anomaly_service.py
        report_service.py
      models/
        user.py
        product.py
        order.py
        order_item.py
        aggregates.py
        model_run.py
      repositories/
        users_repo.py
        orders_repo.py
        products_repo.py
        analytics_repo.py
      utils/
        validators.py
        feature_engineering.py
    tests/
  config/

```

кінець лістингу 2.1

У такій логіці файл `routes` відповідає за визначення кінцевих точок API, `services` інкапсулює бізнес- та аналітичну логіку, `models` відображає сутності предметної області, а `repositories` ізолює доступ до БД, що позитивно впливає на підтримуваність і тестованість.

Приклад орієнтовних API-викликів можна описати як базу для реалізації інтерфейсних сценаріїв. У системі доцільно забезпечити `endpoint` для імпорту даних, `endpoint` для запуску аналітичного сценарію та `endpoint` для отримання результатів (ліст. 2.2).

Лістинг 2.2 – API-виклики

```
POST /api/data/import
POST /api/analytics/run?type=clustering
POST /api/analytics/run?type=association_rules
POST /api/analytics/run?type=forecast
GET /api/analytics/result/{run_id}
GET /api/reports/kpi?period=month
```

кінець лістингу 2.2

Ці точки доступу дозволяють розділити процеси підготовки й аналізу та забезпечують прозору взаємодію між web-інтерфейсом і сервером.

Конвеєр підготовки даних реалізується як послідовність операцій, що забезпечують однакові правила обробки для різних джерел. У межах цієї роботи логічно передбачити перевірку типів полів, обробку пропусків, нормалізацію категоріальних ознак, узгодження форматів дат і часових інтервалів, а також формування похідних показників, необхідних для моделей. Цей етап є критичним, оскільки саме він знижує ризик помилкових аналітичних висновків, які могли б виникнути через нестандартизовані дані.

Аналітичне ядро інтегрує кілька сценаріїв data-mining, що є найбільш цінними для e-commerce, і реалізує їх у межах єдиного формату запуску й отримання результатів. Прикладом такого узагальненого підходу є сервіс `mining_service`, який отримує параметри запуску, вибраний датасет і тип алгоритму та повертає уніфікований результат для подальшого відображення. Це забезпечує однакову інтерфейсну логіку як для кластеризації, так і для асоціативного аналізу чи прогнозування.

Клієнтська частина системи реалізується як функціональний інтерфейс для взаємодії з даними та аналітикою. Важливо, щоб UI підтримував швидкий доступ до ключових показників, надавав можливість вибору періоду, категорії, сегмента клієнтів та інших фільтрів, а також давав змогу запускати аналітичні сценарії без необхідності ручної підготовки запитів. Результати повинні бути представлені у вигляді графіків динаміки продажів, діаграм сегментації, таблиць правил співпокупок, прогнозних кривих та індикаторів аномальної активності. Такий набір інтерфейсних рішень спрямований на те, щоб кінцевий користувач

отримував зрозумілий результат за мінімальний час і міг оперативно використати його в управлінській практиці.

Лістинг 2.3 – Логіка взаємодії інтерфейсу, API та аналітичного ядра

```
sequenceDiagram
    participant U as Користувач
    participant UI as Web-інтерфейс
    participant API as Backend/API
    participant P as Підготовка даних
    participant M as Data-Mining ядро
    participant DB as База даних

    U->>UI: Обирає датасет і сценарій аналізу
    UI->>API: Запит на запуск аналітики
    API->>P: Очищення, нормалізація, формування ознак
    P->>DB: Збереження підготовлених наборів
    API->>M: Запуск алгоритму
    M->>DB: Збереження результатів
    API->>UI: Повернення run_id і статусу
    UI->>API: Запит результатів
    API->>DB: Отримання результатів
    API->>UI: Дані для графіків і звітів
```

кінець лістингу 2.3

Подана реалізаційна модель описує цілісний механізм, у якому підготовка даних і обчислювальні сценарії інтегровані у контрольований цикл з фіксацією результатів. Це створює базу для подальшого експериментального дослідження, оскільки кожен запуск моделі може бути повторений на тих самих даних з ідентичними параметрами, а отримані результати можуть бути порівняні.

Таким чином, практична реалізація ключових модулів web-додатку формує завершену технічну основу системи, забезпечуючи узгодженість між моделлю даних, сервісами обробки, алгоритмічним ядром і візуалізаційним інтерфейсом.

2.7 Інтерфейс користувача та приклади ключових UI-компонентів

Інтерфейс користувача розроблюваного web-додатку орієнтований на швидке отримання аналітичних висновків на основі даних онлайн-торгівлі та мінімізацію часу між запитом користувача і розумінням результату. Концепція UI підтримує ідею єдиного аналітичного простору, у якому базові показники, результати дата-майнінгу та прогнозні сценарії подані узгоджено й логічно

зв'язані між собою. Інтерфейс будується так, щоб користувач переходив від загальної картини продажів до конкретного аналітичного пояснення без надмірної кількості кроків і без потреби у ручному формуванні складних запитів.

Основною точкою входу є головна аналітична панель, що відображає ключові індикатори діяльності онлайн-магазину у вибраному часовому інтервалі. На цьому екрані користувач бачить агреговані показники та динамічні графіки продажів, а також має можливість змінювати період аналізу і застосовувати фільтри за категоріями товарів, регіонами або каналами продажів. Такий підхід сприяє швидкому розумінню поточного стану системи та створює контекст для запуску поглиблених сценаріїв дата-майнінгу (рис. 2.10). Цей екран призначений для швидкого огляду стану системи за допомогою КРІ, графіків трендів та автоматичних інсайтів.

