

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

**Економічний факультет  
Кафедра економічної кібернетики**

**КВАЛІФІКАЦІЙНА РОБОТА БАКАЛАВРА  
«Підтримка маркетингових рішень в електронній комерції на  
основі технологій штучного інтелекту»**

студентки 4 курсу  
спеціальності 051 «Економіка»  
ОПП «Економічна кібернетика»  
денної форми навчання  
Могиль Сніжани Сергіївни

**Науковий керівник:**  
доктор економічних наук,  
професор  
Чорноус Галина Олександрівна

Засвідчую, що у цій дипломній  
роботі немає запозичень із  
праць інших авторів без  
відповідних посилань

Студент \_\_\_\_\_  
(підпис)

Роботу допущено до захисту перед ЕК  
рішенням кафедри економічної кібернетики  
від 05.06.2024 р., протокол № 15

Завідувач кафедри:  
доктор економічних наук, професор  
Ляшенко Олена Ігорівна

\_\_\_\_\_  
(підпис)

## РЕФЕРАТ

Кваліфікаційна робота бакалавра містить: 90 сторінок, 2 таблиці, 20 рисунків, 46 джерел та 5 додатків.

**Ключові слова:** електронна комерція, електронний маркетинг, персоналізація, штучний інтелект, машинне навчання, аналіз відгуків користувачів, сентимент-аналіз, тематичне моделювання.

**Об'єкт дослідження:** маркетингові рішення в сфері електронної комерції.

**Мета дослідження:** розробка теоретичних і практичних засад для підтримки прийняття маркетингових рішень в електронній комерції за допомогою технологій штучного інтелекту на основі аналізу сучасних методів та інструментів ШІ, оцінення їхньої ефективності в аспектах маркетингової діяльності, а також інтеграції ШІ у маркетингові стратегії компаній для підвищення їхньої конкурентоспроможності.

**Методи дослідження:** наукова абстракція, аналіз, синтез, індукція, дедукція, методи системного аналізу, узагальнення, порівняння, кейс-стаді, систематизація, візуалізація, інтелектуальний аналіз даних, методи програмної реалізації тематичного моделювання та сентимент-аналізу в програмному середовищі Python.

**Теоретична значимість дослідження:** полягає в систематизації теоретичних напрацювань та практик використання технологій ШІ в обґрунтуванні маркетингових стратегій в електронній комерції, розробці концепції аналізу відгуків користувачів з використанням методів машинного навчання, а також її практичної реалізації.

**Практична цінність:** розроблено програму для аналізу відгуків користувачів, яка дозволяє глибше зрозуміти потреби та очікування клієнтів, оцінити їхній досвід та загальне ставлення до компанії, продуктів, обслуговування. На основі цих даних в подальшому можна розробляти та удосконалювати маркетингові стратегії, спрямовані на поліпшення взаємодії з клієнтами та підвищення рівня їхньої лояльності.

## ЗМІСТ

ВСТУП.....	4
РОЗДІЛ 1. ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ЗАСАДИ МОДЕЛЮВАННЯ МАРКЕТИНГОВОЇ СТРАТЕГІЇ В ЕЛЕКТРОННІЙ КОМЕРЦІЇ.....	7
1.1. Електронна комерція та електронний маркетинг: поняття, особливості	7
1.2. Визначення та аналіз ключових компонентів маркетингової стратегії.	16
1.3. Роль технологій штучного інтелекту у підтримці маркетингових рішень на платформах електронної комерції .....	22
РОЗДІЛ 2. МЕТОДИ МАШИННОГО НАВЧАННЯ ТА ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ПІДТРИМКИ МАРКЕТИНГОВИХ РІШЕНЬ В ЕЛЕКТРОННІЙ КОМЕРЦІЇ.....	25
2.1. Концепція впровадження технологій штучного інтелекту для оптимізації маркетингових стратегій.....	25
2.2. Методи штучного інтелекту для підтримки запропонованої концепції	36
2.3. Методи машинного навчання в аналізі поведінки користувачів та оптимізації продажів.....	42
РОЗДІЛ 3. РОЗРОБКА ПРОГРАМНОГО ЗАСОБУ ДЛЯ АНАЛІЗУ ВІДГУКІВ КОРИСТУВАЧІВ ДЛЯ ПІДТРИМКИ ПРИЙНЯТТЯ МАРКЕТИНГОВИХ РІШЕНЬ .....	46
3.1. Розуміння бізнес-процесів та постановка цілей дослідження .....	46
3.2. Характеристика етапів реалізації аналізу відгуків користувачів з використанням розробленого інструменту.....	48
3.3. Результати реалізації запропонованого сценарію на конкретному датасеті.....	57
3.4. Упровадження та напрями удосконалення програмного засобу .....	75
ВИСНОВКИ .....	81
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ .....	85
ДОДАТКИ .....	91

## ВСТУП

**Актуальність дослідження.** З огляду на стрімкий розвиток цифрових технологій, неупинне зростання сфери електронної комерції та постійне збільшення обсягу даних, що генеруються в процесі онлайн-торгівлі, виникає необхідність ефективної обробки цих даних та їх аналізу для підвищення конкурентоспроможності підприємств. Новітні технології штучного інтелекту (ШІ) та машинного навчання можуть бути використані для досягнення бізнес-цілей на кожному етапі реалізації маркетингової стратегії, а саме для підвищення рівня персоналізації, розуміння потреб користувачів, прогнозування попиту, автоматизації сегментації та ціноутворення, підтримки покупців та інших задач.

Використання ШІ дозволяє підвищити ефективність рекламних заходів, автоматизувати маркетингові стратегії, зробивши їх релевантними для кожного користувача, забезпечити вищий рівень залученості клієнтів, лояльності аудиторії і збільшення конверсій. Також технології штучного інтелекту можуть бути інструментом здійснення глибокого аналізу, прогнозування поведінки споживачів для прийняття обґрунтованих маркетингових рішень.

Таким чином, дослідження підтримки маркетингових рішень в електронній комерції на основі технологій ШІ можуть стати надзвичайно важливим для забезпечення стійкого розвитку онлайн-бізнесу в умовах цифрової економіки.

**Метою дослідження** є розробка теоретичних і практичних засад для підтримки прийняття маркетингових рішень в електронній комерції за допомогою технологій штучного інтелекту на основі аналізу сучасних методів та інструментів ШІ, оцінення їхньої ефективності в аспектах маркетингової діяльності, а також інтеграції ШІ у маркетингові стратегії компаній для підвищення їхньої конкурентоспроможності.

Мета дослідження зумовила постановку та розв'язання наступних **завдань**:

- Визначити ключові аспекти функціонування е-комерції та е-маркетингу в цифровому середовищі.

- Проаналізувати компоненти маркетингового міксу в контексті електронної комерції.
- Запропонувати концепцію впровадження технологій ШІ для оптимізації маркетингових стратегій.
- Дослідити вплив технологій штучного інтелекту на різні аспекти маркетингової діяльності, такі як таргетинг, сегментація ринку, рекомендаційні системи, аналіз споживчої поведінки та інші.
- Визначити та оцінити можливості штучного інтелекту та машинного навчання для проведення маркетингових досліджень.
- Створити програмне забезпечення для аналізу користувацьких відгуків з використанням методів сентимент-аналізу та тематичного моделювання.
- Провести оцінку результатів аналізу фідбеку користувачів, надати рекомендації з імплементації програми та визначити перспективи її використання для вдосконалення маркетингових стратегій у електронній комерції.

**Об'єктом** дослідження є маркетингові рішення в сфері електронної комерції.

**Предметом** є технології штучного інтелекту, які використовуються для підтримки та оптимізації маркетингових рішень в контексті електронної комерції.

**Методи дослідження.** Методологічною основою дослідження є наукова абстракція, аналіз, синтез, індукція, дедукція, методи системного аналізу, узагальнення, порівняння, кейс-стаді, систематизація, візуалізація, інтелектуальний аналіз даних, а також методи програмної реалізації тематичного моделювання та сентимент-аналізу в програмному середовищі Python.

**Практичне значення одержаних результатів.** Було розроблено програму для аналізу відгуків користувачів, яка дозволяє глибше зрозуміти потреби та очікування клієнтів, оцінити їхній досвід та загальне ставлення до компанії, продуктів, обслуговування. На основі цих даних в подальшому можна

розробляти та удосконалювати маркетингові стратегії, спрямовані на поліпшення взаємодії з клієнтами та підвищення рівня їхньої лояльності.

**Інформаційну базу дослідження** формують наукові дослідження та розробки вітчизняних і зарубіжних дослідників, електронні статті, публікації з обраної проблематики, веб-ресурси для отримання додаткової інформації та обговорення актуальних питань у галузі.

**Структура й обсяг роботи.** Робота складається зі вступу, трьох розділів, висновків і списку використаних джерел, додатків. Основний зміст роботи викладено на 90 сторінках. Список використаних джерел налічує 46 найменувань. У роботі міститься 2 таблиці, 20 рисунків, 5 додатків.

## РОЗДІЛ 1. ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ЗАСАДИ МОДЕЛЮВАННЯ МАРКЕТИНГОВОЇ СТРАТЕГІЇ В ЕЛЕКТРОННІЙ КОМЕРЦІЇ

### 1.1. Електронна комерція та електронний маркетинг: поняття, особливості

**Електронна комерція** (е-комерція, electronic commerce, e-commerce) — це сфера цифрової економіки, яка охоплює всі торгові та фінансові транзакції, що проводяться з використанням комп'ютерних мереж, а також бізнес-процеси, пов'язані з проведенням цих транзакцій [1, 2].

Синонімічно до терміну "електронна комерція" часто застосовується словосполучення "електронний бізнес", також іноді зустрічається термін "e-tail" (електронна роздрібна торгівля) для опису транзакційних процесів, які формують онлайн-роздрібну торгівлю.

Протягом останніх двадцяти років платформи електронної комерції, такі як Amazon та eBay, зіграли значну роль у зростанні онлайн-роздрібною торгівлі. У 2011 році е-комерція становила 5% від загального обсягу роздрібних продажів, за даними Бюро перепису населення США. За другий квартал 2020 року, після початку пандемії COVID-19, її частка зросла до 16,5% від роздрібних продажів. Після цього вона трохи знизилась до близько 15%, коли фізичні магазини знову відкрилися [3].

Є кілька основних типів бізнес-моделей електронної комерції [3, 4]:

**B2B (Business-to-Business)** - обмін товарами, послугами або інформацією між підприємствами, а не між компаніями та індивідуальними споживачами. Прикладами можуть бути онлайн-платформи та обмінні веб-ресурси, що дозволяють підприємствам знаходити продукти, послуги або інформацію, а також проводити онлайн-транзакції через інтерфейси електронного закупівельного менеджменту (e-procurement).

**B2C (Business-to-Consumer)** - підприємства продають товари, послуги або інформацію безпосередньо споживачам. Зазвичай, існують посередники або сервіси, що відповідають за доставку, обслуговування та інші аспекти взаємодії з клієнтами. Цей термін набув популярності під час зародження інтернет-бізнесу

наприкінці 1990-х років. Сьогодні в Інтернеті існує невичерпна кількість онлайн магазинів та торговельних майданчиків, що пропонують найрізноманітніші товари споживачам. Лідером на ринку B2C є Amazon, який є найбільш визнаним серед цих платформ.

**D2C (Direct-to-Consumer)** - бізнес, що виробляє або надає товари та послуги, продає їх безпосередньо кінцевим споживачам онлайн, без посередників або дистриб'юторів, на відміну від B2C електронної комерції.

**C2C (Consumer-to-Consumer)** - споживачі обмінюються продуктами, послугами та інформацією між собою онлайн (наприклад, онлайн-аукціони чи купівля-продаж через оголошення). Зазвичай такі транзакції проводяться через посередника, який надає онлайн-платформу для укладення угод.

Прикладами C2C-платформ є eBay, Craigslist. Facebook Marketplace та Derop - майданчики для перепродажу модних речей, що також забезпечують транзакції C2C. При цьому, оскільки eBay є бізнесом, цю форму електронної комерції також можна назвати consumer-to-business-to-consumer.

**C2B (Consumer-to-Business)** - споживачі роблять свої продукти та послуги доступними онлайн для того, щоб компанії могли подавати заявки та купувати їх. Це ніби обернена модель до традиційної моделі комерції B2C. Популярним прикладом платформи C2B є ринок, що продає безкоштовні фотографії, зображення, медіа та компоненти дизайну, наприклад, iStock, а також дошка оголошень про вакансії.

**B2A (Business-to-Administration)** - транзакції проводяться онлайн між компаніями та державними органами чи публічною адміністрацією. Чимало урядових сфер потребують різних видів електронних послуг або продуктів, які часто стосуються юридичних документів, реєстрів, соціального захисту, фіскальних даних та зайнятості. Ці послуги можуть надаватися компаніями через електронні канали. Останнім часом ринок послуг B2A значно зростають, оскільки робиться акцент на розвиток можливостей електронного урядування.

**C2A (Consumer-to-Administration)** - транзакції відбуваються онлайн між споживачами та державними органами або публічною адміністрацією. Хоча уряд

рідко купує товари чи послуги від осіб, але вони у свою чергу активно користуються електронними засобами у таких сферах як соціальний захист (надання інформації, проведення платежів), оподаткування (подання декларацій та оплата податків), освіта (надання інформації, дистанційне навчання), медицина (запис на прийом, отримання результатів тестів та оплата медичних послуг).

**Мобільна комерція (m-commerce)** - форма електронної комерції, яка дозволяє користувачам отримувати доступ до онлайн-платформ з метою здійснення покупок без використання комп'ютера, використовуючи лише бездротові мобільні пристрої, такі як смартфони чи планшети [5].

Мобільна комерція включає в себе: електронні покупки, мобільний банкінг, цифрові гаманці, такі як Apple Pay, оплату рахунків і підписку на різні сервіси, такі як Spotify.

З часом доставка контенту через бездротові пристрої стала швидшою, безпечнішою та масштабованою. У зв'язку з цим мобільна комерція стрімко розвивається. За даними Forbes 91% покупок онлайн здійснюється саме через смартфон [6]. Про це не варто забувати власникам електронного бізнесу, маркетологам та аналітикам, які працюють у цій сфері, оскільки цей факт суттєво змінює підхід до організації, розробки, оптимізації онлайн-платформ електронної комерції, роблячи акцент на користувацький досвід у першу чергу.

Галузь роздрібної торгівлі та електронної комерції вважається одним з основних каталізаторів практичного впровадження штучного інтелекту в сферу бізнесу. Згідно з останніми маркетинговими дослідженнями (зокрема, від Precedence Research та DataHorizzon), станом на 2022 рік глобальний ринок рішень штучного інтелекту для електронної комерції оцінювався в діапазоні від \$4-6 мільярдів. Протягом найближчих років очікується його зростання з середньорічним темпом в 14-16%, і до 2032 року воно може збільшитися приблизно в 4 рази, досягнувши обсягу від \$18-22 мільярдів [7].

Алгоритми машинного навчання, нейронні мережі та генеративний штучний інтелект дозволяють онлайн-роздрібним торговцям ефективно

вирішувати різноманітні завдання, серед яких персоналізація маркетингу, покращення клієнтського досвіду та сервісу загалом, оптимізація ланцюжків постачання, прогнозування попиту, динамічне ціноутворення, автоматизація багатьох операцій, забезпечення кібербезпеки та конфіденційності даних.

На відміну від галузей фінансів, медицини, освіти, де широке використання штучного інтелекту все ще може стримуватися певними етичними аспектами та ризиками, електронна комерція може отримати найбільшу вигоду від штучного інтелекту вже зараз. Оскільки приклади успішного застосування ШІ в електронній комерції гігантами галузі (eBay, Alibaba, Walmart, Amazon) вже присутні, роздрібні торговці матимуть намір повторити їх успіх і роль штучного інтелекту в електронній комерції тільки зростатиме.

Однією зі стратегічних функцій будь-якого бізнесу є маркетинг. Саме завдяки йому компанія будує і підтримує якісні відносини з клієнтами, визначає потреби споживачів, їх болі, на основі чого створює найкращий продукт для власного покупця, більше того – ефективно просуває його на ринку, підвищуючи конкурентоспроможність, впізнаваність бренду, збільшуючи свої продажі, залучаючи нові аудиторії. Маркетинг – це інструмент, який забезпечує стабільний ріст, можливості масштабування та допомагає бізнесам досягати стратегічних цілей.

В контексті електронної комерції маркетинг також трансформується, набуває нових рис, форм, функцій. Електронний маркетинг, хоч і не має ще повноцінного свого визначення, вже став невід'ємною частиною бізнесу, успішне функціонування якого унеможлиблюється без е-маркетингу. Цей компонент охоплює не лише вузький профіль маркетингових рішень, а ще й вимагає повноцінної співпраці з іншими відділами: виробництва, закупівель, фінансів, кадрів, контролю якості, розробки продукту – аби мати повне уявлення про внутрішнє та зовнішнє середовища, в яких відбуваються процеси, можливість їх оцінити та покращити.

Можна сказати, що електронний маркетинг є стратегією використання сучасних методів комунікації, яку організація або агенство використовує для

перетворення потенційного ринку у реальний. Мережевий маркетинг дає можливість використовувати Інтернет для надання послуг на етапах перед- та післяпродажу, і звісно безпосередньо обслуговування споживачів під час здійснення покупок. Він включає як пошук нових клієнтів, так і обслуговування вже існуючих в системі, працюючи зі всіма бізнес-операціями. Таким чином, е-маркетинг - це процес, в основі якого використання сучасної маркетингової теорії та Інтернет-технологій, їх функцій для задоволення потреб споживачів з метою реалізації цілей розвитку нового маркетингу та підвищення операційного прибутку бізнесу.

Основні характеристики е-комерції та е-маркетингу подібні. Вони є синтезом базових теорій в галузях економіки, маркетингу, фінансів, комп'ютерних наук, інформаційних систем, менеджменту, дослідження операцій, соціології, лінгвістики тощо; використовують сучасні технології зв'язку, комп'ютерних та бездротових мереж, Інтернету, вебу, баз даних, електронних платежів, безпеки тощо. Зміст бізнес-діяльності електронних торгівлі та маркетингу також схожий, оскільки наявна однакова функція просування (товарів, веб-сайтів, інформації). До того ж для обох понять характерні необмеженість операцій у часі та просторі, електронне середовище розгортання та функціонування, обидва мають функцію "без інвентаря" та можливість відображати SCM (ланцюжок постачання), CRM (управління взаєминами з клієнтами), ERP (система управління ресурсами підприємства), змінювати режим операцій підприємства, сприяти реформуванню моделі бізнесу та управління, збільшувати інтеграцію багатofункціональних послуг.

Е-маркетинг як метод управління підприємством для реалізації товарного обміну є необхідною складовою е-комерції. Можна сказати, що електронний маркетинг є важливим засобом для реалізації електронної комерції, а електронна комерція є етапом розвитку електронного маркетингу.

Електронний маркетинг є основною бізнес-діяльністю в електронній комерції і їх поєднання надасть електронній комерції новий зміст, розширить її область застосування, створить інтегровану та комплексну сучасну нову

маркетинг-модель. Успішний розвиток електронного маркетингу вимагає підтримки з боку платформ електронної комерції та зовнішнього базового середовища, такого як банки, торгові центри, закони та регуляції, політика, що стосується електронної комерції - онлайн-бізнесів інтернет-користувачів, необхідні інтернет-ресурси, а також вимагає підтримки внутрішніх базових умов, таких як концепції та стратегії управління та експлуатації, інформаційні технології тощо.

Філософія управління постійно розвивається, методи менеджменту стають більш передовими, інформаційні мережеві технології (особливо бездротові технології) набувають популярності. Електронна торгівля електронних товарів повинна поступово формувати злиття, де провідна мережа та бездротова мережа існують разом, внутрішнє управління та зовнішні торговельні та маркетингові активності координуються, електронна торгівля та електронний маркетинг поступово з'єднуються.

Інтеграція кожного елементу є ключовим аспектом. При цьому мережа грає важливу роль, в основному, вона з'єднується дротовою мережею між внутрішніми підрозділами (модулями), але в основному використовується бездротова мережа між зовнішніми клієнтами для здійснення торговельних та маркетингових операцій, персоналізованих послуг та комунікацій.

Функції кожної складової визначаються потребами. Модуль електронної комерції повинен забезпечувати можливості персоналізації для клієнтів, бізнес-переговорів, онлайн-оплати, розподілу товарів, логістики та післяпродажного обслуговування. Модуль електронного маркетингу має здійснювати маркетинг у пошукових системах, онлайн-рекламу та інші "дротові" методи, а також застосовувати "бездротові" методи, такі як маркетинг через короткі повідомлення. Різноманітні онлайн-стратегії маркетингу є більш гнучкими та інтегрованими. Внутрішній модуль управління повинен комплексно використовувати ERP, SCM, CRM та інші сучасні методи для ефективного управління підприємством та інтеграції з корпоративною культурою та інфраструктурою.

Типи маркетингу електронної комерції охоплюють широкий спектр каналів, таких як:

- **Електронна пошта.** Перевірений часом і надійний маркетинговий інструмент, що найвищу серед каналів зв'язку середню рентабельність інвестицій (ROI), а саме 36 доларів на кожную витрачену грошову одиницю [9]. Маркетологи використовують пошту задля просування продуктів, спілкування з клієнтами, підвищення їх лояльності, стимулювання покупок. Окрім цього, листування може підвищувати рівень довіри до бренду, будувати його впізнаваність та навіть допомагати клієнтам відчувати себе частиною компанії, свою залученість. Так, & Other Stories, можний бренд, застосовує email-маркетинг для публікації надихаючих історій, нових трендів та стилів, які привертають увагу підписників, підвищують їхній інтерес до компанії, у той час як ексклюзивні пропозиції та акції збільшують продажі бренду. Наразі існує чимало сервісів, таких як таких як Intercom, ActiveCampaign та ConvertKit, за допомогою яких можна легко та швидко налаштувати ефективну рекламну кампанію.

- **Контент-маркетинг.** Маркетологи активно використовують стратегію контентного маркетингу для приваблення та утримання конкретної аудиторії. Це досягається шляхом створення цінного та актуального контенту, репутації бренду у відповідних сферах та приведення якісних лідів до магазинів. Контент може бути у формі блог-постів, білих книг, дослідження кейсів, відео, подкастів, шаблонів для завантажень тощо – і це різноманіття типів потребує конкретної стратегії застосування окремо кожного з них. Після визначення стратегії контентного маркетингу можуть бути корисними у пошуці ідей, виявленні прогалин у конкурентоспроможності, аудиті існуючого контенту та контролі результативності інструмент дослідження тем Semrush, аналіз різниці ключових слів та перевірка SEO сторінки.