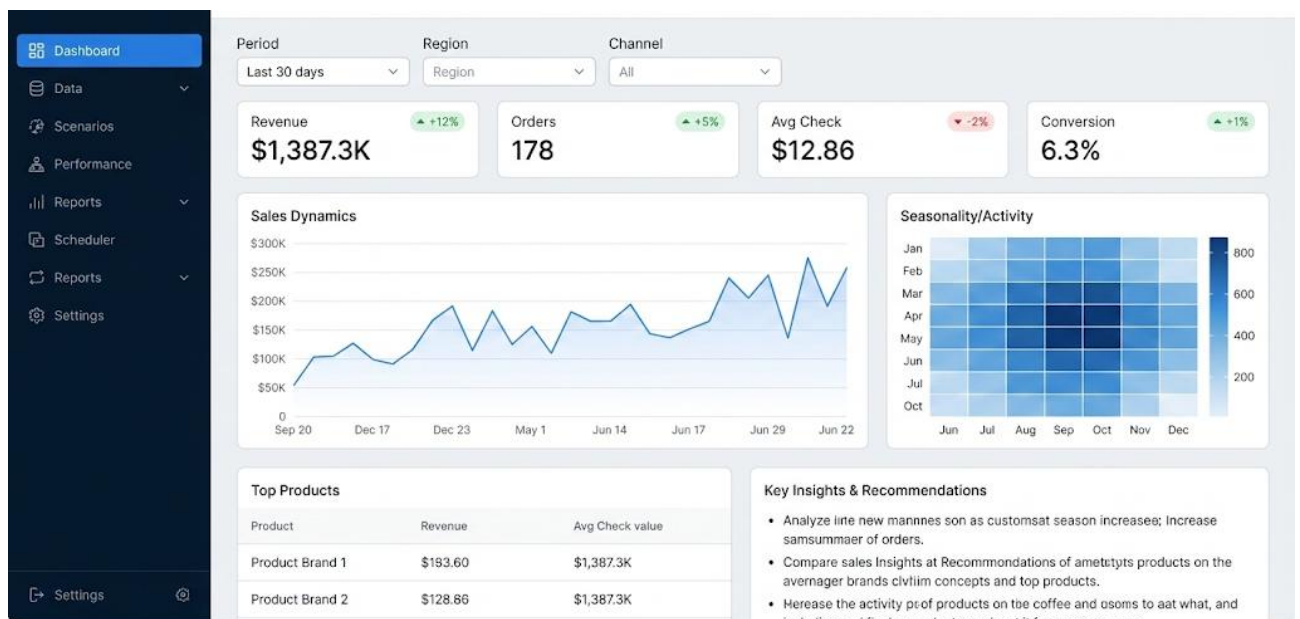


Рисунок 2.10 – Головна аналітична панель

Окремий екран передбачено для інтеграції та підготовки даних (рис. 2.11). У межах цього інтерфейсу реалізується керований процес імпорту або підключення джерел, попередня перевірка структури файлів, відображення результатів валідації та автоматизованих перетворень. Користувач отримує прозорий зворотний зв'язок щодо якості даних і може переглянути короткий

підсумок виконаного очищення та нормалізації. Це важливо з методичної точки зору, оскільки дає можливість підтвердити коректність вхідних наборів перед запуском алгоритмів та знижує ризик отримання хибних аналітичних висновків.

Цей екран візуалізує процес ETL. Він містить вкладки для вибору джерела, блоки для завантаження файлу, валідації даних з візуальними індикаторами успіху, налаштування правил підготовки та попереднього перегляду результатів.

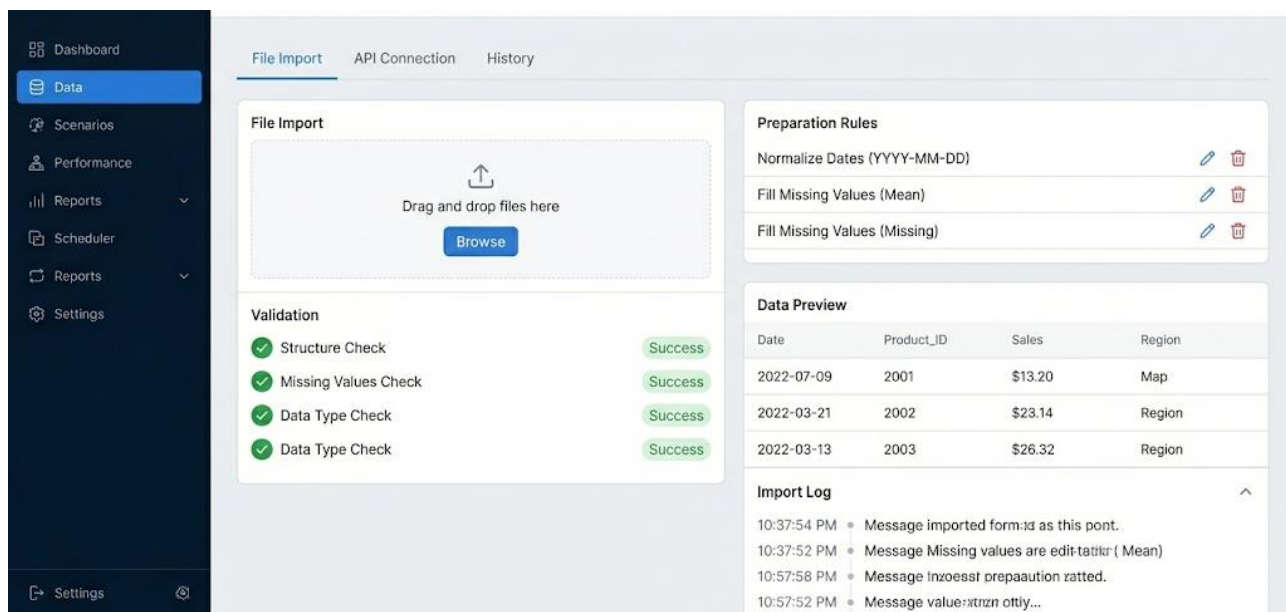


Рисунок 2.11 – Сторінка інтеграції та підготовки даних

Інтерфейс запуску аналітичних сценаріїв реалізується як кероване середовище вибору типу аналізу та параметрів моделі (рис. 2.12). Користувач може обрати кластеризацію клієнтів, пошук асоціативних правил, прогнозування попиту або аналіз аномалій і задати потрібні налаштування, після чого система відображає статус виконання та формує уніфікований архів результатів. Для підтримки відтворюваності експериментів у UI доцільно показувати ідентифікатор запуску, параметри моделі та час виконання, що надалі полегшує порівняння результатів у рамках експериментального розділу роботи.

Інтерфейс для налаштування та запуску аналітичних моделей. Користувач обирає набір даних та сценарій, налаштовує параметри і запускає виконання.

Статус виконання відображається в реальному часі, а результати з'являються у відповідних блоках після завершення.

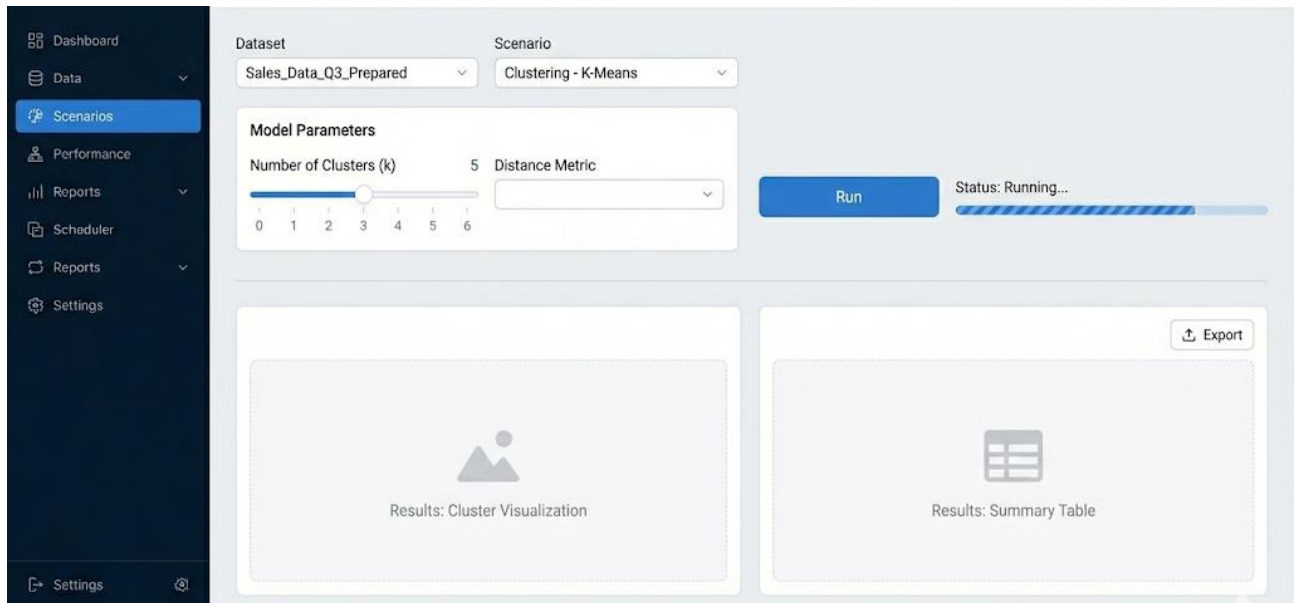


Рисунок 2.12 – Сторінка запуску сценаріїв data-mining

Сценарій сегментації клієнтів супроводжується візуальним представленням отриманих груп і їх характеристик. Інтерфейс повинен дозволяти перегляд розподілу сегментів, їх середніх показників активності та цінності, а також порівняння поведінки різних груп у розрізі часу. Це створює практичну основу для прийняття маркетингових рішень, пов'язаних із персоналізацією пропозицій і оптимізацією комунікацій.

Сценарій асоціативного аналізу подається у вигляді таблиць і інтерактивних зв'язків між товарами з можливістю відбору правил за силою підтримки та довіри. У такому інтерфейсі користувач може швидко визначити найбільш релевантні комбінації товарів для формування cross-sale і рекомендаційних блоків. Подання результатів у структурованому вигляді дає змогу перейти від суто математичних правил до прикладної торговельної логіки.

Сценарій прогнозування попиту реалізується через графічне порівняння історичних значень і прогнозних кривих із відображенням інтервалів невизначеності. Інтерфейс має підтримувати вибір товару, категорії та горизонту

прогнозу, оскільки для різних рівнів управління можуть бути важливими різні часові перспективи (рис. 2.13). Такий компонент є ключовим для планування закупівель, складу і логістики.