- **Пошукова оптимізація (SEO).** SEO-оптимізація включає в себе оптимізацію сторінок товарів, категорій, метатегів, створення високоякісних

зворотних посилань та продукування якісного контенту. Вона сприяє покращенню видимості та позицій сайту у пошукових системах, в контексті електронної торгівлі та маркетингу забезпечує стабільний органічний трафік до магазину. Одним із фундаментальних аспектів будь-якої SEO-стратегії є чітке розуміння пошукових термінів (ключових слів), які використовуються аудиторією конкретного стору для пошуку продуктів, які він пропонує. Наприклад, якщо магазин продає дитячі меблі, і хтось у тій самій локації шукає "купити дитяче ліжечко" у Google, при цьому веб-сайт стору оптимізований під цей запит, він може з'явитися першим у SERP (на сторінці результатів пошуку), збільшуючи ймовірність того, що користувач перейде саме на цей сайт і, можливо, здійснить покупку. Визначити цільові ключові слова та почати впровадження SEO-стратегій, що мають довгострокові переваги для бізнесу, також можна за допомогою онлайн-інструментів, таких як Semrush Keyword Magic Tool.

- **Соціальні медіа.** Маркетинг у соціальних мережах використовує різноманітні платформи, такі як Instagram, Facebook та LinkedIn, для підтримки бренду, продукту або послуги, а також для встановлення взаємовідносин з клієнтами та збільшення трафіку чи обсягів продажів в електронній комерції. Стратегії маркетингу в соціальних мережах охоплюють створення та поширення різноманітного контенту, аналіз конкурентів для розуміння їхніх стратегій, активне привертання та залучення аудиторії, встановлення партнерських відносин з впливовими особистостями, а також постійний аналіз метрик з метою покращення ефективності маркетингових кампаній. Деякі бренди у сфері електронної комерції використовують соціальні мережі для підтримки своєї місії. Наприклад, компанія Tentree, яка спеціалізується на продажу одягу, активно використовує соцмережі для транслювання своєї екологічної місії та позиції, що не лише привертає увагу соціально свідомої цільової аудиторії, а й частково формує її.

- **Платна реклама за кліками (PPC).** Рекламодавці, які використовують метод Pay-Per-Click, оплачують кожен клік на їхні оголошення, які відображаються на пошукових системах або платформах соцмереж. В електронній комерції PPC-оголошення використовуються для збільшення трафіку на веб-сайт, спрямовуючи їх на певні демографічні групи через PPC-платформи. Одна з популярних платформ PPC-реклами є Google Ads. Вона дозволяє створювати платні пошукові оголошення про товари чи послуги на сторінках результатів пошуку. Коли користувачі шукають певне ключове слово, наприклад, "дитяча палатка", вони бачать багато спонсорованих текстових оголошень та торговельних пропозицій, які показують інформацію про товар разом із зображенням прямо на сторінках результатів пошуку за допомогою зображень, що дає клієнтам змогу побачити товар перед тим, як вони клікнуть на оголошення. Google використовує різні фактори для визначення рейтингу оголошення та його розташування в результатах пошуку. Але щоб отримати верхню позицію в оплачених пошукових результатах компанія має зробити найбільшу ставку на певні ключові слова, вказуючи максимальну суму, яку вона готова заплатити за кожен клік на її оголошення. Інструменти, такі як Google for Retail, можуть використовуватися для розміщення списків товарів та створення PPC-оголошення. За допомогою облікового запису Merchant Center, можна безкоштовно розмістити бізнес на Google, отримати персоналізовані відомості та збільшити продажі. Також можна посилатися на платформи електронної комерції, такі як Shopify. Кампанії Performance Max на основі цілей дозволяють відстежувати цілі, пов'язані з конверсією, та досягати їх. Кожен з цих інструментів допомагає створювати та реалізовувати успішну електронну маркетинг-кампанію.

- **Локальний маркетинг** (орієнтований на місцевий ринок). Онлайн-магазини можуть також мати як таргет регіонально обмежену аудиторію, використовуючи інструменти локального маркетингу. Наприклад, роздрібний продавець одягу може використовувати локальний маркетинг для просування

спеціальної розпродажі зимових пальто серед клієнтів у холодних регіонах. Так, The North Face використовував спрямований на погоду таргетинг у своїй рекламі для жіночих походів, показуючи різні версії оголошень в залежності від погодних умов у регіоні того, хто переглядає. Користувачі у холодних регіонах бачили рекламу светра чи худі, тоді як користувачі у теплих - рекламу майки. Локальний маркетинг в електронній комерції використовує різні стратегії, такі як оптимізація веб-сайтів, інструменти типу Google Business Profile, місцеві цитати, зворотні посилання та позитивні відгуки. Локальний алгоритм пошуку Google ранжує бізнеси за їхньою релевантністю, відстанню та популярністю у результатах пошуку.

## 1.2. Визначення та аналіз ключових компонентів маркетингової стратегії

Оскільки електронний бізнес нерідко стикається зі складнощами під час застосування традиційних маркетингових інструментів у віртуальному середовищі, розгляд цих компонентів стає надзвичайно важливим. У даному підрозділі буде проведено детальний розгляд цих ключових компонентів, їх визначення та аналіз у контексті змін, що відбуваються в електронній комерції, зокрема з урахуванням тенденцій електронного бізнесу та інтернет-технологій.

У контексті трансформації маркетингової стратегії на основі сучасних інноваційних технологій було розглянуто концепцію маркетингового міксу для електронної комерції [10].

**Product.** Компонент "Продукт" частково відповідає традиційному концепту цього елементу. Загалом в електронній комерції виділяють декілька типів онлайн товарів та послуг:

1. Інформаційні продукти – це товари, які можна попередньо продемонструвати перед покупкою, такі як аудіо- та відеотовари.

2. Товари, перед придбанням яких зазвичай відбувається збір значної кількості інформації. Це можуть бути автомобілі, комп'ютери, побутова техніка, а також товари з нижчої цінової категорії, покупці яких передують вивчення анотацій та відгуків, наприклад, книги.

3. Товари та послуги, що надаються через Інтернет-канали. Ця категорія включає програмне забезпечення, прокат автомобілів, бронювання готелів, продаж квитків на залізничний чи повітряний транспорт, туристичні послуги тощо.

4. Унікальні товари - товари, які реалізуються через електронні аукціони або спеціалізовані колекційні магазини, наприклад, рідкісні чи колекційні екземпляри, антикваріат, предмети мистецтва, для яких характерна обмежена доступність та висока вартість.

5. Товари остаточного споживання, для яких часто характерна висока цінова чутливість. Це може бути одяг, іграшки, побутові товари та інші предмети широкого вжитку - попит на ці позиції часто залежить від їх вартості, що вимагає особливих підходів у ціновій політиці та рекламних стратегіях.

6. Продукти харчування. В контексті електронної торгівлі важливими аспектами для цієї категорії є логістика доставки, контроль за свіжістю та відповідністю харчових продуктів стандартам безпеки.

Специфіка кожної групи товарів визначає відповідний комплекс маркетингових заходів, однак можна виокремити загальні стратегії розвитку маркетингу в електронному бізнесі.

Однією з характерних можливостей в електронній торгівлі є формування товарних груп електронних продуктів і послуг. Ці товари не мають матеріальної форми, їх суттєві характеристики полягають в реалізації "економічної форми" інформації. У цифровому середовищі інформаційні послуги є невід'ємною частиною торговельних послуг, що особливо актуально в онлайн-взаємодії між продавцем та покупцем.

Унікальним явищем розвитку електронної комерції є активний розподіл товарів, які виступають "віртуальними заміниками". Ці товари - статуси, вподобання, лайки, подарунки, токени – необхідні для залучення уваги користувача та ефективної взаємодії в мережі. Віртуальне середовище також сприяє посиленню феноменів видимого споживання, що вказує на важливість

враховувати соціальні потреби споживачів, такі як залучення, повага та самовираження.

Таким чином, повертаючись до концепції "Продукту" у традиційному розумінні, варто зазначити, що в електронній комерції змінюється споживча поведінка як щодо традиційних товарів, які мають фізичну форму, так і щодо послуг, адже у цифровому середовищі, лишаючись "один на один" з інтерфейсом програмного забезпечення або веб-сайтом, користувач забезпечується анонімністю, дистанціюванням і психологічною, емоційною безпекою (за потреби завжди можна залишити середовище, наприклад, вимкнувши комп'ютер), ізолюваністю від прямих переконливих факторів матеріального середовища.

**Price.** Формування цін на товари в електронній комерції на перший погляд здійснюється за стандартними принципами. Стратегія ціноутворення базується на трьох основних складових: витрати, цінність для споживача та показники конкурентів. Однак досвід світових лідерів у галузі електронної комерції свідчить про виникнення абсолютно нових методів гнучкого та індивідуального ціноутворення.

Згідно з даними Internet Retailer, понад 9715 змін цін зафіксовано на Amazon.com у сезон святкових розпродажів з 24 листопада по 14 грудня [11]. Amazon встановив рекордне значення показника, перевершивши частоту змін цін своїх конкурентів, а саме Best Buy Co., Target Corp., Wal-Mart Stores Inc. та Toys 'R' Us Inc. Віце-президент з маркетингу Amazon стверджує, що компанія може змінювати ціни на товари аж до 10 разів на день, при цьому найбільш гнучкий вплив простежується на вартість побутової техніки та деякого одягу. Приблизно для 20% всіх товарів в онлайн-торгівлі ціни змінюються щодня, для найбільш популярних пропозицій можуть оновлюються кожні кілька хвилин. Ця динаміка в е-комерції залежить переважно від ринкових умов, проте стратегії ціноутворення можуть мати й індивідуальний характер для кожного користувача, що можливо завдяки застосуванню технологій BigData, адже аналізуючи великі обсяги даних про споживачів, їхні покупки, попередні

перегляди, поведінку в цілому, компанії можуть створювати персоналізовані пропозиції та цінові стратегії, адаптувати ціни з метою максимізації прибутку, залучення нових клієнтів або збереження існуючих.

**Place.** Місцем продажу в електронній комерції може бути веб-сайт або комерційний обліковий запис у соціальній мережі, при цьому останній, окрім того, що є основною точкою взаємодії з цільовою аудиторією, також виконує функції просування. У сучасному бізнесі місцем продажу може бути також спільна платформа для групи індивідуальних торговців, що продають товари за фіксованими цінами (aliexpress.com) або через аукціон (ebay.com). Основною характеристикою місця в цьому випадку є максимальна доступність каналів продажу як для покупців, так і для продавців.

**Promotion.** В електронному середовищі просування товарів характеризується найбільш різноманітним виявом можливостей та інструментів, які дозволяють активно включати споживачів у комунікаційний процес і стимулювати їх до конкретних дій (затвердження, реєстрація, завантаження, покупка, рекомендація тощо). У випадку успішної маркетингової кампанії інформація про компанію, продукт чи послугу стає природньо віральною, що забезпечує охоплення, яке за масштабами прирівнюється до медіа-кампаній, та призводить до покращення показників продажу.

Можна виділити наступні інструменти просування для магазинів електронної комерції (усі вони розробляються та розгортаються в мережі та включають різноманітні техніки):

1. Реклама. В електронній комерції охоплює банерну, контекстну, мобільну, а також у вигляді контенту та у соцмережах.

2. Інтелектуальні боти. У продажах типу “person-to-person” функції продавця або консультанта поступово делегуються інтелектуальним ботам, які дозволяють купувати товари, отримуючи персоналізовані рекомендації чи відповіді на питання про продукт/послугу; банківські боти можуть проконсультувати щодо фінансових послуг; боти, які виконують контролюючу функцію, можуть

сповістити про певні події чи зміни в системі (наприклад, затримка рейсу, потреба в обслуговуванні автомобіля тощо).

3. Стимулювання продажів (sales promotion). Включає такі методи, як “call to action” (заклик до дії) в соцмережах, e-mail маркетинг, веб-конференції та вебінари. Окремо варто наголосити на соціальних мережах, які активно використовуються для підвищення лояльності клієнтів, стимулювання їх залученості, поширення інформації серед контактних аудиторій.

4. Формування зв'язків з громадськістю. Для формування зв'язків з громадськістю (public relations), створення та підтримки іміджу бренду використовується маркетинг у соціальних мережах, контентний та реферальний маркетинг. Згідно з дослідженням Nielsen (2015), у всіх вікових групах (покоління Z, мілленіали, покоління X, бумери, мовчазне покоління) найбільшу довіру викликає реклама у вигляді рекомендацій, тоді як онлайн-середовище вважається сприятливим до прямої комунікації покупців як з брендом, так і з іншими споживачами [12].

5. Пошуковий маркетинг (search engine marketing). Є інструментом, який використовується для підтримки просування бренду, забезпечуючи індексацію пошуковою системою, вигідне розташування в пошукових результатах та збільшення органічного трафіку.

**People, Process, Physical Evidence.** Ці компоненти зазвичай фігурують у маркетинговій стратегії організацій, які надають послуги, наприклад, у сфері обслуговування, туризму, готельного та ресторанного бізнесу. Такі компанії часто оперують у сфері, де взаємодія з клієнтами відбувається безпосередньо, тоді велика увага приділяється впливу персоналу, процесів обслуговування та фізичного середовища на споживачів.

У цифровому середовищі усе це замінюється технологією, що реалізована у віртуальному представленні компанії (наприклад, на сайті онлайн-магазину), де людський фактор у процесах обслуговування клієнта мінімальний. З одного боку, це може позитивно впливати на якість сервісу, оскільки технології базуються на алгоритмах, які не характеризуються емоційністю, неухважністю

тощо. З іншого боку, технології обмежені в творчості та не мають емпатії, харизми та інших важливих людських якостей, які викликають у споживача прихильність, довіру, формують його лояльність.

Важливо наголосити, що хоча втручання людини під час безпосереднього обслуговування клієнтів, які здійснюють покупку, і є мінімальним, проте участь людей як керівників віртуальних бізнес-процесів є вирішальною.

Традиційна торгівля характеризується прямим контактом споживача з персоналом компанії. Окрім персоналу, на прийняття рішення покупця можуть також впливати й інші споживачі, наприклад, перебуваючи у фізичному магазині під час вибору товару. Споживач оцінює їх зовнішній вигляд, поведінку (виражену ступенем задоволення), їхні відгуки та думки. Проте електронна комерція, як вже зазначалося раніше, характеризується відсутністю прямого спілкування споживача з персоналом магазину (принаймні до моменту покупки), оскільки комунікація відбувається за допомогою автоматизованої інформаційної системи (AIS), через інтерфейс веб-сайту магазину або ETF (для B2B).

Також на процес придбання товарів можуть впливати сторонні учасники, і їх вплив є набагато сильнішим, іноді вирішальним, хоча і здається непрямим. Форма такого впливу - інформаційні послуги, що надаються безпосередньо торговими компаніями (ETF, інтернет-магазини), незалежними спільнотами (неприбуткові організації захисту прав споживачів в Інтернеті, спеціалізовані портали, які пропонують оцінки продавців, огляди покупок, групи в соціальних мережах тощо).

Фізичні докази (фізичне оточення) в електронній комерції можна поділити на дві складові: традиційні фізичні і віртуальні. Фізичне середовище представлене пунктами доставки, офлайн-магазинами (Points of Sales) і офісами компанії, які, окрім першочергових функцій, також можуть виконувати функцію візитівки бізнес-профілю, офіційного представництва компанії, підсилюючи її присутність на ринку та додатково викликаючи довіру у потенційних та вже існуючих клієнтів. До віртуальних середовищ належать: веб-сайти та сторінки у соціальних мережах, мобільні додатки, які безпосередньо використовуються для

продажу товарів, здійснення онлайн-платежів, реклами, залучення клієнтів, їх обслуговування, а також збору даних і проведення аудиту, аналізу.

### **1.3. Роль технологій штучного інтелекту у підтримці маркетингових рішень на платформах електронної комерції**

У сучасному цифровому світі штучний інтелект відіграє ключову роль у трансформації маркетингових практик на електронних комерційних платформах. Із стрімким розвитком технологій, електронна комерція стає все більш конкурентноспроможною, а ШІ показує себе як надзвичайно потужний інструмент для оптимізації стратегій маркетингу. Цей підрозділ досліджує важливість технологій ШІ та їхню роль у процесах прийняття маркетингових рішень на платформах е-комерції.

Наразі застосування штучного інтелекту в галузі електронної комерції головним чином проявляється в наступних аспектах:

**ШІ-асистент.** ШІ-асистент (чат-бот), головна функція якого - автоматично відповідати на запитання клієнтів, реагувати на прості голосові команди (можливо навіть зображення) та надавати рекомендації щодо товарів за допомогою системи обробки природної мови. Чат-боти можуть допомагати споживачам знайти відповідні товари, перевіряти наявність товарів, порівнювати різні товари і навіть допомагати споживачам здійснити оплату. Якщо виникають скарги або питання, чат-бот також може допомогти клієнтам зв'язатися з підтримкою магазину. Так, у березні 2017 року Alibaba запустила Shop Xiaomi – чат-бот для продавців на Taobao.

Після авторизації та налаштування від адміна чат-боти можуть замінити деякі служби підтримки клієнтів, ефективно знизити витрати на робочу силу, оптимізувати користувацький досвід, покращити якість обслуговування та максимізувати відновлення нічного трафіку, а також допомогти службі підтримки клієнтів вирішувати питання повторних консультацій.

**Системи рекомендацій.** Система рекомендацій є повноцінною рекомендаційною системою, побудованою на базі алгоритмів машинного

навчання. Використання штучного інтелекту дозволяє здійснювати глибоке навчання, статистичне програмування, прогнозування та аналіз поведінки клієнтів на основі великих обсягів даних з метою передбачення, які продукти відповідатимуть інтересам клієнтів. Починаючи з аналізу останніх пошукових запитів потенційних клієнтів, алгоритми машинного навчання в системі рекомендацій можуть відбирати ключові характеристики шуканого продукту та на їх основі генерувати рекомендації, які будуть найбільш релевантні користувачу.

Імплементуючи алгоритми зменшення розмірності, можна вдосконалити рекомендаційну систему, зробити її більш точною і ефективною у формуванні рекомендацій, в також враховувати ширший контекст взаємодії користувачів з системою. Широке застосування цих інструментів на великих електронних комерційних платформах, таких як Amazon, Alibaba, Taobao та JD.com, свідчить про їхню ефективність у визначенні цільової аудиторії та залученні користувачів до покупок через рекомендації.

**Оптимальне ціноутворення.** Корекція цін на короткострокові та довгострокові періоди є доволі непростою задачею навіть для малих бізнесів в сфері е-комерції через гнучкість і динамічність середовищ функціонування. Застосування технологій ШІ та передових алгоритмів глибокого машинного навчання дозволяють практично вирішити проблему автоматизації ціноутворення для різних товарів за різних умов, оскільки здатні швидко опрацьовувати великі обсяги даних, неперервно оцінювати динаміку ринку та видавати результати, спираючись на рейтинг продуктів, вартість доставки, якість обслуговування та інші фактори.

**Позиціонування бренду.** За даними Curiosity at Work, 52% споживачів (серед зумерів ця частка досягає 66%) зацікавлені новітніми технологіями у сфері штучного інтелекту для роботи з цифровими продуктами та їхньою реалізацією, тому бізнес, який активно почне впроваджувати подібні інновації точно приверне увагу користувачів [13]. До того ж з інструментами ШІ можна легко автоматизувати частину маркетингових процесів, використовуючи при

цьому data-driven підхід, та здобувши перевагу серед конкурентів від персоналізованих маркетингових та рекламних стратегій, які повністю відповідають потребам аудиторії.

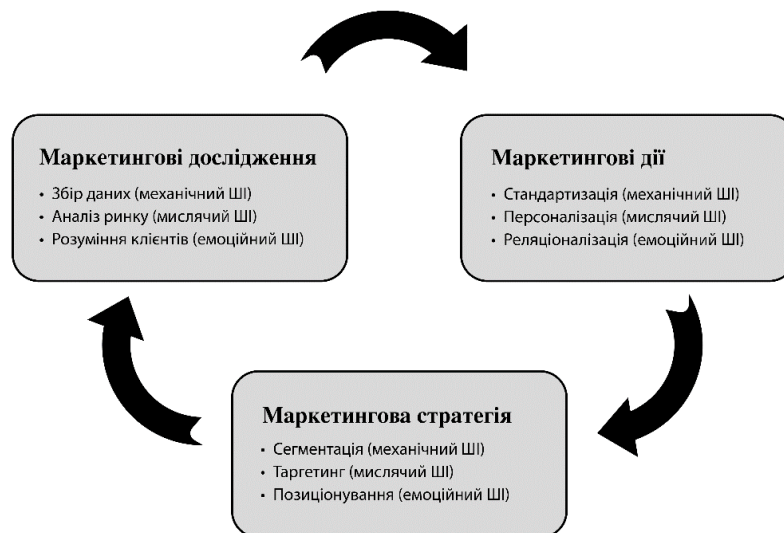
**Вільна генерація контенту.** Сучасний генеративний штучний інтелект дозволяє практично миттєво створити унікальні текстові описи, зображення чи навіть готові веб-сторінки, аудіо- та відеоконтент. Усе це може використовуватися для генерації рекламних матеріалів, написання постів для аудиторії, наповнення інформацією карток продуктів, створення унікального дизайну тощо, зважаючи на конкретні потреби бізнесу, кампанії чи користувачів.

Зі швидким розвитком та постійним прогресом у сфері технологій глибокого навчання, аналізу відео, біометрії, розпізнавання зображень, систем автоматичної обробки запитів та обробки природної мови (NLP) тощо штучний інтелект буде стабільно розвиватися, імплементуватися, адаптуватися під потреби бізнесу і реформувати електронну комерцію у майбутньому.