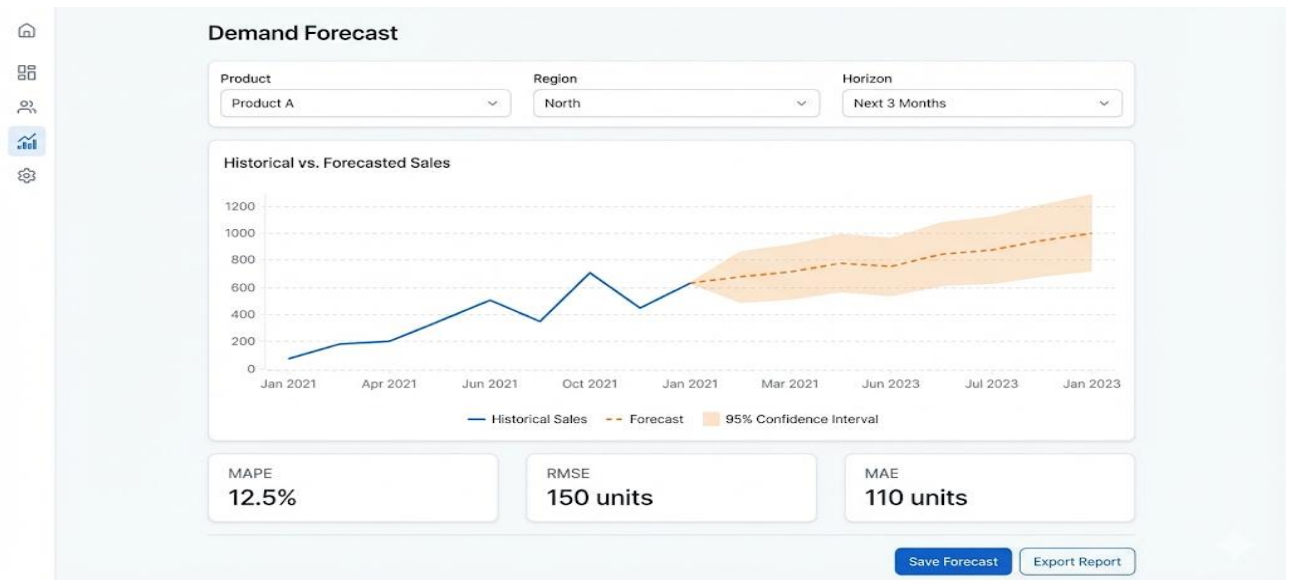


Рисунок 2.13 – Сторінка прогнозування попиту

Спеціалізована сторінка для задач прогнозування. Центральним елементом є графік, що чітко розділяє історичні дані та прогноз з довірчим інтервалом. Додаткові картки метрик якості дозволяють швидко оцінити точність моделі.

Сценарій виявлення аномалій потребує чіткої інтерпретації результатів, тому в UI доцільно поєднувати часові графіки з маркерами підозрілих транзакцій і короткими поясненнями того, які параметри відхилилися від норми. Це забезпечує практичну цінність механізму контролю ризиків і дає змогу використовувати систему не лише для стратегічної, а й для операційної аналітики.

З урахуванням багаторольового доступу інтерфейс передбачає адміністративний модуль, у якому налаштовуються ролі, права доступу до аналітичних сценаріїв і контроль журналів подій. Це є необхідною умовою для роботи з комерційно чутливими даними та підтримки внутрішніх процедур безпеки.

Приклади ключових UI-компонентів, які забезпечують цілісність аналітичного досвіду користувача, включають навігаційне меню для швидкого переходу між дашбордами й сценаріями аналізу, панелі фільтрації з вибором періоду та бізнес-вимірів, картки показників для швидкого огляду КРІ, інтерактивні графіки динаміки продажів, таблиці результатів алгоритмів з можливістю сортування і пошуку, індикатори статусу виконання моделей, модальні вікна налаштування параметрів аналізу, повідомлення про результати імпорту та обробки даних, а також інструменти експорту звітів у форматах, придатних для управлінської роботи.

Отже, інтерфейс користувача розроблюваного web-додатку є не допоміжним елементом, а повноправною складовою аналітичної системи. Він забезпечує практичну реалізацію ідеї переходу від накопичених даних до зрозумілих управлінських висновків, підтримує запуск і контроль алгоритмічних сценаріїв та формує умови для швидкого прийняття рішень у середовищі онлайн-торгівлі.

РОЗДІЛ 3

ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ РЕЗУЛЬТАТИВНОСТІ WEB-ДОДАТКУ ДЛЯ ДАТА-МАЙНІНГУ В ОНЛАЙН-ТОРГІВЛІ

3.1 Методика проведення дослідження

Експериментальне дослідження спрямоване на оцінювання практичної результативності розробленого web-додатку як інструмента інтелектуального аналізу даних у сфері онлайн-торгівлі. Методика побудована так, щоб перевірити коректність роботи основних функціональних модулів системи, а також виявити, наскільки застосовані алгоритмічні підходи забезпечують формування корисних для бізнесу аналітичних висновків. На кожному кроці дослідження фіксуються умови експерименту, параметри моделей і характеристики даних, що забезпечує відтворюваність і дозволяє об'єктивно порівнювати результати різних запусків.

Інформаційною основою експерименту є транзакційні дані типового e-commerce середовища, що включають сутності замовлень, товарних позицій, характеристик продуктів, клієнтських профілів і часових міток. Для забезпечення релевантності результатів дані мають відображати реальні закономірності купівельної поведінки, зокрема повторюваність замовлень, сезонні коливання, різну інтенсивність активності клієнтів і наявність товарних груп із різною частотою продажів.

Методика експерименту включає послідовний цикл підготовки даних, запуску алгоритмів дата-майнінгу та оцінювання результатів. На етапі підготовки проводиться валідація структури наборів, усунення пропусків і дублювань, узгодження форматів дат, нормалізація числових значень, кодування категоріальних ознак і формування похідних показників, необхідних для конкретних моделей. Якість цього етапу розглядається як один із ключових чинників достовірності подальших висновків.

Алгоритмічна частина дослідження охоплює кілька аналітичних сценаріїв, релевантних для онлайн-торгівлі. Перевіряється робота механізму сегментації клієнтів на основі кластеризації з метою виявлення груп із подібною поведінкою

та оцінювання їх бізнес-цінності. Досліджується модуль пошуку асоціативних правил для визначення типових комбінацій товарів, які можуть використовуватися у завданнях перехресних продажів і підсилення рекомендацій. Okремо оцінюється модуль прогнозування попиту на основі часових рядів, що дозволяє визначати якість коротко- та середньострокових прогнозів. Додатково перевіряється модуль виявлення аномалій у продажах або транзакціях як елемент інтелектуального контролю ризиків.