## РОЗДІЛ 2. МЕТОДИ МАШИННОГО НАВЧАННЯ ТА ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ПІДТРИМКИ МАРКЕТИНГОВИХ РІШЕНЬ В ЕЛЕКТРОННІЙ КОМЕРЦІЇ

### 2.1. Концепція впровадження технологій штучного інтелекту для оптимізації маркетингових стратегій

Запропонована нижче концепція ґрунтується на трьохетапному стратегічному фреймворку впровадження технологій ШІ в процесі маркетингу (рис. 1.1), в основу якого покладено цикл Research (маркетингові дослідження) – Strategy (маркетингова стратегія) – Action (маркетингові дії) [14]. Цикл розглядає стратегічне планування як циклічний процес, який включає проведення маркетингових досліджень для розуміння ринку, фірми, конкурентів і клієнтів, розробку стратегій сегментації, цілепокладання і позиціонування та проектування конкретних маркетингових дій для виконання стратегії. Цикл не закінчується на маркетингових діях, їх виконання буде сприяти поверненню до маркетингових досліджень і, таким чином, почнеться новий цикл.



*Рис. 1.1. Застосування ШІ на різних етапах маркетингового циклу Research – Strategy – Action*

*Джерело: сформовано автором на основі [14]*

На рисунку показано, що штучний інтелект може виконувати критичні ролі на всіх трьох етапах стратегічного маркетингу. При цьому ШІ розглядається з

кількох аспектів «розумовості», якими може користуватися маркетолог: механічний, мислячий і емоційний [15, 16].

**Механічний ШІ** призначений для автоматизації повторюваних та рутинних завдань. Наприклад, дистанційне зондування, машинний переклад, алгоритми класифікації, алгоритми кластеризації, зменшення розмірності даних - це деякі поточні технології, які можна вважати механічним ШІ.

**Мислячий ШІ** призначений для обробки зазвичай неструктурованих даних з метою прийти до нових висновків або рішень. Мислячий ШІ добре розпізнає регулярні патерни, шаблони даних, наприклад, в аналізі тексту, розпізнаванні мови та обличчя. Машинне навчання, нейронні мережі та глибоке навчання - це деякі поточні методи, за допомогою яких мислячий ШІ обробляє дані, тоді як IBM Watson, експертні системи та системи рекомендацій - деякі поточні застосування для прийняття рішень.

**Емоційний ШІ** призначений для безпосередньої взаємодії з людиною, і/або для аналізу почуттів та емоцій людей. Деякі поточні технології включають аналіз настрою, обробку природної мови (NLP), перетворення тексту в мову (TTS), рекурентні нейронні мережі (RNN), чатботи для імітації мови людини, віртуальних агентів для взаємодії з людьми та роботи з налаштованим обладнанням для виявлення емоційних сигналів [17].

Кожен тип ШІ здатен вирішувати свої концептуальні завдання на кожному етапі маркетингового циклу, використовуючи сучасні техніки машинного та глибинного навчання (рис. 1.2), які детальніше розглянуто у підрозділі 2.2.

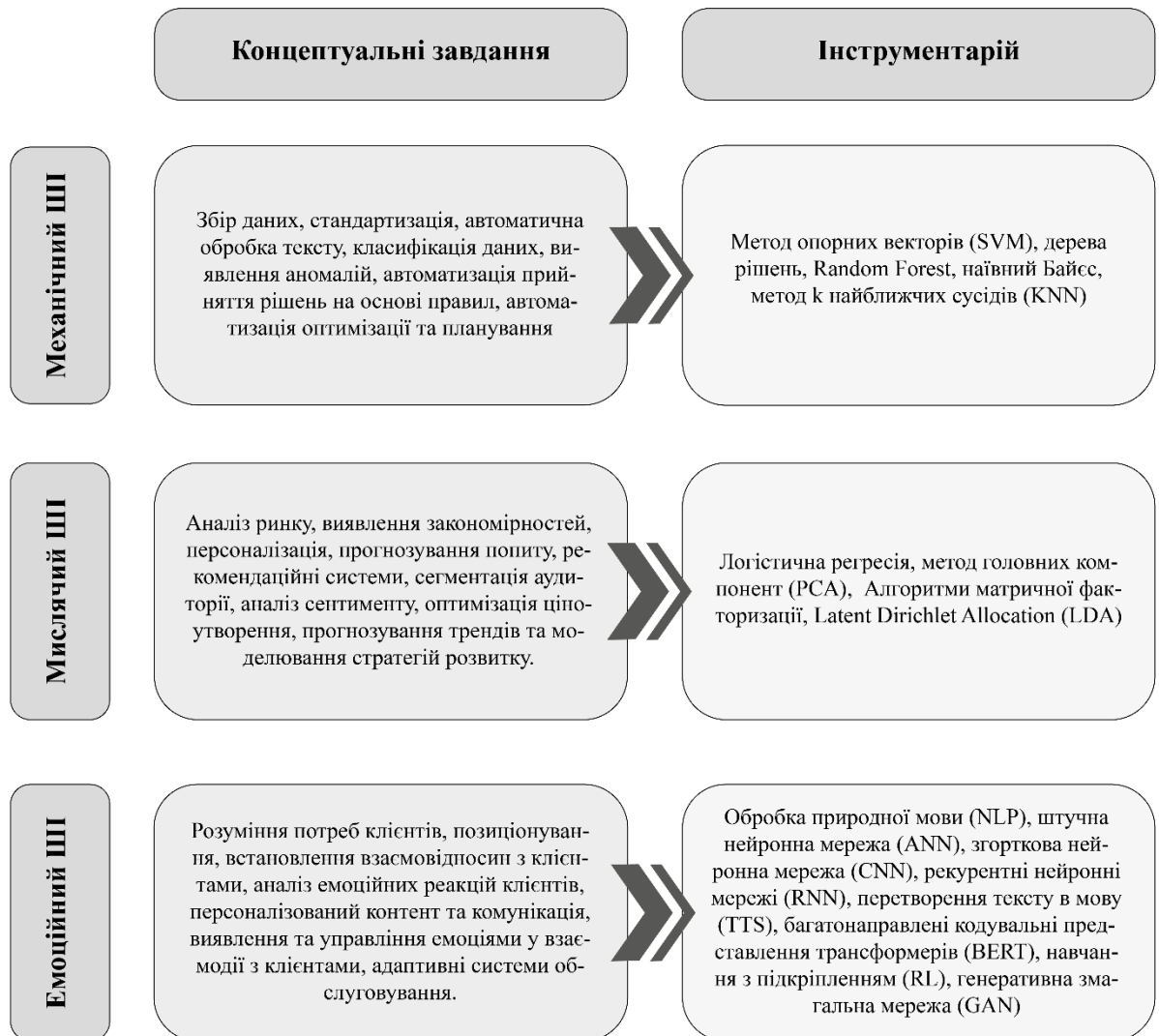


Рис. 1.2. Концептуальні завдання маркетингового циклу та інструментарій інтелектуального аналізу даних для їх вирішення

Джерело: сформовано автором

Повертаючись до фреймворку (рис. 1.1), імплементація технологій ШІ в маркетинг передбачає:

- **на етапі маркетингових досліджень** механічний штучний інтелект використовується для збору даних, мислячий штучний інтелект для аналізу ринку та емоційний штучний інтелект для розуміння потреб клієнтів;
- **на етапі маркетингової стратегії** штучний інтелект використовується для стратегічних рішень щодо сегментації (механічний ШІ), цільової аудиторії (мислячий ШІ) та позиціонування (емоційний ШІ);

- **на етапі маркетингових дій** штучний інтелект використовується для стандартизації, персоналізації та встановлення взаємовідносин з клієнтами, окремо або в синергії. Тут маркетологам доцільно вирішувати, які типи штучного інтелекту використовувати для конкретних маркетингових дій. Наприклад, функції оплати та доставки можуть скористатися перевагами механічного штучного інтелекту, забезпечивши автоматичну оплату та трекінг доставки. Цифровий маркетинг може підвищити рівень персоналізації, використовуючи мислячий ШІ, формуючи, наприклад, різні системи рекомендацій. У контексті клієнтського обслуговування може бути використаний емоційний штучний інтелект: роботи, що вітають клієнтів, чатботи, які надають клієнтське обслуговування тощо.

Щоб збалансувати інтереси маркетологів та клієнтів, концепція використання штучного інтелекту в маркетингових діях розкривається через маркетингові комплекси 4P/4C. Табл. 1.1 визначає стратегічні елементи цього фреймворку.

Таблиця 1.1

Стратегічний фреймворк використання ШІ в маркетингу			
Стратегічне рішення \ Тип ШІ	Механічний (Mechanical AI)	Мислячий (Thinking AI)	Емоційний (Feeling AI)
Маркетингові дослідження	Автоматизація безперервного визначення, відстеження, збору і обробки даних про ринок і клієнтів	Аналіз ринку для виявлення конкурентів і конкурентних переваг	Використання даних поведінку клієнтів, їхнього емоційного аналізу для розуміння існуючих та потенційних потреб і бажань клієнтів
Маркетингова стратегія (STP)	Сегментація	Таргетинг	Позиціонування
Маркетингові дії (4Ps/4Cs)	Стандартизація	Персоналізація	Встановлення взаємозв'язків
Product/Consumer	Автоматизація процесу і результату задоволення потреб і бажань клієнтів	Персоналізація продуктів	Розуміння емоційних потреб клієнта для їх задоволення

Стратегічний фреймворк використання ШІ в маркетингу			
Стратегічне рішення \ Тип ШІ	Механічний (Mechanical AI)	Мислячий (Thinking AI)	Емоційний (Feeling AI)
Price/Cost	Встановлення ціни та автоматизація оплати	Індивідуальні ціни залежно від готовності клієнта платити	Обговорення вартості в діалоговому режимі
Place/Convenience	Керування доступом до продуктів	Персоналізація інтерактивної взаємодії	Персоналізація досвіду для залученості користувачів
Promotion/Communication	Автоматизація спілкування з клієнтами	Персоналізація рекламного контенту	Модифікація спілкування на основі емоцій, уподобань та реакцій клієнта

*Джерело: сформовано автором на основі [14]*

### Етап маркетингових досліджень

На цьому стратегічному етапі можна використовувати механічний штучний інтелект для збору даних, мислячий штучний інтелект для аналізу ринку та емоційний штучний інтелект для розуміння клієнтів.

**Механічний ШІ для збору даних.** Відстеження та збір даних про ринок, оточуюче середовище, фірму, конкурентів, клієнтів є рутинними, повторюваними завданнями, які легко можна автоматизувати з використанням механічного штучного інтелекту.

Наприклад, інтелектуальний аналіз споживачів, включаючи дані про них, їхню діяльність та середовище, якщо користувачі використовують підключені пристрої, можуть бути візуалізовані з Інтернетом речей (IoT), технології та аналітика можуть містити в собі неструктуровані дані, необхідні для маркетингової діяльності, датчики в автомобілі можуть відстежувати поведінку за кермом для визначення страхових премій, а технології, такі як теплові карти, відеоспостереження та сигнальні маячки, можуть бути використані для профілювання та визначення роздрібних покупців.

Також механічний ШІ може бути використаний для полегшення збору експериментальних даних для визначення думок, настроїв споживачів. SurveyMonkey та SurveyCake - це дві комерційні платформи для проведення опитувань, які автоматизують проектування опитувань та збір даних з них.

**Мислячий штучний інтелект для аналізу ринку.** Мислячий штучний інтелект може бути використаний для ідентифікації конкурентів на визначеному ринку або визначення можливих альтернатив у новому ринковому сегменті, а також для отримання інсайтів щодо конкурентних переваг продукту.

Якщо структура ринку відома маркетологам, може використовуватися навчання з учителем, тоді як для нового ринку або виявлення можливих альтернатив доцільно було б застосувати навчання без учителя. На практиці часто використовуються прогностичні аналітичні методи для передбачення тенденцій на ринку та гетерогенних уподобань споживачів. Наприклад, компанія Gap, виробник модного одягу, здійснює прогноз модних тенденцій, які надвичайно мінливі, для кращого задоволення потреб клієнтів, а Amazon - прогнозує майбутні замовлення клієнтів.

Автоматизований аналіз текстів може допомогти отримати нові маркетингові інсайти, алгоритми машинного навчання та класифікація текстів на основі словників - аналізувати набори даних у соцмережах. Специфічні приклади застосування включають створення карт структур ринку для великих роздрібних асортиментів за допомогою мовного моделювання нейронної мережі, шляхом аналізу співвідношень продуктів у кошиках покупців, виявлення копіювальних мобільних додатків, а також допомога в інжинірингу вмісту соціальних медіа за допомогою алгоритмів обробки природної мови, які простежують зв'язки між контентом в соціальних медіа та залученням користувачів.

**Емоційний штучний інтелект для розуміння клієнтів.** Емоційний штучний інтелект може бути використаний для розуміння потреб і бажань існуючих та потенційних клієнтів: хто вони, що вони хочуть та які їхні поточні рішення. Основна відмінність між аналізом ринку та розумінням клієнтів полягає

в тому, що останнє часто включає в себе дані про відчуття, почуття, емоції, вподобання та ставлення користувачів.

Для існуючих клієнтів маркетологи можуть використовувати емоційний штучний інтелект з метою проведення опитування, чи задоволені вони продуктом і чому, а також для аналізу отриманих результатів. Вподобання існуючих клієнтів стабільніші, до того ж можна вільно оперувати даними про минулі та поточні транзакції для глибшого розуміння ситуації. Наприклад, Affectiva співпрацювала з Ford для створення аналізу емоцій AutoEmotive, щоб визначати емоційний стан водіїв.

Для потенційних клієнтів маркетологи можуть використовувати емоційний штучний інтелект для розуміння того, чого вони хочуть і чому вони задоволені конкурентами або зовнішніми альтернативами. Потреби та бажання потенційних клієнтів складніше передбачити, а їхні емоційні дані менш доступні. У маркетинговій практиці Albert AI, платформа маркетингу, що працює на основі штучного інтелекту від Adgorithm, використовувалася компанією RedBalloon для виявлення і залучення нових клієнтів та компанією Harley-Davidson для ідентифікації потенційних споживачів на основі бази даних управління взаємовідносинами з клієнтами компанії, персоналізуючи маркетингову кампанію.

У наукових дослідженнях існують різні підходи до використання емоційного штучного інтелекту для розуміння клієнтів. Наприклад, емоції, виражені споживачами в соціальних медіа (наприклад, відгуки в Інтернеті, твіти), включаючи явний та неявний мовний вираз та моделі дискурсу, можуть бути проаналізовані для розуміння реакції споживачів за допомогою їх власної мови, взаємодія між розмовним штучним інтелектом та клієнтами може бути покращена за допомогою аналітичного відображення для скрипта відповідних послідовностей відповідей, що дозволяє клієнтам відчувати, що вони "розмовляють" з ШІ, а потреби клієнтів можуть бути визначені на основі контенту, створеного користувачем, за допомогою нейронних мереж.

### **Маркетингова стратегія (STP)**

На цьому стратегічному етапі маркетологи можуть використовувати штучний інтелект для трьох ключових стратегічних рішень: сегментації, таргетингу та позиціонування. Однак, перед переходом до конкретних рішень STP, маркетологи повинні визначити загальне стратегічне позиціонування. Хуанг та Раст пропонують технологічно-орієнтований підхід до позиціонування стратегії фірми вздовж вимірів стандартизації–персоналізації та транзакцій–відносин [18]. Так, компанія може прямувати до стратегії товару, що використовує автоматизовану/роботизовану технологію, стратегії відносин, що культивує позитивну цінність існуючих клієнтів, статичної стратегії персоналізації, що використовує аналітику крос-секційних великих даних (наприклад, схожих за думкою клієнтів) для персоналізації, або стратегії адаптивної персоналізації, що використовує історичні дані клієнтів для динамічної персоналізації з часом. Це стратегічне позиціонування керуватиме рішеннями фірм щодо STP. Наприклад, якщо компанія прямує до стратегії статичної персоналізації, їй потрібно мати велику релевантну базу клієнтів дійсних та потенційних клієнтів і дозволити навчанню без вчителя досліджувати патерни вподобань або поведінки покупців як основу таргетингу та позиціонування. Якщо ж це стратегія адаптивної персоналізації, компанія може бажати використовувати навчання з вчителем для продовження аналізу задоволення/незадоволення існуючих клієнтів з часом (може бути невеликий часовий проміжок).

Загалом, на цьому етапі стратегічних рішень більше покладаються на мислячий штучний інтелект та його здатність обробляти дані для прийняття рішень. Проте, коли мислячий штучний інтелект стає абсолютно рутинним, як це часто буває у випадку сегментаційних застосувань, він має багато спільних рис з механічним ШІ, оскільки просто ідентифікує патерни з даних повторно, без включення багато про мету прийняття нових рішень (наприклад, сегментація, але не перенаправлення).

**Сегментація.** Сегментація полягає в розділенні ринку на частини, при цьому клієнти в кожній частині мають унікальні потреби та бажання. Наприклад, використання гендеру для поділу ринку взуття на сегменти чоловічого і жіночого взуття; використання співвідношення ціни-якості для поділу авіаперевізників на бюджетних та преміум тощо. Методи видобування та групування механічного штучного інтелекту дозволяє йому ідентифікувати нові патерни з даних.

Наприклад, текстовий аналіз та машинне навчання можуть бути використані для автоматичної обробки запитів на отримання кредиту для поділу позичальників на хороших (тих, хто поверне кредит) і поганих клієнтів (тих, хто не поверне) [19], автоматизований текстовий аналіз та кореспонденційний аналіз можуть бути використані для психографічної сегментації споживачів на ринку мистецтва [20] видобування даних може бути використане для отримання туристичних сегментів на основі значення напрямків для споживачів, що краще, ніж класичні методи кластеризації [21], і роздрібні клієнти можуть бути мікро-сегментовані на основі їхніх вподобань для персоналізованих рекомендацій [22].

**Таргетинг.** Таргетинг - це вибір правильних сегментів, на яких фірма зосереджує свої маркетингові заходи. Розділення ринку є більш механічним і може бути автоматично здійснене механічним ШІ, за умови наявності відповідних даних. Проте вибір правильного сегмента вимагає знань в галузі, суджень і інтуїції. Для таргетингу використовуються різні технології, такі як пошукові системи, що використовують ключові слова і історію переглядів для таргетингу споживачів, а також платформи соціальних медіа, які використовують інтереси, контент та зв'язки для таргетингу користувачів соціальних медіа. У даному випадку ШІ може рекомендувати різні потенційні цільові групи для подальшого кінцевого рішення менеджером, та застосовувати прогностичне моделювання для вибору сегмента.

**Позиціонування.** Позиціонування демонструє найкращі характеристики, конкурентні переваги продукту клієнтам. Цей термін часто асоціюється з позиціонуванням бренду або рекламним позиціонуванням для підтримки бажаного ставлення до бізнесу.

У порівнянні з механічною сегментацією і мислячим таргетингом, позиціонування більше стосується того, щоб звертатися до клієнтів так, аби зачепити емоційно, зазвичай з використанням слогану чи влучного висловлювання в рекламній кампанії.

Успішне позиціонування дозволяє брендам закріпитися у свідомості клієнтів та існувати на ринку протягом тривалого часу. Такими прикладами є "Just do it" від Nike, "Be different" від Apple та "I'm loving it" від McDonalds. Використання емоційного штучного інтелекту є ідеальним для цього стратегічного рішення, щоб допомогти розробити влучні слогани, які резонують з таргет-клієнтами.

### **Маркетингові дії**

На цьому стратегічному етапі маркетологи можуть використовувати механічний штучний інтелект для стандартизації, мислячий штучний інтелект для персоналізації та емоційний штучний інтелект для встановлення взаємовідносин з клієнтами. Залежно від того, яка перевага бажана, маркетолог може використовувати кілька ШІ разом або поокремо.

**Наслідки для маркетингових досліджень.** На цьому етапі маркетологи повинні вирішити, (1) як використовувати штучний інтелект для ідентифікації конкурентів (включаючи конкурентів у тій же галузі та поза нею) та для розуміння конкурентних переваг (тобто того, як продукт може краще задовольняти потреби клієнтів), та (2) як використовувати штучний інтелект для виявлення та розуміння існуючих і потенційних клієнтів (тобто, хто вони є, що вони хочуть та які є їх поточні рішення) та для розуміння їх уподобань та почуттів. Для фірм, які приймають теоретично обґрунтований підхід до маркетингових стратегій, дані та інтелект, що виникають на цьому етапі, відіграють критичну роль.

На даний момент опитування, експерименти, інтерв'ю, панелі та дані про продажі все ще є основними методами, якими користуються маркетологи для отримання даних, навіть якщо адміністрування цих методів може бути частково автоматизованим або спрощеним за допомогою технологій. Опитування та

експериментальні методи, як правило, більш теоретично орієнтовані, тоді як інші методи тяжіють до збору даних. Маркетологи часто також покладаються на дані, що синдикуються третіми сторонами, особливо для зовнішніх даних, які складно зібрати компанії. Ці дані, як правило, збираються періодично, після факту (після споживання), і не під час генерації даних.

У сфору чергу нові технології автоматизують більшу частину процесі збору даних, з'єднуючи технології, такі як IoT, соціальні мережі, мобільні додатки, датчики, розумні годинники, Fitbit тощо. Ці підходи дозволяють відстежувати та отримувати дані в реальному часі. Таким чином, вони знаходяться в контексті, формуються під час досвіду споживання, напряду стосуються клієнта.

**Обмеження механічного штучного інтелекту.** Хоча поточний механічний штучний інтелект вміє самостійно збирати та обробляти дані з різних джерел, часто втрачається контекст даних, особливо коли йдеться про емоційні дані. Автоматизований процес збору даних також має загрози, пов'язані з конфіденційністю та захищеністю особистих даних користувачів.

Експерт зі штучного інтелекту з компанії Dell, виступаючи на конференції з обслуговування клієнтів у 2019 році, вказав, що хоча моделювання емоцій не є складним завданням, складність полягає в зборі та аналізі емоційних даних. Наприклад, під час взаємодії з обслуговуванням клієнтів зафіксовано лише зміст і настрої розмови, а не її контекст. Поряд з тим, коли клієнт виражає роздратування, його манера спілкування може різнитися в залежності від різних факторів, таких як наявність співрозмовників, погодні умови або транспортні ускладнення. Навіть якщо голосовий аналіз може визначити настрої голосу клієнта, важко встановити причину його емоційного стану та надати її коректно агентам з обслуговування.

**Недоліки мислячого штучного інтелекту.** Поточний мислячий штучний інтелект може бути необ'єктивним та непрозорим, що може призвести до викривлених рекомендацій або несподіваних наслідків чи рішень, проблем відповідальності, оскільки маркетологи не зможуть ефективно відстежувати та пояснювати дії штучного інтелекту.

Непрозорий штучний інтелект також може призводити до викривлення результатів через неточні вхідні дані. Навіть без умисних відхилень, алгоритми можуть продукувати результати, які містять дискримінаційні або гендерні відхилення, що може порушувати принципи справедливості та толерантності в маркетингових дослідженнях і заходах.

**Обмеження використання емоційного штучного інтелекту** полягає в тому, що ще немає реальних машин, які можуть стовідсотково коректно визначати, як діяти та реагувати на людські емоції. Тому заміна використання механічного і розумового штучного інтелекту на емоційний інтелект може призвести до непередбачених наслідків.

Ще одним важливим аспектом є те, що клієнти можуть не бути готовими до взаємодії з емоційним штучним інтелектом. Лоу та ін. (2019) виявили, що багато клієнтів відключаються від чат-ботів, як тільки вони розуміють, що розмовляють не з людьми.

## **2.2. Методи штучного інтелекту для підтримки запропонованої концепції**

На кожному етапі маркетингової концепції, розглянутої в попередньому підрозділі, а саме Research (маркетингові дослідження), Strategy (маркетингова стратегія), Action (маркетингові дії) - можуть застосовуватися різні методи штучного інтелекту для підтримки стратегій [8].

Так, **на етапі досліджень** для збору даних може використовуватися розпізнавання візуальної інформації (так званий комп'ютерний зір), фрод-моніторинг, веб-скрейпінг, Інтернет речей (IoT), автоматизовані опитування, безпосередньо для аналізу ринку корисною може виявитися предиктивна аналітика, яку можна віднести до мислячого ШІ, а для розуміння клієнтів варто імплементувати інструменти conversational commerce з елементами емоційної розумовості ШІ.

**На етапі маркетингових дій** персоналізація та встановлення зв'язків з громадськістю здійснюється з використанням гіперперсоналізації, векторного пошуку, генеративних ШІ, інтелектуальних ботів.

**На етапі стратегічного планування** також доцільм є використання результатів предиктивної аналітики, генерації контенту, включно з маркетинговими матеріалами, фрод-моніторингу для сегментації.

Деякі з цих методів надалі розглянуто детальніше.

**Розпізнавання візуальної інформації.** Ця технологія (відома також як комп'ютерний зір) дозволяє машинам розпізнавати та розуміти візуальні образи майже на рівні людського сприйняття.

Застосування комп'ютерного зору в сфері електронної комерції може повністю змінити досвід пошуку товарів в мережі. Цілком реальною стане ситуація, коли користувачу не потрібно буде більше виконувати жодних пошукових запитів: просто навести камеру смартфона на фото товару, який його зацікавив, або навіть на сам товар, і ШІ миттєво знайде цей продукт та альтернативи до нього.

Комп'ютерний зір може бути використаний також для автоматизації таких завдань, як контроль якості, категоризація продуктів, віртуальна примірка, обробка повернень, зручний пошук, надання віртуальної допомоги.

Інструменти візуального пошуку на основі штучного інтелекту вже успішно впроваджені такими ритейлерами, як Target та ІКЕА. Крім того, ІКЕА поєднує можливості комп'ютерного зору з іммерсивними технологіями, такими як розширена реальність (AR). Це дає їй користувачам неперевершений досвід шопінгу - можливість візуалізувати меблі та предмети домашнього декору у власному приміщенні, не виходячи з дому.

**Фрод-моніторинг.** Алгоритми машинного навчання довели свою ефективність у боротьбі з фінансовими шахрайствами, так як здатні обробляти величезні обсяги транзакцій у реальному часі та виявляти аномальну поведінку користувачів.

Штучні нейронні мережі та алгоритми машинного навчання можуть автоматично виявляти основні патерни в поведінці покупців та відслідковувати потенційно шкідливі транзакції. Ці транзакції часто відбуваються за участю різних пристроїв та каналів, мають незвичайну географію чи час здійснення. За

допомогою інструментів обробки природної мови можна аналізувати спілкування з користувачем та виявляти контент, створений шахраями з використанням ШІ.

Серед основних переваг такого фрод-моніторингу можливість аналізувати великі обсяги даних у режимі реального часу з мінімальним втручанням фахівців 24 години на добу 7 днів на тиждень. Також часто такі системи моніторингу дозволяють виявляти приховані закономірності, які б людина банально не помітила, а також вони здатні до постійного вдосконалення та навчання на нових даних.

Наприклад, платформа електронної комерції Shopify використовує алгоритми машинного навчання для виявлення потенційно небезпечних дій. Система приділяє увагу таким факторам, як місце здійснення транзакції, пристрій, що використовується для здійснення операції, а також історія поведінки користувача. Подібні системи також використовуються в Amazon, Alibaba, PayPal, Mastercard та інших провідних платіжних системах.

**Предиктивна аналітика.** Алгоритми машинного та глибокого навчання вже активно застосовуються для аналізу даних й побудови прогнозів, для сфери електронної торгівлі це може бути предиктивна аналітика щодо динаміки цін, транзакцій, користувацької поведінки тощо.

Предиктивна аналітика на базі ШІ ґрунтується на обробці великих даних за допомогою алгоритмів машинного/глибокого навчання та методів статистичного аналізу. Таке застосування ШІ в e-commerce дозволяє визначити потреби, інтереси та запити користувачів, розпізнати патерни шопінгу в історії покупок, спрогнозувати тренди ринку, попит на продукцію, динаміку формування цін тощо. Відтак ритейлер може ефективно планувати свої товарні запаси, вчасно створювати нові перспективні продукти, формувати маркетингову стратегію на основі визначених уподобань, інтересів та очікувань аудиторії та визначати оптимальні ціни на товари й послуги в реальному часі, загалом ефективно оцінювати та управляти ризиками, мінімізувати вплив зовнішніх факторів, форс-мажорів тощо.

Яскравим прикладом успішного предиктивного аналізу стало співробітництво DHL та IBM у 2017 році, коли їхня прогностична модель штучного інтелекту вчасно передбачила вибухове зростання попиту на спінери - "хайпові" механічні іграшки. Що цікаво, у своїх прогнозах модель спиралася на динаміку переглядів вірусних відео і пошукових запитів. Наразі логістичний оператор DHL розвиває ШІ-систему Global Trade Barometer, яка зможе прогнозувати тренди у глобальній торгівлі, використовуючи при цьому близько 240 млн змінних та дані 7 найбільших економік світу.

**Conversational commerce.** ШІ відкриває нову епоху в історії онлайн-торгівлі, оскільки дозволяє продавцям відступити від звичного UX платформ електронної комерції, в основі якого каталог товарів та корзина покупок. На зміну їм можуть прийти цифрові продавці-консультанти, які надаватимуть користувачам товари у формі живого спілкування: так само просто, як це роблять консультанти у звичайних магазинах.

Великі мовні моделі (BMM або LLM від англ. large language model) вже сьогодні дозволяють роздрібним торговцям створювати віртуальних консультантів, які можуть відповідати на запитання щодо товарів та послуг не гірше за людину. При цьому вони працюють у режимі 24/7 і здатні навчатися, отримуючи зворотний зв'язок від клієнтів і адаптуючись під індивідуальні уподобання кожного користувача.

Типовим прикладом такого консультанта сьогодні є голосовий асистент Alexa від Amazon, який дозволяє користувачам отримувати рекомендації, замовляти товари та послуги онлайн у формі живого спілкування. Компанія з продажу косметики Sephora розробляє ШІ-чатбота Virtual Artist, який надає користувачам поради під час шопінгу. Навіть франшиза Domino's Pizza експериментує з інтеграцією мовного ШІ у свою систему обробки замовлень.

**Підтримка користувачів.** Системи чатботів на основі штучного інтелекту дозволяють роздрібним торговцям створити віртуальний кол-центр, який буде доступний користувачам цілодобово, забезпечуючи швидке та ефективне реагування на їх запити.

Незважаючи на це, якісна підтримка користувачів все ще може потребувати втручання людей (принаймні, у найближчій перспективі). Однак технології ШІ дозволяють підприємствам значно скоротити витрати на кол-центр, покращити оперативність, знизити проблеми мовного бар'єру та підвищити загальну якість обслуговування.

Одним із типових прикладів повноцінного ШІ-бота користувацької підтримки є голосовий асистент Isha, який сприяє корпорації Sony у наданні сервісу в Індії, адаптуючись до багатомовного та мультикультурного середовища цієї країни. Щодо нашого досвіду у компанії WEZOM, нещодавно було інтегровано ШІ у систему користувацької підтримки для виробника кліматичної техніки Cooper&Hunter. Розпізнавання голосу користувачів суттєво прискорило обробку запитів у службі підтримки компанії.

**Гіперперсоналізація.** Хоча механізми персоналізованих рекомендацій вже давно використовуються такими гігантами як Google та Facebook, розвиток штучного інтелекту дозволяє ритейлерам створювати унікальні профілі для кожного користувача, уникаючи прив'язки до категорій аудиторії.

Гіперперсоналізація надає кожному користувачеві індивідуальні рекомендації та контент, які належним чином відповідають його поведінці, запитам та уподобанням. При цьому, портрет користувача формується комплексно, з використанням інформації про його місце проживання, демографічні дані, ну і звісно історії пошуку та покупок.

Цей підхід дозволяє ритейлерам максимізувати задоволення та лояльність споживачів, надаючи їм персоналізований сервіс. Світові лідери у сфері роздрібної торгівлі, такі як Amazon та Alibaba, вже успішно використовують подібні механізми. Наприклад, Starbucks реалізував у своєму додатку механізм підбору унікальних пропозицій на основі активності, історії покупок та особистих уподобань користувача.

Кожен бізнес прагне максимально задовольнити потреби та інтереси своїх споживачів, а маючи індивідуальний підхід до кожного це вдається легше та ефективніше. Персоналізація збільшує конверсії, оскільки користувачам

рекомендуються продукти, які їм дійсно цікаві, підвищує середній чек і прибутковість, так як сприяє крос-селінгу та ап-селінгу, покращує лояльність клієнтів, допомагає компаніям глибше розуміти поведінку користувачів.

**Векторний пошук.** Моделі NLP (Natural Language Processing) можуть покращити пошуковий досвід користувачів, орієнтуючись на релевантність. Навіть у випадках, коли покупцю важко чітко сформулювати свої потреби, він може розпочати пошук з понять, що йому близькі. Це схоже на той спосіб, яким ми доходимо до висновків у розмові з іншими людьми.

Таким чином, імплементація векторного пошуку може сприяти зростанню коефіцієнта клікабельності (CTR), підвищивши релевантність, індивідуальність результатів пошуку, забезпечити підтримку багатомовності, розширені можливості фільтрації, пошук за зображеннями.

У порівнянні з традиційним методом, який базується на ключових словах, векторний пошук ґрунтується на математичному моделюванні мови, що дозволяє враховувати не лише окремі слова, а й їхні контексти.

Векторний пошук у сфері електронної комерції розвивається динамічно. Наприклад, Amazon використовує свою власну платформу OpenSearch. Іншим прикладом є російський векторного пошуку, який входить до пакету управління базами даних MongoDB Atlas. Ці інструменти легко поєднуються з іншими рішеннями штучного інтелекту, такими як чат-боти, засоби комп'ютерного зору та інші.

**Генерація контенту.** Створення текстового, візуального, аудіо- та відеоконтенту для маркетингу, описів продукту, лендингів і каталогів – мабуть найчастіше згадують саме це завдання, коли говорять про автоматизацію процесів з використанням ШІ.

Генеративний ШІ, такий як ChatGPT, може використовуватися для ведення блогів і соцмереж. Багато брендів уже використовують його для управління сторінками в соцмережах.

Також ШІ може бути використаний у дизайні і візуалізації пропозицій. Наприклад, бренд J'evan використовує ШІ для розробки дизайну ювелірних виробів, заощаджуючи значний час дизайнерам компанії.

ШІ перемагає людину, якщо й не у креативності, то однозначно у швидкості генерації контенту, що дозволяє оперативно створювати описи продуктів, заповнювати картки товарів чи отримувати іншу унікальну інформацію. Також генеративний ШІ спрощує оптимізацію контенту для пошукових систем, забезпечуючи легкий доступ до SEO-вмісту онлайн-магазинів. За допомогою ШІ можна створювати контент у форматах, які складно або неможливо зробити вручну. Також ШІ може бути використаний для дизайну та розробки продуктів, пристосовуючись під завдання конкретного бізнесу.

Якщо спробувати зазирнути у майбутнє, цілком реальною здається картина, коли інтеграція ШІ і сфери е-комерції буде настільки глибокою, що зможе створити персональний, зручний та повністю безпечний досвід покупок для кожного користувача. Клієнти матимуть цифрові облікові записи, які повністю відображатимуть їх зовнішність, стиль і вподобання. Вони зможуть приміряти різні товари та отримувати поради від ШІ-асистентів, зможуть використовувати голосові та жестові команди для управління інтерфейсом онлайн-магазину, замінюючи традиційні методи пошуку. Засоби ШІ співпрацюватимуть з технологіями розширеної та доповненої реальності, дозволяючи клієнтам відчувати продукти перед покупкою. А продавці зможуть прогнозувати замовлення клієнтів і готувати їх до доставки, ще до того, як замовлення буде зроблено. Що цікаво, інструменти ШІ потенційно зможуть аналізувати емоції користувачів і адаптувати комунікацію відповідно до їх стану.

### **2.3. Методи машинного навчання в аналізі поведінки користувачів та оптимізації продажів**

У цьому підрозділі запропоновано короткий огляд технік і алгоритмів машинного та глибокого навчання, які можуть бути використані в дослідженнях для підтримки маркетингових рішень в електронній комерції.

## Техніки машинного навчання

1. **Метод опорних векторів (SVM)** є моделлю машинного навчання, яка використовується для класифікації та регресії. SVM працює, визначаючи оптимальну гіперплощину, яка максимізує відстань між різними класами, що визначається критичними точками даних, відомими як опорні вектори. Може обробляти як лінійно роздільні, так і нелінійно роздільні дані, є особливо ефективною для бінарних та навіть багатокласових задач класифікації.

2. **Дерева рішень** використовуються для задач прогнозування, сегментуючи простір предикторів на прості області для аналізу. Має деревоподібну структуру для прийняття рішень на основі значень ознак. У кожному внутрішньому вузлі дерева застосовується критерій прийняття рішення, щоб визначити найкращу ознаку та поріг для розбиття даних [30]. У задачах класифікації кожен вузол представляє класову мітку, тоді як у задачах регресії вузли містять прогнозоване безперервне значення у цьому підмножині.

3. **Random Forest** є методом ансамблевого навчання, який комбінує декілька дерев рішень для здійснення прогнозів. Він підвищує точність та запобігає перенавчанню шляхом тренування множини дерев на різних підвбірках даних та агрегування прогнозів окремих дерев [31].

4. **Наївний Байєс** базується на припущенні, що ознаки є незалежними та наївно не пов'язаними між собою. Він використовує теорему Байєса для обчислення апостеріорних ймовірностей класів на основі спостережуваних значень ознак. Наївний Байєс широко відомий своєю простотою та ефективністю у навчанні та прогнозуванні, що робить модель достатньо популярною у застосуванні [32].

5. **Логістична регресія** використовує логістичну або сигмоїдну функцію для оцінки ймовірностей належності вхідних даних до різних класів та забезпечує прості та інтерпретовані підходи до задач класифікації, що дозволяють здійснювати точні та ймовірнісні прогнози.

6. **Метод головних компонент (PCA)** є лінійною моделлю, що використовується для відображення високовимірних вхідних ознак у простір нижчої вимірності, зазвичай відомий як латентні фактори або головні

компоненти. PCA має на меті трансформувати початкові дані у набір ортогональних компонент, які пояснюють максимальну варіацію у даних [33].

7. **Алгоритми матричної факторизації** працюють шляхом розкладання початкової матриці на дві або більше матриць нижчої вимірності, які представляють латентні фактори. Ці алгоритми мають на меті знайти представлення даних нижчого рангу шляхом виявлення прихованої структури або шаблонів у матриці [34].

8. **Метод k найближчих сусідів (KNN)** є непараметричним алгоритмом, який прогнозує класову мітку (для класифікації) або цільове значення (для регресії) тестового випадку на основі його схожості з k найближчими сусідами у тренувальних даних. У класифікації класова мітка визначається більшістю голосів серед сусідів, тоді як у регресії береться середнє (або зважене середнє) цільових значень [35].

### **Техніки глибокого навчання**

1. **Штучна нейронна мережа (ANN)** — це обчислювальна модель, натхненна структурою та функціональністю біологічних нейронних мереж у людському мозку. Вона складається з взаємопов'язаних штучних нейронів або вузлів, організованих у шари, включаючи вхідний шар, приховані шари та вихідний шар. Зв'язки між нейронами мають асоційовані ваги, які коригуються ітеративно шляхом поширення помилки від вихідного шару до вхідного, керуючись визначеною метою або функцією втрат [36].

2. **Згорткова нейронна мережа (CNN)** складається з згорткових шарів, які застосовують фільтри для вилучення ознак з вхідних даних, а потім наступні шари зменшують просторові розміри. Вони демонструють виняткову продуктивність у задачах класифікації зображень, виявлення об'єктів та сегментації зображень [37].

3. **Рекурентні нейронні мережі (RNN)** призначені для обробки послідовних даних і використовують рекурентні зв'язки, що дозволяють інформації передаватися через різні часові гепи. Ключова особливість RNN — це їх рекурентні зв'язки, які створюють циклічну структуру, що дозволяє інформації проходити по циклах, а мережі зберігати форму пам'яті або контексту для обробки та запам'ятовування інформації з попередніх кроків [38].

4. **Навчання з підкріпленням (RL)** передбачає навчання агента через взаємодію з середовищем, отримання зворотного зв'язку у вигляді винагород або покарань на основі його дій, і вивчення відповідності між станами та діями, які максимізують очікувану сумарну винагороду з часом [39].

5. **Генеративна змагальна мережа (GAN)** складається з генеративної мережі та дискримінативної мережі, які беруть участь у змагальній грі. Генератор намагається створювати синтетичні зразки даних, тоді як дискримінатор намагається розрізнити справжні та фальшиві зразки. Завдяки ітеративному тренуванню у цьому змагальному процесі, GAN-мережі виявили надзвичайні можливості у задачах генерації зображень, перекладу зображень та генерації тексту [40].

6. **Багатонаправлені кодувальні представлення трансформерів (BERT)** є потужною попередньо навченою мовною моделлю, введеною Google у 2018 році. BERT навчається у багатонаправленому режимі, прогнозуючи відсутні слова, враховуючи як попередній, так і наступний контекст, що дозволяє краще розуміти загальне речення або документ. Здатність BERT захоплювати контекстуальну інформацію та використовувати попереднє навчання сприяла значним досягненням у розумінні та генерації людської мови [41, 42].

7. **Latent Dirichlet Allocation (LDA)** — це статистична модель тематичного моделювання, яка використовується для виявлення прихованих тем у великих колекціях текстів. LDA розглядає кожен документ як суміш тем, а кожен тему — як розподіл слів. Вона використовує байєсівський підхід для визначення ймовірностей, що кожен документ належить до певної теми, а кожне слово належить до теми. Ця модель широко використовується в задачах кластеризації текстів, аналізу тем та інформаційного пошуку [43].

Методи глибокого та машинного навчання швидко розвиваються, з новими архітектурами, алгоритмами, техніками та здатністю вивчати складні представлення даних вони значно просунули галузь штучного інтелекту. Наразі ці технології дозволяють не лише ефективно аналізувати великі обсяги даних, але й знаходити приховані закономірності патерни та інсайти, які раніше залишалися поза увагою дослідників та які можуть бути використані для побудови чи коригування стратегій, обґрунтуванні бізнес-рішень.

### **РОЗДІЛ 3. РОЗРОБКА ПРОГРАМНОГО ЗАСОБУ ДЛЯ АНАЛІЗУ ВІДГУКІВ КОРИСТУВАЧІВ ДЛЯ ПІДТРИМКИ ПРИЙНЯТТЯ МАРКЕТИНГОВИХ РІШЕНЬ**

#### **3.1. Розуміння бізнес-процесів та постановка цілей дослідження**

Метою дослідження є створення програмного засобу (ПЗ) для аналізу текстових даних відгуків користувачів, який допоможе маркетинговій команді автоматизувати цей процес, отримувати інсайти з даних і приймати обґрунтовані рішення щодо покращення продуктів та послуг.

Основні цілі включають:

- автоматизацію обробки та аналізу великого обсягу текстових даних відгуків;
- більш якісне розуміння споживачів та їхніх потреб;
- покращення якості обслуговування клієнтів.

Оцінка відгуків користувачів є критично важливою для будь-якого бізнесу, оскільки надає цінну інформацію, інсайти, які можуть суттєво вплинути на стратегічні та оперативні рішення компанії.

Перш за все, відгуки варто аналізувати, аби краще зрозуміти потреби, бажання, очікування користувачів і на основі цього раціонально спрямовувати зусилля на важливі для клієнтів аспекти та підвищувати рівень їхньої задоволеності. Також окремо варто виділити важливість швидкого виявлення проблем, згаданих у відгуках, оскільки оперативне реагування на них та вирішення може вберегти бізнес від втрати клієнтів, прибутків та іміджу.

Відгуки часто містять інформацію про дефекти, які не були виявлені внутрішнім тестуванням, чи негативний досвід користувача у взаємодії з продуктом. Також користувачі часто неявно або напряду можуть подавати ідеї для вдосконалення діяльності бізнесу, реалізуючи які, компанія підвищує якість свої послуг, забезпечує свою інновіційність та конкурентоспроможність на ринку.

Також дуже важливим фактором для успішного функціонування бізнесу є формування та підтримка лояльності клієнтів. Реагуючи на відгуки, враховуючи думки користувачів, компанія демонструє готовність сприймати критику, вести відкриту комунікацію, брати на себе відповідальність як за успіх, так і невдачі – усе це зміцнює довіру до бренду.

Окрім вище переліченого, інсайти з відгуків користувачів можуть допомогти маркетологам сформулювати меседжі, які враховують найбільш важливі аспекти продукту чи сервісу, а відповідно будуть резонувати з аудиторією. Позитивні відгуки можуть виступати соціальним доказом якості продукту чи послуги, бути використаними у матеріалах, що пропонуються користувачам в рекламних цілях. До того ж аналіз ревію може підтвердити чи спростувати маркетингові гіпотези, виявити слабкі та сильні сторони відносно конкурентів, відслідковувати зміни у вподобаннях користувачів, трендах, це дозволить будувати стратегії на основі реальних даних, фактів, а не припущень, завжди лишатися в контексті поточних ринкових умов.