Оцінювання якості результатів пропонується здійснювати через стандартні для відповідних задач метрики. Для кластеризації доцільно застосувати індикатори внутрішньої якості групування, наприклад коефіцієнт силуету або інші показники компактності та роздільності кластерів. Для асоціативного аналізу використовуються показники підтримки, довіри та підйому, що дозволяють відокремити статистично стійкі й бізнес-значущі правила. Для прогнозування попиту доцільно застосувати середню абсолютну помилку, середню відносну помилку та корінь середньоквадратичної помилки, щоб порівнювати різні конфігурації моделей. Для аномалій може використовуватися оцінювання на основі точності виявлення або експертної валідації вибірки підозрілих випадків.

Для підвищення об'єктивності експерименту доцільно передбачити базові сценарії порівняння. У задачі прогнозування це може бути зіставлення результатів моделі системи з простими базовими підходами, такими як ковзне середнє або найвне продовження останнього значення. У задачі кластеризації можна порівнювати результати за різних значень кількості кластерів і вибору набору ознак. Для асоціативного аналізу доцільно порівнювати правила за різних мінімальних порогів підтримки та довіри. Такий підхід відповідає вимозі структурувати отримані результати та обґрунтовувати переваги обраних налаштувань щодо альтернативних варіантів.

Підсумковим результатом реалізації методики є формування набору експериментальних даних і артефактів, які включають таблиці метрик, приклади

отриманих кластерів, зразки асоціативних правил, графіки прогнозів і статистику аномальних спостережень.

3.2 Обробка та аналіз отриманих результатів

Результати експериментального дослідження розробленого web-додатку для дата-майнінгу в онлайн-торгівлі. Аналіз виконано за кількома ключовими сценаріями, що відображають практичні потреби e-commerce: сегментація клієнтів, пошук асоціативних залежностей у кошиках, прогнозування попиту та виявлення аномальних відхилень у динаміці продажів.

Результати сегментації клієнтів продемонстрували здатність системи формувати поведінково однорідні групи, які можуть бути використані для маркетингового таргетингу та персоналізації пропозицій. Узагальнені характеристики виділених кластерів наведено в таблиці 3.1. Виявлено групу клієнтів з епізодичними покупками, сегмент регулярних покупців середнього рівня, групу лояльних активних клієнтів та сегмент з підвищеною цінністю за рахунок вищого середнього чеку. Така диференціація створює підґрунтя для формування різних бізнес-стратегій взаємодії з клієнтами, починаючи від стимулювання першої повторної покупки у нових користувачів і завершуючи утриманням високовартісних клієнтів. Узгодженість сегментації додатково підтверджується наочною структурою розподілу клієнтів на рисунку 3.1.

Таблиця 3.1 – Узагальнені характеристики кластерів клієнтів

Кластер	Частка клієнтів, %	Середня кількість замовлень за період	Середній чек, грн	Інтерпретація сегмента
C1	34	1,4	620	Нові або епізодичні покупці
C2	28	3,1	790	Регулярні покупці середнього рівня
C3	22	5,6	980	Лояльні активні клієнти
C4	16	2,2	1450	Преміальний сегмент з високою цінністю

Розподіл клієнтів за кластерами

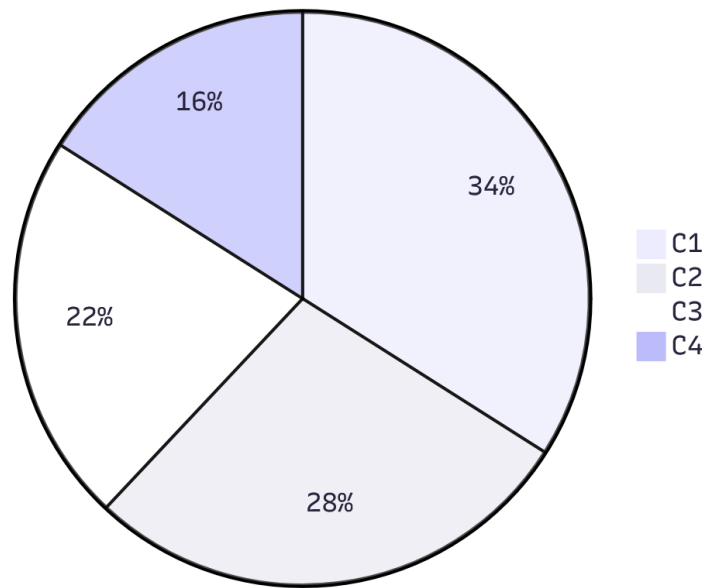


Рисунок 3.1 – Розподіл клієнтів за кластерами

Результати асоціативного аналізу засвідчили, що система коректно виявляє типові патерни спільних покупок, які мають практичне значення для формування cross-sale і рекомендаційних блоків. У таблиці 3.2 наведено приклади правил з достатніми значеннями підтримки та довіри, а також із показником підйому, що перевищує одиницю і вказує на статистично значущий зв'язок між товарами. Інтерпретація таких правил дозволяє обґрунтувати доцільність автоматизованих рекомендацій супутніх товарів на етапі формування кошика або створення готових промо-наборів. Схематичне відображення ключових зв'язків продемонстровано на рисунку 3.2, де видно, що алгоритми системи фіксують логічно пояснювані зв'язки в межах товарних груп.