Етапи, необхідні для реалізації дослідження, представлені на рисунку 3.1 та прокоментовані детальніше у підрозділі 3.2.

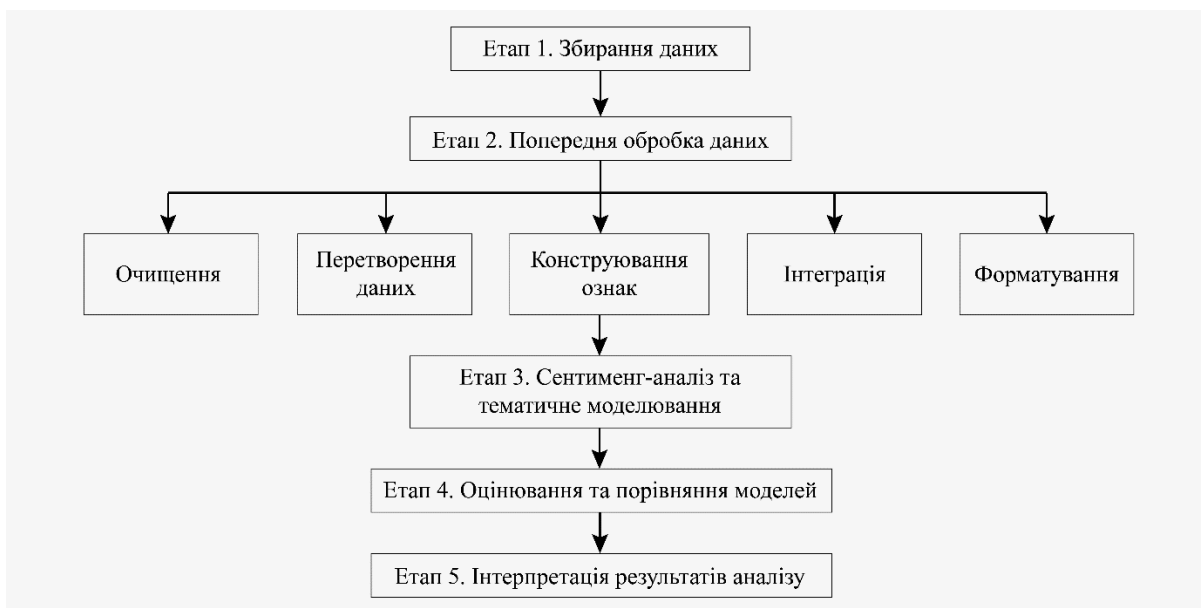


Рис. 3.1. Етапи реалізації аналізу відгуків користувачів з використанням запропонованого ПЗ

Джерело: створено автором

Отже, результатом дослідження є розробка інструменту (програмного засобу), яка дозволить спростити та частково автоматизувати процес аналізу відгуків користувачів продуктів/послуг на онлайн комерційних платформах та стане обґрунтованою базою для прийняття рішень бізнесом щодо розвитку продуктів, маркетингових стратегій та покращення сервісу обслуговування клієнтів.

### **3.2. Характеристика етапів реалізації аналізу відгуків користувачів з використанням розробленого інструменту**

Жоден бізнес, особливо той, що тримає фокус на функціонуванні онлайн, не може обійтися без системи відгуків користувачів. Бізнес в електронній комерції може збирати відгуки користувачів через різноманітні канали та інструменти, такі як вбудовані системи відгуків на вебсайті, автоматизовані електронні листи після покупки, соціальні мережі, спеціалізовані платформи для відгуків (наприклад, Trustpilot або Google Reviews), мобільні додатки з функціоналом для відгуків, чат-боти та онлайн чати, програми лояльності з бонусами за відгуки, QR-коди на продуктах, опитування у фізичних магазинах, а також інтеграцію з аналітичними інструментами для збору та аналізу даних. Які б інструменти не було б обрано, функціонуючи в інтернет-просторі, найчастіше магазин має вже готову базу даних з відгуками користувачів, яка динамічно оновлюється.

#### **Етап 1. Збирання даних**

Передбачає збір даних та завантаження фахівцем датасету з відгуками користувачів у середовище програми.

#### **Етап 2. Попередня обробка даних**

Цей етап є одним з найважливіших з точки зору якості аналізу та включає в себе кілька підетапів, які допомагають підготувати дані для подальшого моделювання. В контексті text-mining відгуків користувачів доцільним буде здійснити наступне.

**Фаза 2.1. Очищення.** Перш за все, це звісно очищення даних, тобто видалення спеціальних символів, чисел, HTML-тегів, якщо такі є, та виправлення

орфографічних помилок у відгуках. Спеціальні символи можуть спотворювати результати текстового аналізу, як і граматичні помилки користувачів, HTML-теги не несуть корисної інформації, тому позбавившись цих пунктів можна отримати чистіші та спрощені дані для подальшого моделювання. Для цих цілей використовувалися такі бібліотеки, як Regular Expressions (re) - для роботи з регулярними виразами, дозволяє здійснювати пошук шаблонів та модифікувати, замінювати їх у тексті; TextBlob - для обробки тексту, яка базується на NLTK і Pattern та надає простий інтерфейс для виконання типових операцій з обробки тексту, таких як розбір синтаксису, аналіз настроїв та корекція орфографії; BeautifulSoup - для парсингу HTML та XML документів, дозволяє легко витягати дані з HTML-тегів та структурувати їх у зручний для аналізу формат.

**Фаза 2.2. Перетворення даних.** Наступний крок – це перетворення даних, яке включає в себе видалення рідкісних слів та лематизацію. Перше здійснюється для зменшення шуму у даних, а лематизація дозволяє зменшити розмір словника та покращити узагальнення моделі. Лематизація у комп'ютерній лінгвістиці — це алгоритмічний процес визначення леми слова на основі його передбачуваного значення. На відміну від стемінгу, лематизація залежить від правильного визначення передбачуваної частини мови та значення слова в реченні, а також у ширшому контексті, що оточує це речення, наприклад, у сусідніх реченнях або навіть у цілому документі. Простими словами на цьому етапі слова приводяться до початкової форми. Для перетворення даних було використано Counter - клас з модуля collections, який використовується для підрахунку кількості унікальних елементів у колекції (наприклад, у списку слів відгуків) та WordNetLemmatizer з бібліотеки NLTK – поширеного інструменту для обробки природньої мови в Python.

**Фаза 2.3. Конструювання ознак.** Фаза конструювання ознак передбачає створення нових ознак чи характеристик з наявних даних для покращення ефективності, точності та надійності моделі машинного навчання. Це здійснюється також задля можливості виявити неочевидні закономірності чи шаблони в даних, які не були помічені раніше, а ще відбір найбільш

інформативних ознак дозволяє зменшити кількість вхідних змінних, що спрощує модель і знижує ризик перенавчання.

Конструювання ознак передбачає генерацію нових ознак на основі наявних, агрегацію, трансформацію (логарифмічна, поліноміальна), one-hot encoding (перетворення категоріальних ознак у набір бінарних ознак), бінінг (розбиття числових ознак на категорії або бінарні значення), видалення низьковаріативних чи корельованих ознак тощо. У контексті аналізу відгуків користувачів в процесі роботи програми були створені такі ознаки, як частота появи певних слів та фраз у тексті, оцінка емоційного забарвлення коментарів тощо. Проте безпосередньо перед аналізом нова ознака була сформована в результаті видалення пунктуації, стоп-слів і загальних малоінформативних коротких слів, що дозволило зосередитися на ключових словах в подальшому. На цьому етапі застосовувалися інструменти з nltk бібліотеки що вже згадувалася раніше, (містить інструменти для роботи з текстом, такі як токенізація, стемінг, лематизація, видалення стоп-слів та інші), а саме stopwords, WordNetLemmatizer, regexp\_tokenize.

Окремо варто виділити такий підхід до обробки тексту, як токенізація. Токенізація - це процес розбиття тексту на окремі частини, які називаються токенами. Токени можуть бути словами, фразами, числами або іншими елементами тексту, залежно від призначення та вимог задачі. У випадку з регулярними виразами (regexp\_tokenize), токенізація виконується з використання певних шаблонів, заданих за допомогою регулярних виразів і це дає можливість доволі гнучко визначати, які частини тексту вважатимуть окремими токенами.

**Фаза 2.4. Інтеграція.** Наступною фазою підготовки даних може бути їхня інтеграція, тобто об'єднання даних з різних джерел або таблиць, проте якщо всі дані зберігаються в одному DataFrame, цей етап не вимагає додаткових дій.

**Фаза 2.5. Форматування.** Останній крок підготовки даних – їх форматування, коли дані приводяться до формату, який є зручним для подальшого аналізу. В даному контексті текст може бути приведено до нижнього

регістру, аби забезпечити консистентність даних, спростити порівняння та аналіз.

### **Етап 3. Сентимент-аналіз та тематичне моделювання**

Написана програма передбачає здійснення аналізу настрою клієнтів, метрик та трендів у відгуках, визначення найчастіше вживаних словосполучень, тематичне моделювання. Для реалізації цілей проекту було обрано моделі VADER, Roberta та Latent Dirichlet Allocation (LDA).

**Vader (Valence Aware Dictionary and Sentiment Reasoner)** — це інструмент аналізу настроїв на основі правил і словника слів, попередньо навчена модель, спеціально розроблена для аналізу текстів у соціальних мережах.

Vader використовує оцінку валентності для кожного слова, щоб визначити його позитивність чи негативність. Показник валентності коливається від -4 до +4, де -4 є найбільш негативним, а +4 - найбільш позитивним. Також модель враховує інтенсивність настрою, яку можна визначити за допомогою великих літер і пунктуації. Так, усі великі літери або знаки оклику можуть вказувати на сильніші почуття. У цьому контексті форматування даних, тобто приведення їх до нижнього регістру у попередній фазі можна упустити.

Серед аналогів можна виділити такі інструменти, як TextBlob та SentiWordNet. TextBlob підтримує полярність та суб'єктивність тексту, SentiWordNet - лексичний ресурс для аналізу думок, проте усі вони можуть не враховувати контекстуальні нюанси та взаємодії між словами. Тоді як Vader забезпечує швидкий і ефективний аналіз настрою для великих обсягів тексту, особливо корисний для соціальних медіа та відгуків, аналізу коротких, неформальних текстів [44].

**RoBERTa (Robustly optimized BERT approach)** – модель обробки природної мови (NLP), розроблена Facebook AI, яка базується на архітектурі Bidirectional Encoder Representations from Transformers та була представлена в 2019 році як покращена версія BERT. RoBERTa використовує архітектуру трансформера, дещо змінену у підходах до навчання, є двонаправленою

моделлю, яка враховує контекст з обох боків слова для точнішого розуміння тексту. Також модель була натренована на значно більшому обсязі текстових даних порівняно з BERT, оптимізувала гіперпараметри навчання, включаючи розмір пакета та кількість, при цьому використовувала корпус даних обсягом 160 ГБ, включаючи дані з новин, книг, Вікіпедії та інших джерел. Також RoBERTa покращила підхід до маскування слів (Masked Language Modeling) шляхом випадкового маскування словосполучень під час навчання, що дозволило моделі краще розуміти контекст [45].

Загалом RoBERTa демонструє одну з найвищих точностей серед моделей NLP на багатьох бенчмарках (показники, що дозволяють оцінити ефективність бізнесу через порівняння з кращими та середніми показниками на ринку), таких як GLUE, RACE та SQuAD. Це робить її ідеальною для завдань сентимент аналізу, де точність і розуміння контексту є критичними. Наразі модель активно підтримується та розвивається спільнотою, а отже є актуальною, живою, з можливостями інтеграції з новими методиками в сфері NLP. Серед можливих аналогів – власне BERT, XLNet (більш складні методи навчання, потребує значніших обчислювальних потужностей), GPT (Generative Pre-trained Transformer - хоч і є генеративним ШІ, проте може використовуватися для подібних завдань).

Тематичне моделювання (topic modeling) — це класифікаційна модель з навчанням без вчителя, схожа на кластеризацію числових даних, яка знаходить природні групи елементів (теми), навіть коли ми не знаємо точно, що шукаємо.

Документ може належати до кількох тем, подібно до нечіткої кластеризації (soft clustering), де кожна точка даних належить більше ніж до одного кластеру.

Тематичне моделювання надає методи для автоматичного впорядкування, розуміння, пошуку та підсумовування великих електронних архівів. Це може бути корисним у виявленні прихованих тем у колекції текстових даних, класифікації документів за виявленими темами тощо.

**Latent Dirichlet Allocation (LDA)** - генеративна модель, яка використовується в машинному навчанні та інформаційному пошуці і дозволяє

пояснювати результати спостережень за допомогою прихованих (латентних) груп, що може допомогти виявити причини схожості деяких частин даних. Наприклад, якщо спостереженнями є слова, зібрані в документи, LDA передбачає, що кожен документ є сумішшю невеликої кількості тем і що поява кожного слова пов'язана з однією з тем документа.

Подібний підхід має схожість з латентно-семантичним аналізом (pLSA), проте в LDA передбачається, що розподіл тем має в своїй основі розподіл Діріхле. На практиці це дійсно дає в результаті більш коректний набір тем.

#### **Етап 4. Оцінювання та порівняння результатів моделювання**

Для оцінки моделей VADER та RoBERTa була побудована візуалізація (рис. 3.2), на якій представлені взаємозв'язки між оцінками полярності (негативної, нейтральної та позитивної), отриманими за допомогою двох різних моделей аналізу тональності тексту. Оцінки полярності для кожного відгуку кодуються кольором та відповідно до рейтингу (1-5).

Цей графік дозволяє дослідити кореляції між різними метриками тональності, наданими моделями, а також зрозуміти, як ці метрики співвідносяться з рейтингами відгуків.

Перш за все, взаємозв'язки між негативними, нейтральними та позитивними оцінками, отриманими за допомогою VADER та RoBERTa, показують певну міру узгодженості між цими двома моделями. Наприклад, спостерігається позитивна кореляція між негативними оцінками, наданими обома моделями (`vader_neg` та `roberta_neg`), що свідчить про те, що обидві моделі загалом погоджуються у визначенні негативної тональності відгуків. Аналогічну картину можна спостерігати для позитивних оцінок (`vader_pos` та `roberta_pos`). Однак нейтральні оцінки (`vader_neu` та `roberta_neu`) демонструють менш явну кореляцію, що може вказувати на різне трактування контенту як нейтрального двома моделями.

Внутрішні взаємозв'язки між оцінками VADER та RoBERTa також можуть надати цінну інформацію. Наприклад, негативна кореляція між негативними та позитивними оцінками в обох моделях є очікуваною, оскільки збільшення

негативної тональності зазвичай супроводжується зменшенням позитивної. Нейтральні оцінки мають слабкіші взаємозв'язки з негативними та позитивними, що відповідає очікуванням, оскільки нейтральні відгуки не повинні мати сильну негативну або позитивну тональність.

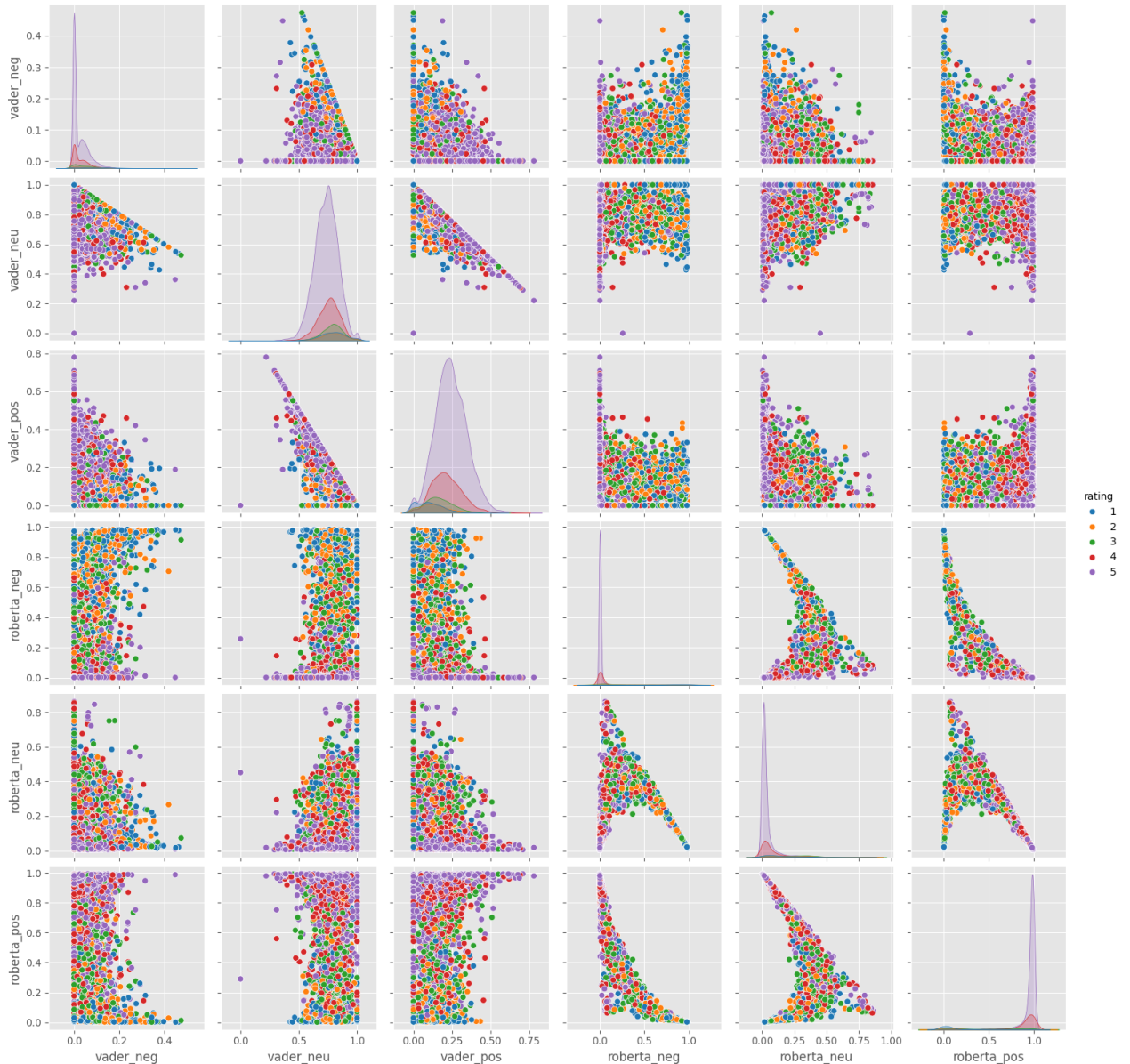


Рис. 3.2. Взаємозв'язки між оцінками полярності VADER та RoBERTa

Джерело: створено автором у середовищі Python

Діагональні графіки, що показують розподіли (KDE) для кожної змінної, надають додатковий контекст щодо розподілу оцінок тональності. Вони показують, що як для VADER, так і для RoBERTa, багато відгуків мають нульові значення для негативної тональності, тобто відсутність негативної тональності.

Позитивні оцінки мають ширший розподіл, таким чином можна говорити про наявність більшої кількості позитивних тональностей у відгуках. Нейтральні оцінки мають більш рівномірний розподіл, особливо для VADER.

Аналіз візуалізації взаємозв'язків між оцінками полярності, отриманими за допомогою моделей VADER та RoBERTa, не надає однозначних підстав для вибору однієї з моделей як переважаючої. Обидві демонструють певну узгодженість у визначенні негативної та позитивної тональності, що свідчить про надійність їхніх результатів. Однак, розбіжності у нейтральних оцінках можуть вказувати на різне трактування нейтрального контенту двома моделями, що може вимагати подальшого дослідження.

З одного боку, модель VADER є більш простою та швидкою у використанні, тож може бути зручнішою для аналізу великих обсягів даних у режимі реального часу. Водночас, RoBERTa, базуючись на глибокому навчанні та трансформерах, пропонує глибше розуміння контексту, що може бути критично важливим для текстів з багатозначністю та складною структурою.

З огляду на це, використання обох моделей одночасно може бути виправданим. Такий комбінований підхід дозволить використати як швидкість та ефективність VADER, так і глибоке розуміння контексту RoBERTa, що дозволить забезпечити більш збалансований аналіз тональності, мінімізуючи можливі упущення та забезпечуючи глибше розуміння відгуків.

### **Етап 5. Інтерпретація результатів аналізу**

Передбачає безпосередньо економічну інтерпретацію отриманих результатів аналізу та видобування інсайтів, що мають практичне значення для бізнесу. На цьому етапі маркетологи аналізують результати роботи програми, а саме візуалізації, графіки, рейтинги, розподіли відгуків за настроями, найбільш часто вживані слова та біграми, результати тематичного моделювання та відгуки з незвичним поєднанням рейтингу та емоційного забарвлення. Завдяки цьому стає можливим глибше розуміння переваг та недоліків продуктів, очікувань та потреб клієнтів, а також тенденцій на ринку.

До того ж цей етап дозволяє реалізувати ідеї концепції трьохетапного стратегічного фреймворку впровадження технологій ШІ в процесі маркетингу Research – Strategy – Action, представленого у підрозділі 2.1.

NLP володіє унікальною здатністю аналізувати, розуміти та генерувати текстову інформацію, що робить його незамінним інструментом для вирішення концептуальних завдань на кожному з етапів. Так, за допомогою методів NLP на етапі маркетингових досліджень можна аналізувати великі обсяги текстової інформації з різних джерел, таких як соціальні медіа, відгуки, новини, статті тощо та визначати тенденції, настрої та думки споживачів, що є важливим для розуміння ринку; на етапі стратегії – можна використовувати для аналізу споживацьких поведінкових патернів, створення деталізованих профілів аудиторії та персоналізації маркетингових пропозицій; на етапі дій - для автоматизації комунікації з клієнтами через чат-боти, створення контенту, який відповідає їхнім потребам тощо.

Запропонована програма та аналіз її результатів характерні саме для етапу Research, на якому проводяться маркетингові дослідження задля розуміння ринку, фірми, конкурентів і клієнтів. Інсайти, отримані під час аналізу, формують рекомендації, які стають базою для рішень на наступних етапах циклу.

На етапі Strategy ці рекомендації використовуються для розробки стратегій сегментації, цілепокладання і позиціонування. Наприклад, розуміння найбільш поширених проблем та побажань клієнтів допомагає визначити ключові сегменти ринку та адаптувати продуктову пропозицію під їхні потреби.

На етапі Action ці стратегії перетворюються у конкретні маркетингові дії, спрямовані на виконання стратегії. Впровадження таких дій може включати запуск нових кампаній, покращення клієнтського досвіду, коригування продукту чи послуги. Виконання маркетингових дій, у свою чергу, сприяє поверненню до маркетингових досліджень, починаючи новий цикл.

Таким чином, програма не лише допомагає здійснювати маркетингові дослідження на етапі Research, але й закладає основи для стратегічного

планування та подальших дій, створюючи замкнений цикл безперервного вдосконалення маркетингових процесів.

### **3.3. Результати реалізації запропонованого сценарію на конкретному датасеті**

Для тестування функціоналу програми було обрано базу даних Sephora Skincare Reviews, та взято частину, яка містить більш ніж 100 тис відгуків користувачів за період кінця квітня 2022 року до другої половини березня 2023, тобто майже за рік.

У базі даних наявно 19 змінних різних типів. До цілочисельних або строкових даних належать поля, які виступають ідентифікаторами, тобто `author_id`, `product_id`. Числові дані включають `rating`, `helpfulness`, `total_feedback_count`, `total_neg_feedback_count` та `total_pos_feedback_count`, `price_usd` – усі, окрім `helpfulness` та `price_usd` є цілочисельними. Логічні бінарні (boolean) змінні представлені лише полем `is_recommended`. `submission_time` - дані типу дата і час. І нарешті текстові дані включають поля `review_text`, `review_title`, `skin_tone`, `eye_color`, `skin_type`, `hair_color`, `product_name` та `brand_name`.

За попередньою оцінкою найбільш багатообіцяючими та такими, що безпосередньо будуть використовуватися в текстовому аналізі є показники `rating`, `review_text`, `submission_time` (для оцінки динаміки). Також цікавими та інформативними, проте допоміжними, є змінні `is_recommended` (показує, чи рекомендує користувач продукт - так або ні), `total_feedback_count` (загальна кількість зворотного зв'язку, отриманого відгуком), `total_neg_feedback_count` (кількість негативного зворотного зв'язку, отриманого відгуком), `total_pos_feedback_count` (кількість позитивного зворотного зв'язку, отриманого відгуком) та `helpfulness` ( $\text{total\_pos\_feedback\_count} / \text{total\_feedback\_count}$  – відображає, наскільки відгук був корисним іншим споживачам). В глобальному контексті менш потрібними здаються поля `author_id`, `eye_color`, `skin_type`, `hair_color` (індивідуальні характеристики автора коментаря), `product_id`, `product_name`, `brand_name`, `price_usd` (якості товару) – усе це може допомогти

розібрати та проаналізувати конкретний кейс, проте, якщо ми говоримо про загальний огляд стану задоволеності користувачів через текстовий майнінг ці поля можна упустити.

До наявних даних було застосовано інструменти базової статистики (див. Додаток А). Щодо якості даних, то такі факти, як незначна кількість пропущених значень (пропуски переважно в `is_recommended`, `helpfulness`, який може бути прорахованих за формулою та серед змінних, які вказують індивідуальні характеристики клієнта), більшість полів мають валідні значення, розподіл даних виглядає рівномірним, а унікальність значень свідчить про різноманіття та потенційну інформативність бази.

Більшість відгуків (70%) мають заголовок, що вказує на позитивне ставлення користувачів до продуктів, такий як "Love it". Більшість відгуків рекомендують товари (86%), а середній їхній рейтинг становить 4.35 з максимальною оцінкою 5. Це свідчить про переважно позитивні досвід покупки та загальне враження про бренд.

З огляду на середній рівень корисності відгуків (0,76) та велику кількість відгуків з пропущеними даними щодо цього аспекту (44%), цілком можливо є необхідність удосконалення процесу надання корисного зворотнього зв'язку іншими користувачами, стимулювання їх до більш розгорнутого та детального фідбеку.

Також відгуки включають інформацію про тип шкіри (87% відгуків), колір очей (88%) та волосся (83%) користувачів. Цікавою статистикою є те, що більшість користувачів, які залишають відгуки, мають світлу шкіру (26%) та карі очі (48%), абсолютна більшість споживачів має комбіновану шкіру. Ця інформація може бути корисною для виокремлення більш вузьких таргет-груп та розуміння їхніх потреб з подальшою під них адаптацією товарів, послуг та маркетингових стратегій.

**Оцінювання** результатів роботи програми буде відбуватися на основі бази даних відгуків Sephora, яка має більш ніж 100000 записів. Для отримання

конкретних результатів за вхідні дані було взято лише ті, які відповідають першим трьом місяцям 2023 року.

**Number of reviews per week.** Графік "Number of reviews per week" (рис. 3.3) демонструє динаміку кількості відгуків користувачів на сайті Sephora з початку 2023 року за 3 місяці. Кожна точка на графіку представляє кількість відгуків, залишених користувачами за відповідний тиждень.

На початку року, протягом першого тижня січня, спостерігається низька кількість фідбеку користувачів, яка різко зростає на другий тиждень, досягнувши більше тисячі коментарів. Упродовж наступних кількох тижнів кількість відгуків дещо знижується, але залишається на відносно високому рівні.

На початку лютого кількість відгуків знижується до приблизно 700-800, а потім поступово зростає до понад 1000 відгуків на початку березня. Після цього спостерігається новий спад до середини місяця, за яким слідує ще одне значне зростання до піку у понад 1300 відгуків на тиждень у середині березня. Після цього кількість відгуків різко зменшується до кінця місяця.

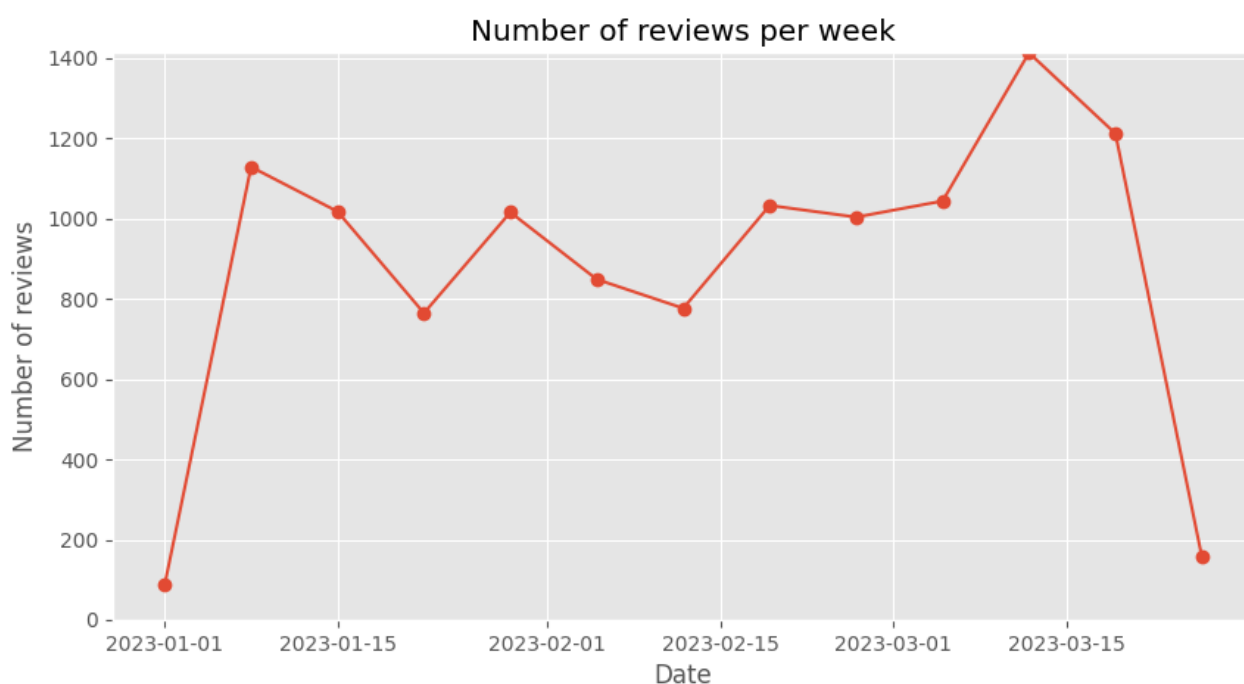


Рис. 3.3. Щотижнева кількість відгуків Sephora за січень - першу половину березня 2023 року

Джерело: створено автором у середовищі Python

**Distribution of customer rating.** Діаграма (рис. 3.4) показує розподіл відгуків користувачів Sephora за рейтингом за перші три місяці 2023 року. Вона представлена у вигляді кругової діаграми, де кожен сегмент відповідає певному рейтингу (від 1 до 5 зірок) та його частці у загальній кількості відгуків.

З цієї візуалізації можна зробити висновок, що переважна більшість (68%) відгуків Sephora мають найвищий рейтинг - 5 зірок. 90% відгуків Sephora мають рейтинг 4-5, що свідчить про переважну задоволеність клієнтів продуктами та послугами Sephora.

Лише 3% відгуків мають найнижчий рейтинг (1-2 зірки), тобто насправді дуже мало клієнтів виявляють незадоволеність.

Враховуючи вище зазначене, компанія може додатково переглянути ті відгуки, які не дотягнули до 5 (22%), аби визначити, які аспекти можна покращити, щоб збільшити кількість максимально задоволених користувачів. Також варто звернути особливу увагу на відгуки з найнижчою оцікою, аби оперативно вирішити проблеми, з якими стикаються клієнти та запобігти подібних ситуацій у майбутньому.

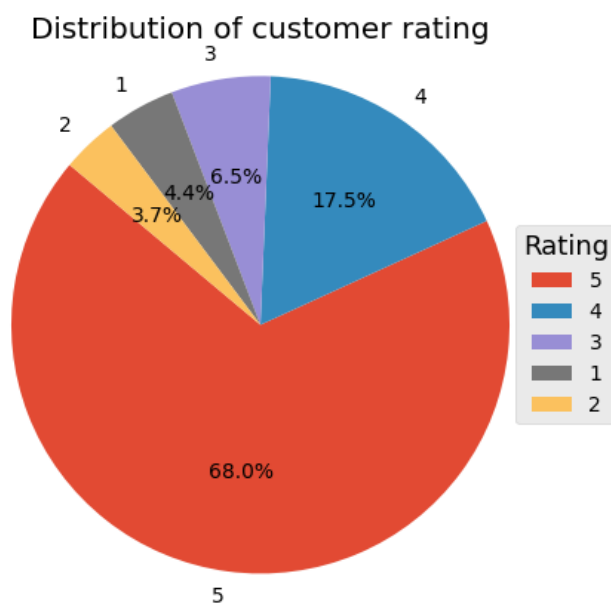
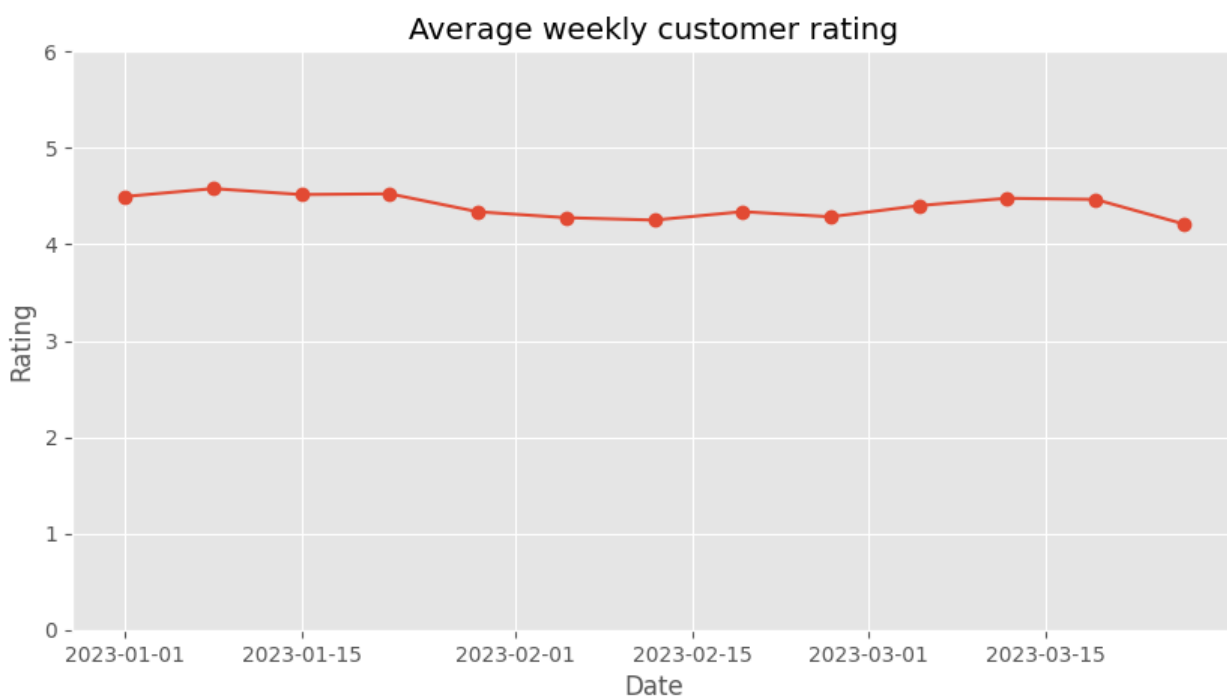


Рис. 3.4. Розподіл відгуків Sephora за рейтингом впродовж січня – першої половини березня 2023 року

Джерело: створено автором у середовищі Python

**Average weekly customer rating.** Графік (рис. 3.5) ілюструє середній тижневий рейтинг відгуків Sephora за перші три місяці 2023 року. Вісь X представляє тижні, а вісь Y - середній рейтинг.

У перші три місяці 2023 року середній тижневий рейтинг клієнтів Sephora незначно коливався у межах від 4,2 до 4,7. Найвище значення показника зафіксований на 12-му тижні (кінець березня), що може бути пов'язано з сезонною активізацією попиту на косметичні продукти.



*Рис. 3.5. Щотижневий середній рейтинг відгуків Sephora впродовж січня – першої половини березня 2023 року*

*Джерело: створено автором у середовищі Python*

**Average weekly sentiment score.** Графік (рис. 3.6) відображає тижневі усереднені показники емоційного забарвлення відгуків користувачів. Вісь X представляє тижні, а вісь Y - емоційне забарвлення від -1 (максимально негативне) до 1 (максимально позитивне).

Аналогічно до попередньої візуалізації протягом перших трьох місяців 2023 року стабільно спостерігається загалом позитивний досвід клієнтів. Знову ж таки, весняний період, ймовірно, сприятливо вплинув на емоційну тональність

їхніх відгуків, оскільки у цей період ілюструється певне зростання показника sentiment score.

Sephora може дослідити, чи є певні продукти або послуги, які особливо популярні у весняний період та розробити маркетингові кампанії, які зосереджені на сезонних вподобаннях користувачів.

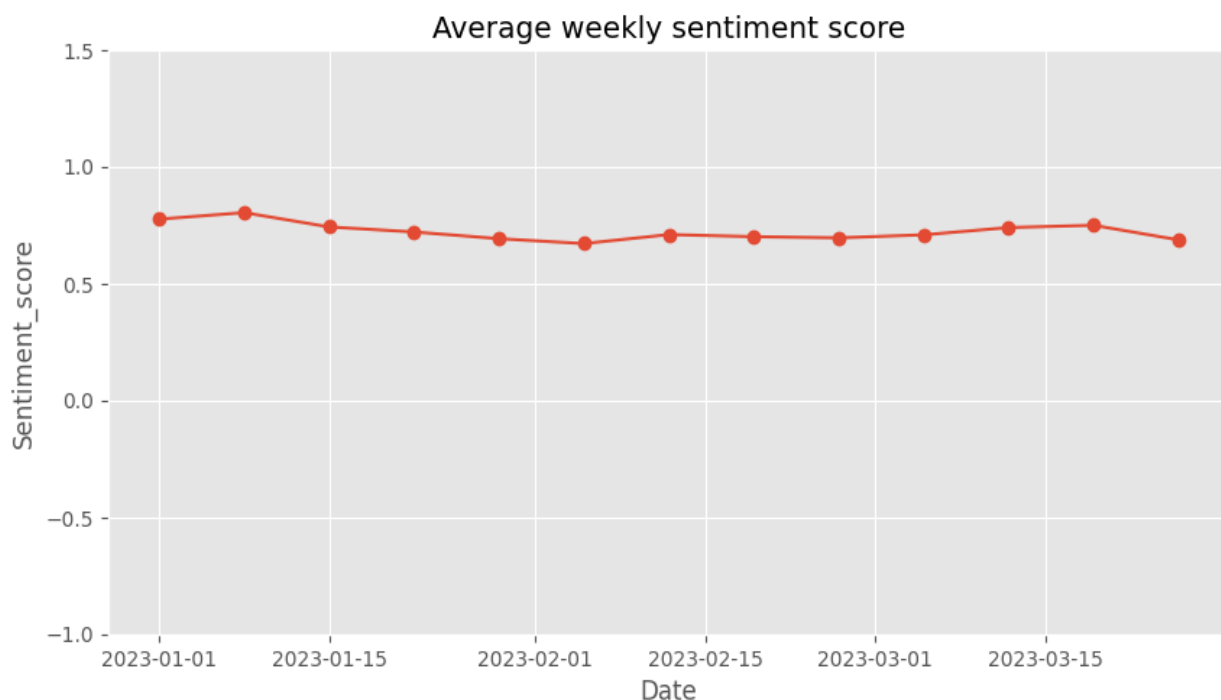


Рис. 3.6. Щотижневий показник тональності відгуків Sephora впродовж січня – першої половини березня 2023 року

Джерело: створено автором у середовищі Python

**The most common and high-frequency words in the review data.** Наступна візуалізація представляє собою хмару слів, що найчастіше зустрічаються у відгуках на продукцію Sephora за перші три місяці 2023 року (рис. 3.7). Розмір слова відповідає частоті його вживання.

Загалом відображення дає уявлення про те, що клієнти Sephora думають про продукти, послуги та який загальний досвід, очікування та потреби мають. Визначення ключових слів та фраз може бути опорною точкою у формуванні маркетингових матеріалів, які були б близькими, цікавими і релевантними для клієнтів. Ну і звісно такий аналіз дозволяє ідентифікувати потенційні або наявні

проблеми чи аспекти, які потребують вирішення, покращення чи особливої уваги.

Дуже важливо використовувати ці візуалізації у поєднанні з іншими методами дослідження ринку, щоб отримати більш комплексне уявлення, а також слідкувати за динамікою змін у хмарі слів, аби розуміти, як міняються думки і настрої клієнтів з часом. До того ж подібне візуальне представлення може бути дуже ефективним у стимулюванні дискусій чи нових ідей під час командних обговорень маркетингової діяльності.

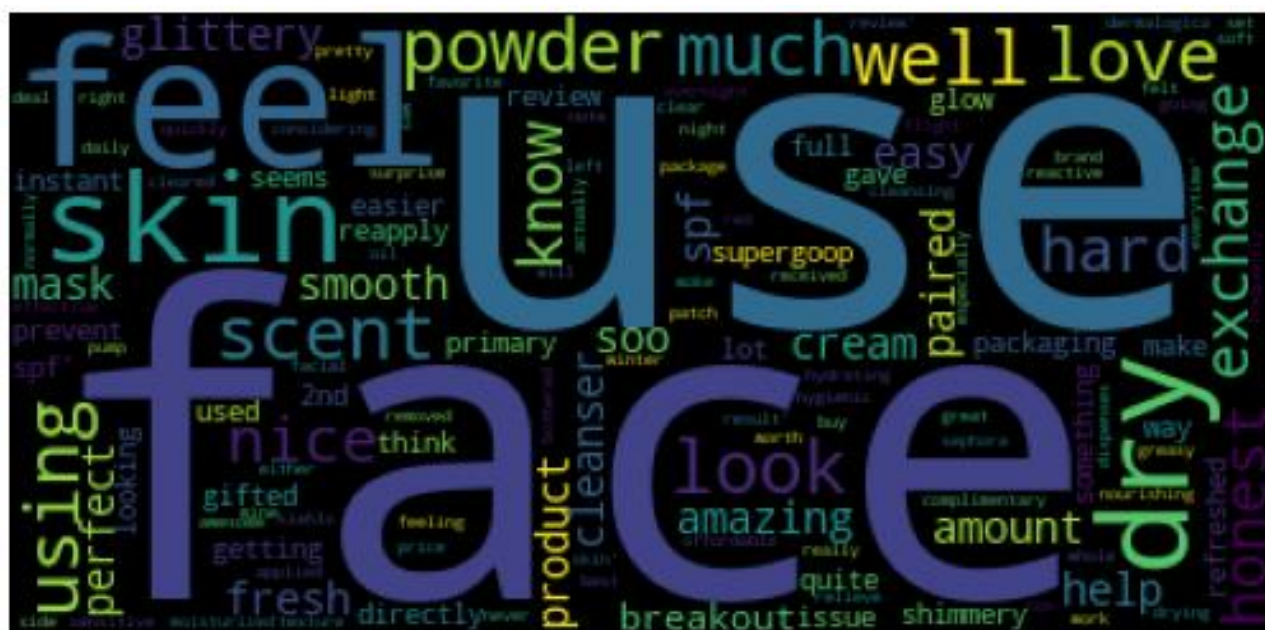


Рис. 3.7. Найбільш поширені слова у відгуках Sephora впродовж січня – першої половини березня 2023 року

Джерело: створено автором у середовищі Python

Аналізуючи конкретно цю візуалізацію, можна виділити наступні думки та рекомендації.

Популярність слів "face", "skin", "dry", "powder", "mask" вказує на високу зацікавленість користувачів у продуктах для догляду за шкірою та макіяжу. Маркетологи можуть використати зацікавленість у цих товарах для просування нових засобів цієї категорії, підкреслюючи їхні унікальні властивості та ефективність.

Використання слів "smooth", "perfect", "amazing", "glow", "glittery", "shimmery" свідчить про те, що користувачі високо цінують продукти, які роблять шкіру гладенькою та сяючою. Розробка нових формул саме таких продуктів може стати ключовим фактором для підвищення їхньої привабливості.