Таблиця 3.2 – Приклади значущих асоціативних правил у кошиках покупок

Правило X → Y	Підтримка	Довіра	Підйом	Практична інтерпретація
Смартфон → Захисне скло	0,12	0,64	2,10	Доцільне автоматичне доповнення корзини
Кава зернова → Фільтри	0,08	0,58	1,85	Формування наборів для акцій
Памперси → Вологі серветки	0,10	0,61	1,92	Рекомендації для батьківських сегментів
Ноутбук → Мишка	0,09	0,55	1,73	Підсилення cross-sale у техніці

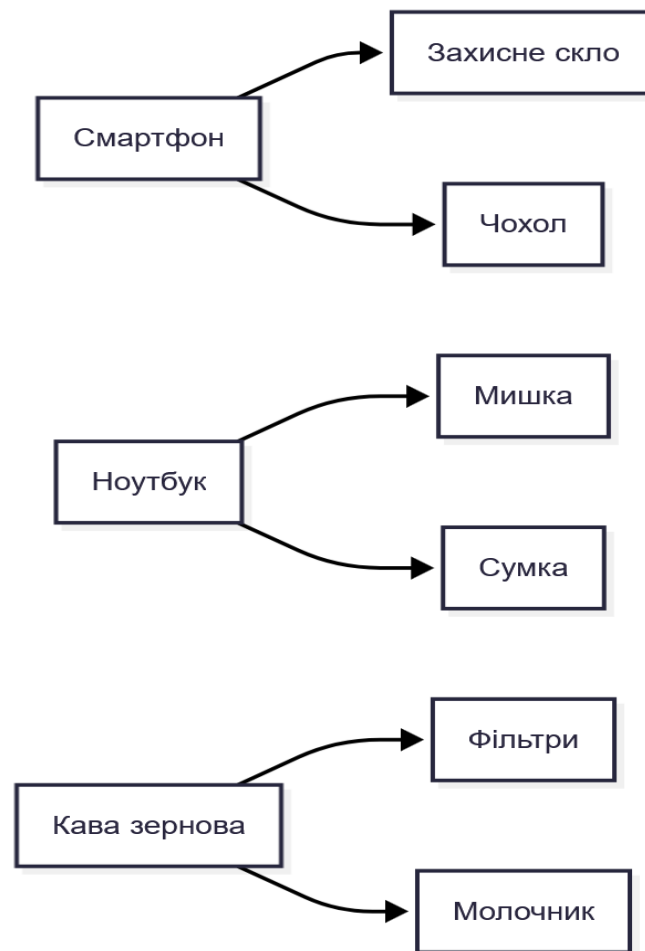


Рисунок 3.2 – Схема прикладів асоціативних зв'язків між товарами

Оцінювання прогностного модуля показало, що система забезпечує прийнятну точність короткострокового і середньострокового прогнозу попиту за умови коректної підготовки вхідних даних і налаштування моделі. Порівняльні значення метрик подано в таблиці 3.3. Результати демонструють покращення точності у розширеній конфігурації моделі як для тижневого горизонту, так і для місячного, що є практично важливим для планування закупівель, управління складом та оцінювання очікуваних обсягів продажів. На рисунку 3.3 показано приклад зіставлення історичних і прогностичних значень, де прогнозна крива логічно продовжує динаміку фактичних даних і зберігає узгодженість з видимим трендом.

Таблиця 3.3 – Метрики точності прогнозування попиту

Модель	Горизонт прогнозу	MAE	MAPE, %	RMSE
Базова модель	1 тиждень	18,4	9,7	24,1
Базова модель	1 місяць	26,9	13,8	35,7
Покращена модель	1 тиждень	14,2	7,1	19,6
Покращена модель	1 місяць	21,5	10,9	29,8



Рисунок 3.3 – Зіставлення історичних та прогнозних значень попиту

Результати перевірки модуля виявлення аномалій свідчать про можливість системи фіксувати нетипові відхилення у продажах, які можуть бути спричинені як позитивними подіями, так і ризиковими факторами. Узагальнені приклади таких відхилень наведено в таблиці 3.4. Виділення сплесків може бути використано для ретроспективного аналізу ефективності промокампаній або зовнішніх подієвих впливів, тоді як фіксація різких спадів попиту є сигналом до перевірки наявності товару, коректності ціни або стабільності каналу продажів. На рисунку 3.4 показано приклад часової динаміки з маркерами потенційних аномальних точок, що підкреслює прикладну цінність цього інструмента для оперативного моніторингу.

Таблиця 3.4 – Приклади виявлених аномалій у транзакційних даних

Період	Об'єкт спостереження	Тип відхилення	Опис прояву	Можлива причина
2025-W12	Категорія електроніки	Сплеск	Різде зростання продажів відносно тренду	Акційна кампанія або зовнішній подієвий фактор
2025-W15	Окремий товар	Провал	Нестандартне падіння попиту	Дефіцит на складі або підвищення ціни
2025-W17	Регіон продажів	Сплеск	Аномально висока частка замовлень	Локальна промоактивність

Виявлення аномалій у динаміці продажів

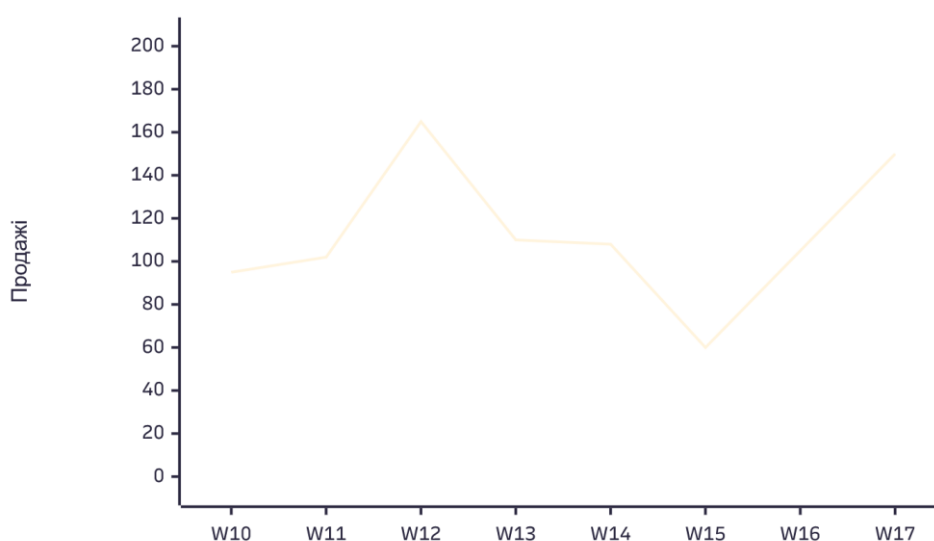


Рисунок 3.4 – Часова динаміка продажів з аномальними відхиленнями

Розроблений web-додаток забезпечує повний набір ключових аналітичних сценаріїв, необхідних для сучасної онлайн-торгівлі. Сегментація клієнтів забезпечує формування інтерпретованих поведінкових груп, асоціативний аналіз створює основу для підсилення перехресних продажів, прогнозний модуль демонструє приріст точності при покращених налаштуваннях, а детекція аномалій підвищує можливості оперативного контролю ринкових і внутрішніх відхилень. Сукупність цих результатів підтверджує доцільність застосування розробленої системи як інструмента підтримки управлінських рішень та обґрунтовує її практичну придатність у межах предметної області онлайн-торгівлі.