Слова "breakout", "dry", "sensitive" вказують на те, що деякі користувачі стикаються з проблемами висипання чи сухості шкіри. Важливо розробляти продукти, які мінімізують ці побічні ефекти, та акцентувати увагу на них, а також завжди додатково пропонувати варіанти для чутливої шкіри.

Часте вживання "spf", "night", "daily", "reapply" говорить про те, що клієнти нерідко шукають вузькоспеціалізовані товари. Важливо рекламувати продукти призначені для денного та нічного використання, а також засоби з SPF для захисту від сонця, це може дуже позитивно вплинути на продажі.

Також неодноразово у відгуках користувачів зустрічається "packaging", тобто зовнішній вигляд продуктів та їх упаковка також важливі для користувачів. Привабливий дизайн товару, естетичне пакування може стимулювати продажі, оскільки покращить сприйняття товарів покупцями.

Слова "love", "favorite", "nice", "amazing" демонструють загальний позитивний досвід користувачів. Маркетологи можуть використовувати відгуки із цими словами для створення рекламних матеріалів, оскільки підтримка та поширення таких відгуків сприятиме підвищенню довіри до бренду.

Те, що слово "exchange" потрапило у цю візуалізацію, може вказувати на доволі часті випадки необхідності обміну продуктів, тож можливо варто переглянути та оптимізувати політику повернень/обмінів, аби зменшити негативний досвід користувачів та підвищити їхню лояльність.

Слово "review" вказує на активне залишення фідбеку, і дуже важливо заохочувати користувачів продовжувати це робити, оскільки їхні відгуки є цінним джерелом зворотного зв'язку та мають вплив на потенційних покупців.

### **Common words around 3 most frequent words**

У контексті трьох найпоширеніших слів – в даному випадку "face", "use" та "feel" – розглянуто супутні слова, що дозволяє краще зрозуміти, як користувачі

взаємодіють із продукцією та які аспекти є для них найважливішими. Результати подані для зручності у табл. 3.1.

Слова, що найчастіше зустрічаються поруч із "face", такі як "felt", "left", "wash", "put", "wake", "mask", "apply", "throughout", "day", "applied" підкреслюють ритуали, пов'язані з доглядом за шкірою обличчя. Зокрема, слова "wash" і "apply" свідчать про регулярне очищення та нанесення засобів, а "mask" підкреслює популярність використання масок для обличчя. Крім того, слова "throughout" і "day" вказують на використання продуктів протягом усього дня, а слова "felt" і "applied" – на відчуття після застосування.

Таблиця 3.1

Поширені слова навколо 3 найбільш часто вживаних слів у відгуках Sephora впродовж січня – першої половини березня 2023 року					
"Face"		"Use"		"Feel"	
Word	Frequency	Word	Frequency	Word	Frequency
felt	0.014778824	night	0.037038796	makes	0.057002366
left	0.009875367	easy	0.024635656	like	0.02980115
wash	0.009350705	continue	0.016140113	moisturized	0.016305309
put	0.0064895893	daily	0.015379807	made	0.015641825
wake	0.005325566	every	0.014157552	amazing	0.008268606
mask	0.005301419	morning	0.011848609	refreshed	0.0073151523
apply	0.004794844	everyday	0.0076380414	smoother	0.0072192103
throughout	0.004092709	twice	0.007075454	nice	0.0066972193
day	0.0040472313	evening	0.0059707332	hydrated	0.006625606
applied	0.0039499165	day	0.005834504	greasy	0.004752131

Джерело: створено автором у середовищі Python

Слова, що супроводжують "use" - "night", "easy", "continue", "daily", "every", "morning", "everyday", "twice", "evening", "day" - вказують на режим та частоту

використання продуктів. Користувачі активно інтегрують продукти у свої щоденні ритуали догляду за шкірою, включаючи ранковий, вечірній та нічний догляд. Слова "daily" та "everyday" підкреслюють, що багато продуктів використовуються щоденно, а слово "easy" вказує на важливість зручності і простоти використання.

Слова навколо "feel" - "makes", "like", "moisturized", "made", "amazing", "refreshed", "smoother", "nice", "hydrated", "greasy" – переважно підкреслюють враження та відчуття користувачів після застосування продуктів. Так, "moisturized", "smoother" та "hydrated" вказують на зволоження та гладкість шкіри. Позитивне ставлення відображене у словах "amazing", "refreshed", "nice", а ось "greasy" вказує на можливе негативне відчуття жирності після використання, на що потрібно звернути увагу та можливо покращити формулу продукту.

Таким чином, у розробці рекламної кампанії можна зацентувати увагу на ефективності щоденних ритуалів догляду за шкірою, тривалості ефекту від продуктів, зручності використання, особливо в контексті продуктів, які використовуються специфічно зранку, протягом дня чи ввечері.

### **Top 20 Bigram Frequency**

Важливо аналізувати не лише окремі слова, а ще й словосполучення. Окремі слова можуть мати різні значення залежно від контексту, в якому вони використовуються, і лише аналіз біграм, триграм та більших наборів дає змогу точно визначити, про що йде мова.

*Таблиця 3.2*

<b>20 найбільш часто вживаних біграм у відгуках Sephora впродовж січня – першої половини березня 2023 року</b>	
<b>Bigram</b>	<b>Frequency</b>
skin feel	1165
dry skin	950

<b>20 найбільш часто вживаних біграм у відгуках Sephora впродовж січня – першої половини березня 2023 року</b>	
<b>Bigram</b>	<b>Frequency</b>
sensitive skin	710
fine line	643
feel like	604
eye cream	588
make skin	587
skin look	545
honest review	510
highly recommend	482
leaf skin	471
skin feeling	463
skincare routine	428
exchange honest	411
dark circle	354
received product	351
white cast	331
long way	322
skin care	299
using product	282

*Джерело: створено автором у середовищі Python*

Аналізуючи отримані результати для Sephora (табл 3.2), можна дійти наступних висновків.

- Біграми "highly recommend" (482), "honest review" (510) і "exchange honest" (411) свідчать про високий рівень задоволеності користувачів, прозорість та чесність у відгуках та обслуговуванні. Це важливий показник довіри до бренду та його продукції.
- Біграми "dry skin" (950), "sensitive skin" (710), "dark circle" (354), і "white cast" (331) підкреслюють, що споживачі часто стикаються з проблемами сухості, чутливості шкіри, темними колами під очима та білим нальотом від засобів. Це на пряму вказує на болі користувача та потреби їх вирішення, що може використовуватися у розробці та просуванні косметичних продуктів.
- Біграми "skin feel" (1165), "feel like" (604), "skin feeling" (463) вказують на важливість відчуттів на шкірі після нанесення засобів для користувачів.
- Біграми "skincare routine" (428) і "skin care" (299) свідчать про те, що багато покупців інтегрують продукти в свій щоденний догляд. Виходячи з цього, можна розробляти набори засобів та пропонувати рекомендації щодо їхнього комплексного використання.
- Біграма "eye cream" (588) може свідчити про популярність конкретних типів продуктів, в даному разі крему для повік, що дозволяє пропонувати акцент у рекламних кампаніях на специфічні продукти.

### **Тематичне моделювання**

В результаті тематичного моделювання було виділено 5 основних тем слів. Вони мають наступний розподіл:

**Topic #1:** exfoliator, exfoliant, exfoliate, thank, new, contain, sensitive, irritate, smell, feel, bal, bottle, packaging, oil, skin, face, cleanser, like, use, clean, makeup, product, cleansing, gentle, dry, love, used, acne, really, oily, gel, balm, time, remove, wash, great, moisturizer, work, outset, device, ingredient, make, way, away, facial, tried, took (рис. 3.8).

Відгуки у цій темі зосереджені на продуктах для очищення та ексфолюації шкіри. Користувачі обговорюють різні аспекти, такі як формула, упаковка,







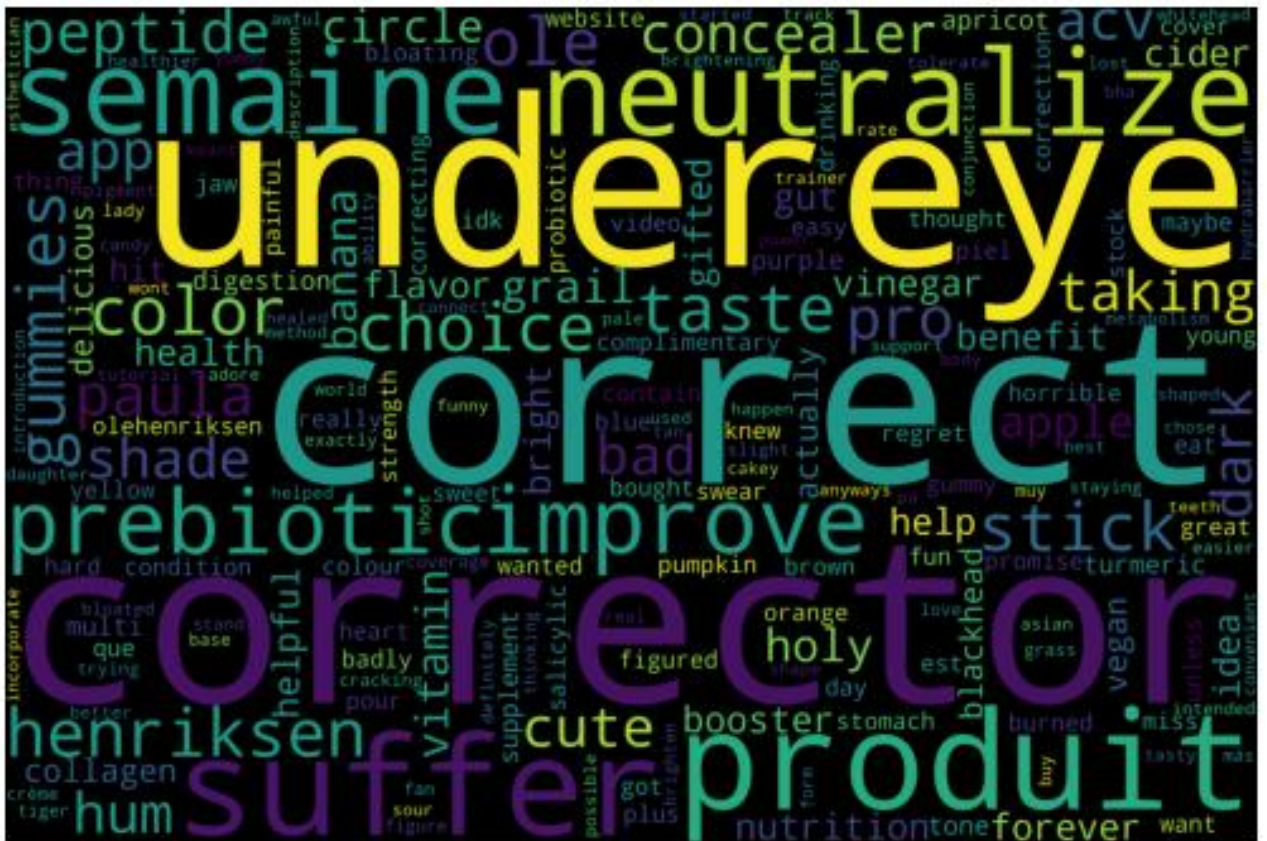


Рис. 3.11. Хмара слів для четвертої теми

Джерело: створено автором у середовищі Python

**Topic #5:** moisturizer, kieh1, moisturize, jart, ceramide, provide, warm, soothe, dissappoint, toner, soft, skincare, perfect, shydrating, leaf, moisturem, hydratioin, barrier, strengthen, penetrate, moisturizedskin, feeling, winter, smooth, super, amazing, plump (рис. 3.12).

Остання тема обговорює засоби для зволоження, їх текстуру, здатність покращувати бар'єрну функцію шкіри.

Загалом, проаналізувавши, про що говорять користувачі різних типів продукції, можна зрозуміти напрямки покращення продукту, аби повністю задовольнити їхню потребу в ідеальних засобах. Так, згідно з очікуваннями покупців, окрім своїх базових функцій, засоби для очищення шкіри мають мати ніжну формулу, приємний аромат, не подразнювати чутливу шкіру; засоби від сонця та для освітлення шкіри – бути легкими у текстурі, не лишати відчуття жирності після нанесення; антивікові продукти купуються переважно за



1. Використання словосполучень "skincare routine" і "skin care" свідчить про важливість продуктів для щоденного догляду. Хорошою ідеєю було б персоналізувати процес формування рекомендацій наборів продуктів для кожного покупця, виходячи з його типу шкіри та вже наявного щоденного догляду або ж досвіду. Для цього можна було б створити онлайн-опитування, бота-консультанта чи тест на визначення ідеального набору для рутини догляду за шкірою.

2. Біграми "honest review" і "highly recommend" вказують на важливість щирих відгуків. Компанія може залучити інфлуенсерів, які користуються довірою аудиторії для чесних оглядів продуктів Sephora. Можна навіть створити серію відео, в яких користувачі з різними потребами, типами шкіри та ретиною діляться власним досвідом.

3. Чимало слів вказувало на бажання користувачів мати інноваційні продукти за новітніми формулами та технологіями, тож було б доцільно вказати це в рекламних матеріалах.

4. Коистувачів також хвилювало пакування, зовнішній вигляд, упаковка та подача. Можна долучитися до еко-трендів, зробити пакування не лише сучасним і привабливим, а ще й екологічно чистим, також висвітлювати це під час просування.

5. Часто покупці обговорювали спосіб застосування продуктів, наголошуючи на важливості їхнього правильного використання. Аби цей аспект зробити більш прозорим можна організувати серію онлайн-вебінарів чи майстер-класів з професійними косметологами, візажистами, які б показали крок за кроком, як правильно користуватися засобами Sephora, надали б загальні рекомендації. Або ж компанія може завести профіль у соцмережах, таких як TikTok чи Instagram та знімати короткі відео-рекомендації, наприклад, "Як правильно використовувати зволожуючу маску" або "Ранкова рутинна догляду за шкірою обличчя" тощо.

6. Отримавши позитивний досвід, користувач частіше схильний не поділитися фідбеком від замовлення, ніж у разі, коли щось йде не так. При цьому Sephora має переважно хороші відгуки, які ще й позитивно забарвленні (часто

використовуються слова "favorite", "nice", "amazing", "love", "favorite" і такі інші). Потрібно стимулювати покупців лишати фідбек, наприклад, пропонуючи знижку 10% на наступну покупку за відгук про продукт, аби постійно мати релевантні відгуки.

7. Більшість рев'юерів мають проблемну шкіру ("dry skin", "sensitive skin"), тож доцільно було б розширити асортимент продуктів з активними компонентами, які допомагають користувачам з сухістю, чутливістю шкіри, висипаннями, темними колами, пігментацією тощо. Просувати ці продукти та рекомендації окремим таргет-групам.

8. Слова "smooth", "perfect", "hydrated" вказують на бажання користувачів досягти ідеального стану шкіри. Це не лише про зовнішній вигляд, а ще й про самопочуття, здоров'я, любов до себе, тож можна подумати про партнерства з фітнес-студіями та СПА-салонами, де продукти Sephora будуть інтегровані в процедури, або ж еко-лавками чи нутриціологами. Також запропонувати знижки на такі послуги, і, навпаки, надати партнерську знижку від закладів, з якими співпрацюватиме компанія.

9. Слогани для рекламних кампаній можуть бути створені з використанням аналізу найбільш часто вживаних слів користувачів. Наприклад, «Чутлива шкіра? Ми знаємо, як про неї подбати. Sephora – турбота без компромісів», «Секрет вашої ідеальної шкіри – зволоження та SPF від Sephora», «Сяйте яскравіше з новими продуктами Sephora», «Весняний догляд з Sephora – розкрийте природне сяйво вашої шкіри» - усі ці слогани враховують ключові слова, які найчастіше зустрічалися у позитивних відгуках ("гладенька", "сяюча", "догляд", "чутлива шкіра", "зволоження", "захист", "ефект", "новий"), таким чином підкреслюють потенційний позитивний досвід та акцентують на перевагах продукції, важливих для користувачів.

### **3.4. Упровадження та напрями удосконалення програмного засобу**

Запропонований програмний засіб буде цікавим, в першу чергу, компаніям, що функціонують у сфері е-комерції, маркетологам та менеджерам роботи з

клієнтами, так як її функціонал дозволяє проаналізувати великі обсяги текстових даних відгуків користувачів та отримати краще розуміння тенденції відгуків, рівня задоволеності клієнтів, ключових тем, що обговорюються, виявити негативні та позитивні настрої, що може допомогти бізнесу покращити продукти, послуги та стратегії взаємодії з клієнтами.

Для коректної роботи програми та забезпечення можливості аналізу відгуків користувачів з використанням сучасних методів обробки та візуалізації даних необхідними є наступні компоненти:

- середовище Python (версія 3.6 або вище);
- встановлення пакетів Flask, Pandas, Matplotlib і Seaborn, NLTK, SpaCy, Scikit-learn, Transformers, Wordcloud, Genism, Collections (встановлюються в терміналі чи командному рядку командою `pip install [назва пакету]`);
- обчислювальні потужності - комп'ютер із сучасним процесором та достатньою оперативною пам'яттю (рекомендується 8GB RAM або більше) для обробки великих обсягів даних;
- база даних з відгуками у форматі CSV (обов'язково мають бути колонки з `submission_time`, `author_id`, `review_text`, `rating`);
- папка з файлами програми, вона має структуру, представлену на рис. 3.13 (код файлів можна переглянути у додатках Б-Д);

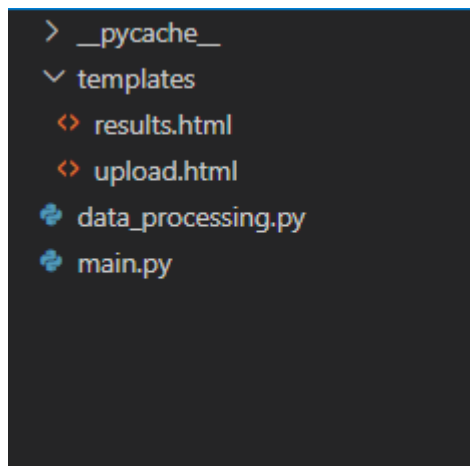


Рис. 3.13. Структура файлів програми

Джерело: створено автором у середовищі Python

- запуск в терміналі файлу main.py проекту, після цього Flask підготує локальний сервер та фронтенд-частину, посилання потрібно відкрити у вкладці браузера (рис. 3.14), після чого програма буде готовою до роботи – відкриється перший її екран з проханням обрати файл (рис. 3.15);

```

snizh@LAPTOP-OMOHIA5A MINGW64 ~/OneDrive/... (master)
$ python -u "c:\Users\snizh\OneDrive\...\main.py"
[nltk_data] Downloading package vader_lexicon to
[nltk_data] C:\Users\snizh\AppData\Roaming\nltk_data...
[nltk_data] Package vader_lexicon is already up-to-date!
[nltk_data] Downloading package punkt to
[nltk_data] C:\Users\snizh\AppData\Roaming\nltk_data...
[nltk_data] Package punkt is already up-to-date!
[nltk_data] Downloading package stopwords to
[nltk_data] C:\Users\snizh\AppData\Roaming\nltk_data...
[nltk_data] Package stopwords is already up-to-date!
[nltk_data] Downloading package wordnet to
[nltk_data] C:\Users\snizh\AppData\Roaming\nltk_data...
[nltk_data] Package wordnet is already up-to-date!
[nltk_data] Downloading package omw-1.4 to
[nltk_data] C:\Users\snizh\AppData\Roaming\nltk_data...
[nltk_data] Package omw-1.4 is already up-to-date!
* Serving Flask app 'main'
* Debug mode: on
WARNING: This is a development server. Do not use it in a production deployment. Use a production WSGI server instead.
* Running on http://127.0.0.1:5000
Press CTRL+C to quit

```

Рис. 3.14. Запуск локального сервера

Джерело: створено автором у середовищі Python

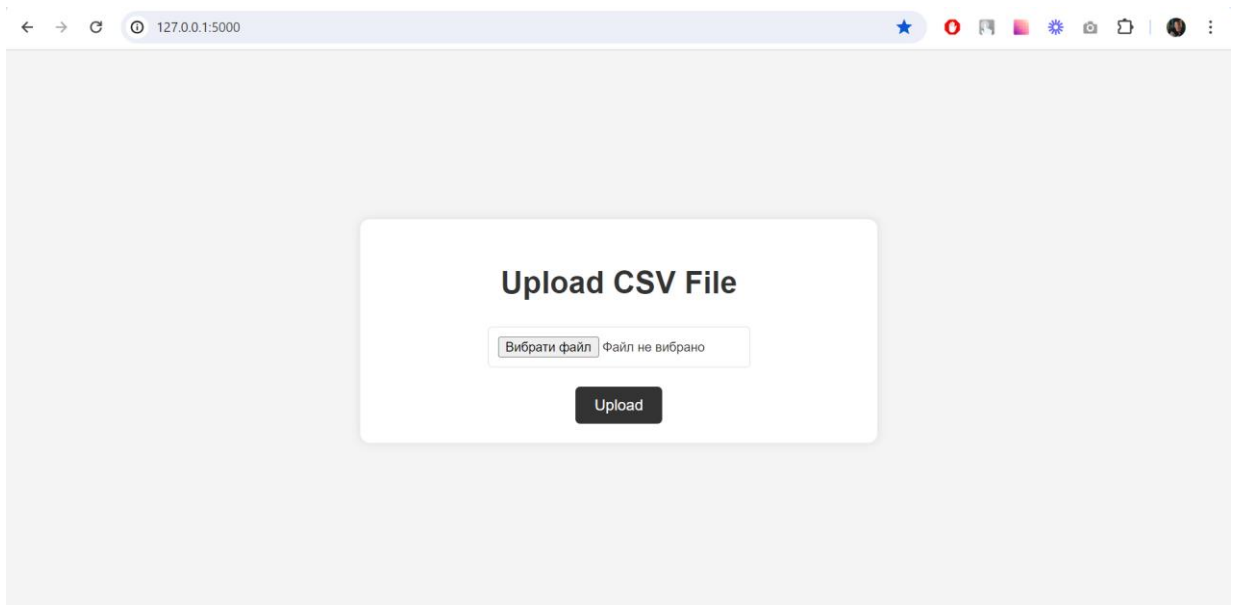


Рис. 3.15. Перший екран програми

Джерело: створено автором у середовищі Python

- коли фахівець обере файл та натисне кнопку “Upload” програма почне аналіз, по завершенні обробки даних, користувача перекине на сторінку

результатів (рис. 3.16) з усіма візуалізаціями для подальшого огляду та аналізу, пошуку інсайтів;

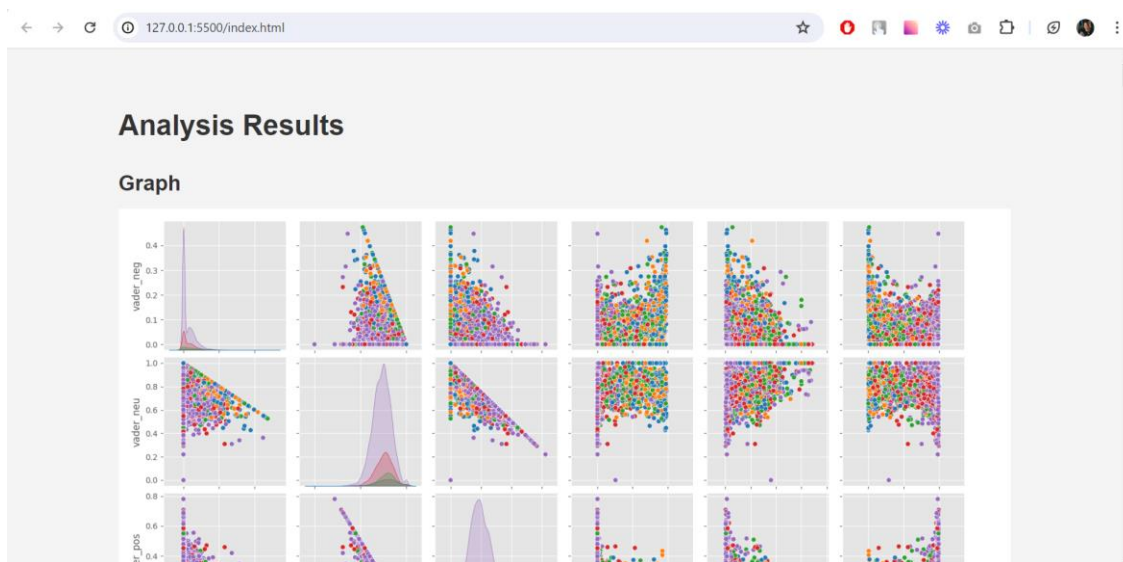


Рис. 3.16. Перший екран результатів роботи програми

Джерело: створено автором у середовищі Python

- якщо виникне потреба зберегти сторінку результатів, це можна зробити, викликавши контекстне меню та обравши опцію «Зберегти як» (рис. 3.17), у відкні «Збереження файлу» обрати необхідний шлях та натиснути «Зберегти» (рис. 3.18). Збережений в html форматі файл буде можна буде відкрити у браузері для повторного перегляду.