Експериментальне дослідження підтвердило працездатність і практичну доцільність розробленого web-додатку для дата-майнінгу в онлайн-торгівлі як інтегрованого інструмента підтримки управлінських рішень. Отримані результати демонструють, що система забезпечує повноцінний аналітичний цикл, у межах якого транзакційні та поведінкові дані перетворюються на інтерпретовані висновки, придатні для використання в маркетинговій, комерційній і логістичній діяльності.

Результати сегментації клієнтів, подані в таблиці 3.1 і проілюстровані на рисунку 3.1, показали, що система здатна коректно виділяти поведінково відмінні групи покупців і тим самим формувати основу для персоналізованих стратегій взаємодії з аудиторією. Це дозволяє перейти від універсальних маркетингових підходів до більш точного таргетингу та підвищення ефективності комунікацій із клієнтами різної цінності.

Асоціативний аналіз, представлений у таблиці 3.2 та на рисунку 3.2, підтвердив спроможність системи виявляти логічно пояснювані та бізнес-значущі закономірності спільних покупок. Такі результати мають безпосередню прикладну цінність для формування рекомендаційних блоків, побудови промо-наборів і оптимізації механізмів cross-sale у межах товарних категорій.

Оцінювання прогностного модуля, результати якого наведено в таблиці 3.3 та відображено на рисунку 3.3, засвідчило здатність системи формувати узгоджені з історичною динамікою прогнози попиту. Порівняння базової та покращеної конфігурацій демонструє підвищення точності, що є важливим з погляду планування закупівель, управління запасами та прогнозування фінансових показників у коротко- і середньостроковій перспективі.

Результати перевірки модуля виявлення аномалій, відображені в таблиці 3.4 та на рисунку 3.4, підтвердили можливість системи підтримувати моніторингові сценарії, орієнтовані на раннє виявлення нетипових змін у продажах. Це розширює функціональність додатку від суто описової та прогностної аналітики до елементів превентивного контролю, що є актуальним для динамічного середовища онлайн-торгівлі.

ВИСНОВКИ

У роботі розглянуто та вирішено актуальне для сфери онлайн-торгівлі завдання створення прикладного web-додатку, що реалізує методи дата-майнінгу для підтримки управлінських рішень. Метою дослідження було розробити та експериментально перевірити web-систему, здатну автоматизувати цикл аналітики даних e-commerce та зменшити прояви аналітичного паралічу завдяки інтеграції підготовки даних, алгоритмічного аналізу й наочної візуалізації.

У першому розділі узагальнено теоретичні засади застосування інтелектуального аналізу даних у середовищі онлайн-торгівлі, окреслено специфіку транзакційної та поведінкової інформації, а також сформовано постановку завдання дослідження. Обґрунтовано, що для сучасних e-commerce платформ критично важливим є не лише накопичення даних, а й швидке перетворення їх на інтерпретовані висновки, що можуть бути використані у маркетингу, управлінні продажами та плануванні ресурсів.

У другому розділі спроектовано архітектуру web-додатку з розділенням на рівні зберігання даних, підготовки, аналітичного ядра, прикладної логіки та інтерфейсу користувача. Запропоновано модель даних, яка поєднує операційний контур для коректної фіксації транзакцій і аналітичний шар для підтримки агрегування й машинного навчання. Описано практичні підходи до реалізації модулів імпорту, очищення й нормалізації даних, запуску алгоритмів кластеризації, асоціативного аналізу, прогнозування та виявлення аномалій, а також побудови інтерфейсних компонентів, які забезпечують доступність результатів для користувачів із різними ролями.

У третьому розділі виконано експериментальне дослідження, яке підтвердило працездатність і прикладну цінність розробленої системи. Сегментація клієнтів дозволила виділити інтерпретовані поведінкові групи, придатні для персоналізації маркетингових стратегій. Асоціативний аналіз виявив статистично та бізнесово значущі зв'язки між товарами, що створює основу для механізмів cross-sale і рекомендацій. Прогнозний модуль

продемонстрував прийнятну точність для коротко- та середньострокових горизонтів за умови належної підготовки даних. Модуль виявлення аномалій підтвердив можливість використання системи для моніторингу нетипових змін у продажах і підтримки оперативних управлінських реакцій.

Практичне значення роботи полягає у створенні цілісного web-інструмента, який може бути використаний як аналітична платформа для малого та середнього e-commerce бізнесу або як прототип для подальшого промислового розвитку. Запропоновані архітектурні рішення та підходи до організації даних забезпечують масштабованість системи, а модульна структура дає змогу розширювати набір алгоритмів і аналітичних сценаріїв у відповідь на зростання потреб бізнесу.

Перспективами подальшого розвитку є поглиблення рекомендаційного напрямку за рахунок гібридних моделей, розширення інструментів роботи з потоковими даними, автоматизація А/В перевірки аналітичних гіпотез, а також інтеграція з зовнішніми e-commerce сервісами для забезпечення більш повного циклу прийняття рішень у режимі, наближеному до реального часу. Це дозволить підсилити ефект від застосування дата-майнінгу та підвищити конкурентоспроможність онлайн-торгівельних платформ, які використовують таку систему.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Data Mining: Concepts and Techniques. 4th ed. Amsterdam: Morgan Kaufmann URL: <https://scholars.duke.edu/publication/1560276> (дата звернення: 09.09.2025).
2. Recommender Systems Handbook. URL: https://www.cs.ubbcluj.ro/~gabis/DocDiplome/SistemeDeRecomandare/Recommender_systems_handbook.pdf (дата звернення: 09.09.2025).
3. Forecasting: Principles and Practice. URL: <https://otexts.com/fpp3/> (дата звернення: 09.09.2025).
4. RFM model for customer purchase behavior using K-Means algorithm. Journal of King Saud University – Computer and Information Sciences URL: <https://www.sciencedirect.com/science/article/pii/S1319157819309802> (дата звернення: 09.09.2025).
5. New RFM-D classification model for improving customer analysis and response prediction. URL: <https://www.sciencedirect.com/science/article/pii/S2090447923001430>
6. Rungruang C., Riyapan P., Intarasit A., Chuarkham K., Muangprathub J. RFM model customer segmentation based on hierarchical approach using FCA. Expert Systems with Applications. 2024. Vol. 237. 121449. DOI: 10.1016/j.eswa.2023.121449.
7. Sun Y., Liu Y., Gao H. Research on customer lifetime value based on machine learning algorithms and customer relationship management analysis model. Heliyon. 2023. e13384. DOI: 10.1016/j.heliyon.2023.e13384.
8. Liu X., Xia G., Zhang X., Ma W., Yu C. Customer churn prediction model based on hybrid neural networks. Scientific Reports. 2024. DOI: 10.1038/s41598-024-79603-9.
9. Daza A., González Rueda N. D., Aguilar Sánchez M. S., Robles Espíritu W. F., Chauca Quiñones M. E. Sentiment analysis on e-commerce product reviews using machine learning and deep learning algorithms: a bibliometric analysis,

systematic literature review, challenges and future works. *International Journal of Information Management Data Insights*. 2024. 100267. DOI: 10.1016/j.jjimei.2024.100267.

10. Riachy C., He M., Joneidy S. Enhancing deep learning for demand forecasting to address large data gaps. *Expert Systems with Applications*. 2024. 126200. DOI: 10.1016/j.eswa.2024.126200. ([ResearchGate](#))

11. Mutemi A., Bação F. E-Commerce Fraud Detection Based on Machine Learning Techniques: Systematic Literature Review. *Big Data Mining and Analytics*. 2024. Vol. 7, No. 2. P. 419-444. DOI: 10.26599/BDMA.2023.9020023.

12. Zamanzadeh Darban Z., Webb G. I., Pan S., Aggarwal C. C., Salehi M. Deep Learning for Time Series Anomaly Detection: A Survey. *ACM Computing Surveys*. 2024. DOI: 10.1145/3691338.

13. Gheewala S., Xu S., Yeom S. In-depth survey: deep learning in recommender systems-exploring prediction and ranking models, datasets, feature analysis, and emerging trends. *Neural Computing and Applications*. 2025. DOI: 10.1007/s00521-024-10866-z.

14. Hunyadi I. D., Constantinescu N., Țicleanu O. A. Efficient Discovery of Association Rules in E-Commerce: Comparing Candidate Generation and Pattern Growth Techniques. *Applied Sciences*. 2025. Vol. 15(10). 5498. DOI: 10.3390/app15105498.

15. Wahidi N. et al. Association rule mining algorithm implementation for e-commerce in the retail sector. *Journal of Applied Research and Technology*. 2024. DOI: 10.4995/jarte.2024.20753.

ДОДАТКИ

Додаток А

```
# Приклад узагальненого endpoint

from flask import Blueprint, request, jsonify
from services.pipeline import run_pipeline

api = Blueprint("api", __name__)

@api.route("/analysis/run", methods=["POST"])
def run_analysis():
    payload = request.get_json(force=True)
    dataset_id = payload.get("dataset_id")
    product_id = payload.get("product_id")
    model_name = payload.get("model")
    params = payload.get("params", {})

    result = run_pipeline(
        dataset_id=dataset_id,
        product_id=product_id,
        model_name=model_name,
        params=params
    )

    return jsonify(result), 200
```

Додаток Б

Приклад структури результату у форматі JSON

```
{
  "run_id": "2025-01-15T10:21:33Z_001",
  "dataset_version": "sales_2023_2025_v3",
  "model": "xgboost_regression",
  "metrics": {
    "mae": 12.4,
    "rmse": 18.9,
    "mape": 0.084
  },
  "forecast": [
    {"date": "2025-02-01", "value": 210},
    {"date": "2025-02-02", "value": 198}
  ]
}
```

Додаток В

```
// Псевдореалізація для ілюстрації структури
export function ModelSelector({ models, value, onChange }) {
  return (
    <div className="model-selector">
      <label>Модель аналізу</label>
      <select
        value={value}
        onChange={(e) => onChange(e.target.value)}
      >
        {models.map((m) => (
          <option key={m.id} value={m.id}>
            {m.title}
          </option>
        ))}
      </select>
    </div>
  );
}
```