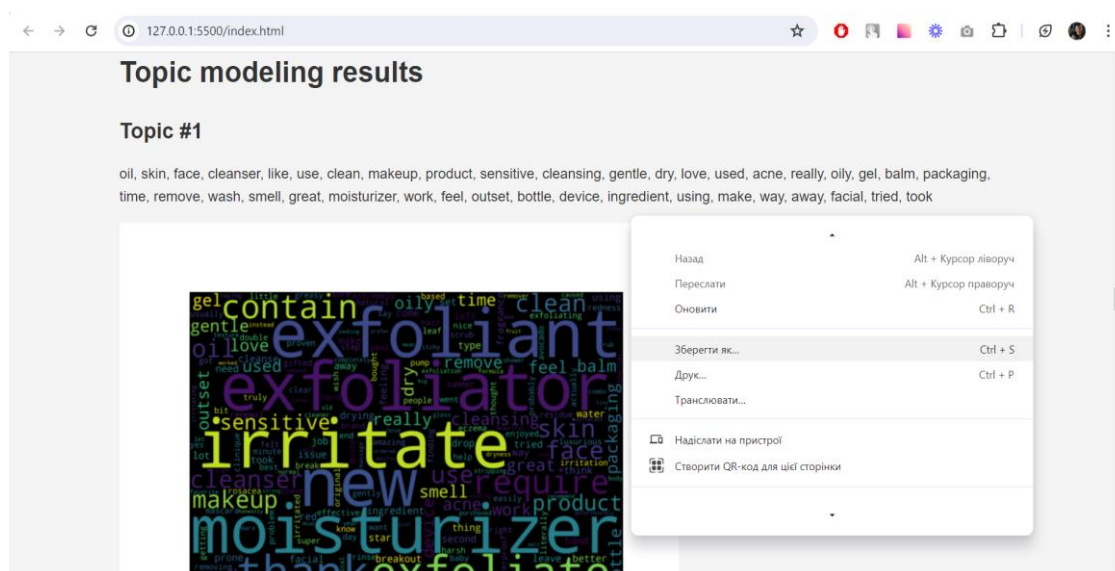


Рис. 3.17. Контекстне меню на сторінці результатів роботи програми

Джерело: створено автором у середовищі Python

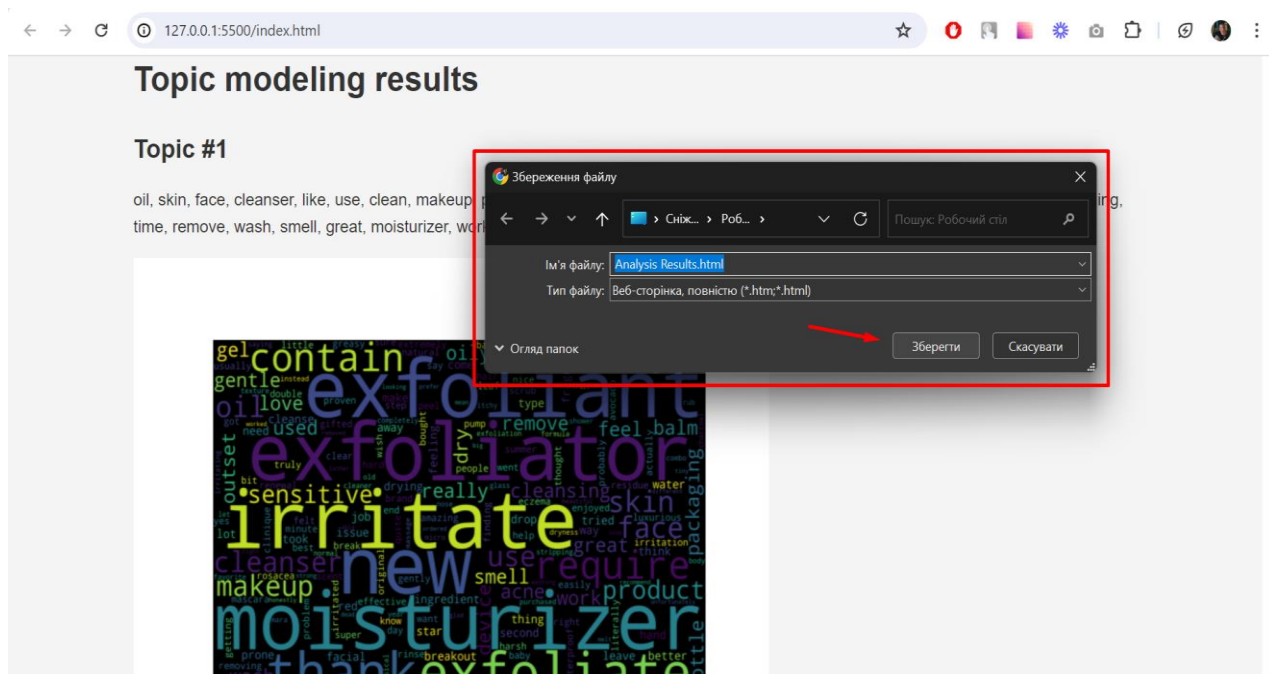


Рис. 3.18. Збереження сторінки результатів в html форматі

Джерело: створено автором у середовищі Python

Загалом програма є доволі простою у використанні для фахівця, який має хоча б мінімальний досвід роботи з Python. Проте на етапі впровадження доцільно провести кілька вебінарів з командою аналітиків та маркетологів, на яких продемонструвати роботу програми, відповісти на питання, які виникнуть, та обговорити подальші можливості підтримки.

Розробка програми та її підтримка є процесом неперервного удосконалення, аби вона відповідала поточним потребам користувачів і вирішувала актуальні завдання бізнесу. Нижче представлено ключові напрями покращень та точки росту програми з аналізу відгуків користувачів, які сприятимуть поліпшенню її функціональності, ефективності та користувацького досвіду.

- додавання мультимовності, що дозволить використовувати програму для сторінок, які функціонують на різних маркетсах;
- додавання інтерактивності візуалізацій;
- додаткові налаштування параметрів аналізу для користувача перед початком роботи програми;

- налаштування програми для автоматичного завантаження нових відгуків з різних джерел (наприклад, через API магазинів) та регулярного оновлення аналітичних звітів;
- інтеграція з системами управління клієнтами (CRM), маркетинговими платформами та інструментами бізнес-аналітики для покращення використання результатів аналізу;
- додавання можливості деталізації та фільтрації результатів за різними параметрами (наприклад, за датою, продуктом, категорією);
- розгортання програми на хмарних платформах для забезпечення масштабованості та доступності.

Ці напрями розвитку враховують потреби користувачів та вимоги ринку, а також дозволять програмі стати більш потужною, гнучкою та конкурентоспроможною в майбутньому.

З поточної версії програми маркетолог отримає широкий спектр вихідних даних, що включає результати аналізу відгуків користувачів, графіки тенденцій, рейтинги та розуміння настроїв. Ці дані можуть стати основою для стратегічного планування, розробки та вдосконалення маркетингових стратегій та комунікацій, спрямованих на поліпшення відносин з клієнтами та підвищення їхньої лояльності. Крім того, вони є важливим інструментом для контент-маркетингу, адже можуть допомогти у створенні матеріалів, які відповідатимуть потребам клієнтів.

Також, ці дані можуть бути корисними для відстеження негативних відгуків та забезпечення відповідної реакції на них, що дозволить підвищити рівень задоволення клієнтів та зберегти позитивну репутацію бренду.

Крім того, зібрані дані можуть бути використані для створення презентацій та звітів, які будуть представлені керівництву компанії, демонструючи маркетингові рішення щодо подальшої діяльності, стратегій, обґрунтовані результатами аналізу, та рекомендації щодо наступних дій.

## ВИСНОВКИ

У кваліфікаційній роботі було здійснено дослідження процесу підтримки прийняття маркетингових рішень у сфері електронної комерції з використанням технологій штучного інтелекту та машинного навчання.

В ході дослідження було виконано поставлені завдання та отримано наступні теоретичні та практичні висновки:

1. У цифровому середовищі ключовими аспектами функціонування е-комерції та е-маркетингу є мобільна доступність, використання штучного інтелекту для оптимізації процесів та персоналізації, трансформація маркетингових стратегій з акцентом на користувацький досвід та інтеграція між е-комерції та е-маркетингом, що сприяє створенню комплексних маркетинг-моделей для ефективного управління бізнесом та підвищення його результативності.

2. Аналіз компонентів маркетингового міксу в контексті електронної комерції показав, що компоненти Продукт, Ціна, Місце та Просування функціонують згідно зі стандартними принципами, проте зазнають модифікаційних впливів сфери електронної торгівлі. У маркетинговій концепції бізнесу, що функціонує в сфері е-комерції, продукт охоплює також електронні товари та послуги, формування ціни стає гнучкішим та більш персоналізованим, зокрема за допомогою аналізу Big Data та динамічних стратегій ціноутворення. Місцем продажу стають віртуальні платформи, а просування включає діджитал рекламу, інтелектуальні боти та встановлення зв'язків з громадськістю через соціальні мережі.

3. Штучний інтелект та технології обробки природної мови (NLP) зокрема відіграють ключову роль у трансформації маркетингових стратегій на платформах електронної комерції, забезпечуючи автоматизацію обслуговування клієнтів, персоналізацію рекомендацій, оптимізацію ціноутворення та ефективне позиціонування бренду. Використання ШІ підвищує конкурентоспроможність

бізнесу, знижує витрати на робочу силу, покращує користувацький досвід і є основою для прийняття точніших і релевантніших маркетингових рішень.

4. На етапі маркетингових досліджень ШІ використовується для збору даних, аналізу ринку, розуміння потреб клієнтів; на етапі стратегічного планування допомагає з сегментацією, таргетуванням, позиціонуванням; на етапі маркетингових дій покриває процеси персоналізації, стандартизації та взаємодії з клієнтами. При цьому задіюється механічний, мислячий та емоційний ШІ для вирішення концептуальних завдань на кожному з етапів.

5. Штучний інтелект (ШІ) пропонує широкий спектр методів для підтримки маркетингових рішень в електронній комерції. Для аналізу даних може застосовуватися розпізнавання візуальної інформації, фрод-моніторинг, предиктивна аналітика, аналіз емоційних даних користувачів; до процесів формування стратегії і безпосередньо маркетингових дій ШІ можуть бути залучені в контексті генерації контенту, гіперперсоналізації, оптимізації кампаній тощо.

6. Технології машинного навчання надають потужну базу для проведення маркетингових досліджень та ефективної обробки великих обсягів структурованих та неструктурованих даних зокрема, що дозволяє виявляти в них латентні зв'язки та патерни, а також спостерігати зміни в динаміці. При цьому найбільшу цінність інструменти мають для персоналізації досвіду користувача, що робить функціонування бізнесу конкурентноспроможним на ринку.

7. Аналіз відгуків користувачів є критично важливим для бізнесу, оскільки надає можливість виявляти дефекти, які не були помічені внутрішнім тестуванням, отримувати ідеї для вдосконалення продуктів і послуг, формувати лояльність клієнтів і підтримувати конкурентоспроможність на ринку. Реагуючи на відгуки, компанія демонструє готовність сприймати критику і вести відкриту комунікацію, що зміцнює довіру до бренду. Інсайти з відгуків допомагають маркетологам створювати релевантні меседжі, підтверджувати або

спростовувати маркетингові гіпотези та будувати стратегії на основі реальних даних, забезпечуючи бізнесу стійку перевагу на ринку. Основні цілі практичного дослідження включають автоматизацію обробки великого обсягу текстових даних відгуків, глибше розуміння споживачів і їхніх потреб, а також покращення якості обслуговування клієнтів бізнесу, що функціонує у сфері е-комерції.

8. Етапи реалізації аналізу відгуків користувачів з використанням запропонованого інструменту забезпечать системний підхід до обробки та аналізу даних. На першому етапі здійснюється завантаження датасету з відгуками, що дозволяє швидко ввести в обробку великі обсяги текстових даних. На другому етапі, включаючи фази очищення, перетворення, конструювання ознак, інтеграції та форматування даних, відбувається підготовка інформації для подальшого аналізу, що підвищує якість результатів. Третій етап передбачає застосування сентимент-аналізу та тематичного моделювання, що дозволяє виявляти ключові тенденції, настрої та теми у відгуках користувачів. Четвертий етап оцінювання моделей гарантує їхню ефективність та точність, а на п'ятому етапі інтерпретація результатів надає можливість отримувати цінні інсайти, які можна використати для прийняття обґрунтованих бізнес-рішень щодо покращення продуктів та послуг.

9. Було проведено реалізацію можливого сценарію за допомогою запропонованого програмного засобу на основі датасету магазину косметики Sephora. Аналіз відгуків користувачів показав, що основні очікування клієнтів щодо різних типів продукції включають м'які формули та приємний аромат для засобів очищення шкіри, легку текстуру для сонцезахисних засобів, високий вміст активних компонентів у антивікових продуктах тощо. Використовуючи ці результати, компанія може розробити нові формули чи вдосконалити існуючі продукти, створити маркетингові кампанії, що підкреслюють важливі для споживачів переваги, поліпшити комунікацію з клієнтами, розробити персоналізовані пропозиції та спеціальні промо-акції. Наприклад, важливість

щирих відгуків можна підсилити через співпрацю з інфлюенсерами, а екологічність упаковки – шляхом використання еко-френдлі матеріалів у просуванні бренду. Онлайн-опитування та боти-консультанти можуть допомогти персоналізувати рекомендації, а серії вебінарів або короткі навчальні відео продемонструвати правильне використання продуктів, що виявилось важливим для покупців. Таким чином, Sephora може сприяти покращенню клієнтського досвіду, підвищенню лояльності клієнтів та своєї конкурентоспроможності на ринку.

10. У ході практичного дослідження було розроблено програмний засіб у середовищі Python, яка використовуючи методи сентимент-аналізу та тематичного моделювання, дозволяє аналізувати великі обсяги текстових даних користувацьких відгуків. Після роботи програми формується веб-сторінка результатів з візуалізаціями, на основі якої маркетологи та аналітики можуть зробити висновки про загальні тенденції, настрої користувачів, ідентифікувати основні проблеми, з якими стикаються покупці, та переваги, що є сильними сторонами бренду чи продукту. Ці матеріали можуть використовуватися для подальших звітів, стратегій, презентацій, забезпечуючи data-driven підхід у прийнятті маркетингових рішень.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Schneide G. Electronic Commerce. // Course Technology. – 2008. URL: <https://erkashif.wordpress.com/wp-content/uploads/2015/12/ecom-schneider.pdf>
2. Мельник О. В. Електронна комерція як складова частина електронного бізнесу. // Третя українська Інтернет-конференція «Сучасність, наука, час. Взаємодія та взаємовплив». URL: <http://intkont.org>
3. Hashemi-Pour C. What is e-commerce? / Hashemi-Pour C., Lutkevich B. – 2023. URL: <https://www.techtarget.com/searchcio/definition/e-commerce>
4. Jain I. An Overview of Electronic Commerce (e-Commerce). / Jain I., Malviya B., Arya S. // Society of Business and Management, – 2021. URL: [https://www.researchgate.net/publication/351775073\\_An\\_Overview\\_of\\_Electronic\\_Commerce\\_e-Commerce](https://www.researchgate.net/publication/351775073_An_Overview_of_Electronic_Commerce_e-Commerce)
5. Bloomenthal A. Mobile Commerce: Definition, Benefits, Examples, and Trends. – 2022. URL: <https://www.investopedia.com/terms/m/mobile-commerce.asp>
6. Snyder K. 35 E-Commerce Statistics of 2024. // Forbes Advisor. – 2024. URL: <https://www.forbes.com/advisor/business/ecommerce-statistics/>
7. Precedence Research. Artificial Intelligence in E-commerce Market. – 2024. URL: <https://www.precedenceresearch.com/artificial-intelligence-in-e-commerce-market>
8. Паталяк Є. Як ШІ змінює електронну комерцію: 8 прикладів. – 2024. URL: <https://wezom.com.ua/ua/blog/yak-shi-zminyuje-elektronnu-komertsiyu-8-prikladiv>
9. Litmus. Email Marketing ROI: What leads to better returns? URL: <https://www.litmus.com/resources/email-marketing-roi>
10. Suyanto A. M. A. Marketing mix on purchase intention and its impact on the decision to purchase something products. / Suyanto A. M. A., Dewi D.G. – 2023.

URL:

[https://www.researchgate.net/publication/374537381\\_Marketing\\_Mix\\_on\\_Purchase\\_Intention\\_and\\_its\\_Impact\\_on\\_the\\_Decision\\_to\\_Purchase\\_Somethinc\\_Products](https://www.researchgate.net/publication/374537381_Marketing_Mix_on_Purchase_Intention_and_its_Impact_on_the_Decision_to_Purchase_Somethinc_Products)

11. Rueter Th. The price is right – then it's not. – 2014. URL: <https://www.internetretailer.com/2014/08/04/price-rightthen-its-not>
12. Nielsen. Still recommended by friends and relatives the most authentic advertising according to consumers. The most trusted on brand websites. – 2015. URL: <https://www.nielsen.com/news-center/2015/still-recommended-by-friends-and-relatives-the-most-authentic-advertising-according-to-consumers-the-most-trusted-on-brand-websites/>
13. Curiosity at Work. Colette Des Georges 25 stats about AI in customer experience that show how consumers really feel. URL: <https://www.surveymonkey.com/curiosity/25-stats-about-ai-in-customer-experience-that-show-how-consumers-really-feel/>
14. Huang M. H. A strategic framework for artificial intelligence in marketing. / Huang M. H., Rust R. T. – 2020. URL: <https://link.springer.com/article/10.1007/s11747-020-00749-9#Tab3>
15. Huang M. H. Artificial intelligence in service. / Huang M. H., Rust R. T. // Journal of Service Research. – 2018. URL: <https://journals.sagepub.com/doi/full/10.1177/1094670517752459>
16. Huang M. H. The feeling economy: Managing in the next generation of artificial intelligence (AI). / Huang M. H., Rust R. T., Maksimovic V. // California Management Review. – 2019. URL: <https://journals.sagepub.com/doi/full/10.1177/0008125619863436>
17. McDuff D. Designing emotionally sentient agents. Communications of the ACM. / McDuff D., Czerwinski M. – 2018. URL: <https://dl.acm.org/doi/abs/10.1145/3186591>

18. Huang M. H. Technology-driven service strategy. / Huang M. H., Rust R. T. // Journal of the Academy of Marketing Science. – 2017. URL: <https://link.springer.com/article/10.1007/s11747-017-0545-6>
19. Netzer O. When words sweat: Identifying signals for loan default in the text of loan applications. / Netzer O., Lemaire A., Herzenstein M. // Journal of Marketing Research. – 2019. URL: <https://journals.sagepub.com/doi/full/10.1177/0022243719852959>
20. Pitt C. S. New approaches to psychographic consumer segmentation: Exploring fine art collectors using artificial intelligence, automated text analysis and correspondence analysis. / Pitt C. S., Bal A. S., Plangger K. // European Journal of Marketing. – 2020. URL: <https://doi.org/10.1108/EJM-01-2019-0083>
21. Valls A. Using ontology-based clustering to understand the push and pull factors for British tourists visiting a Mediterranean coastal destination. / Valls A., Gibert K., Orellana A., Anton-Clave S. // Information & Management. – 2018. URL:
22. Dekimpe M. Retailing and retailing research in the age of big data analytics. International Journal of Research in Marketing. – 2020. URL: <https://www.sciencedirect.com/science/article/pii/S016781161930062X>
23. Chornous G. Optimization of Portfolio Marketing Strategy Based on Data Science Technologies. / Chornous G., Farenjuk Y. // Selected Papers of the VIII International Scientific Conference “Information Technology and Implementation” (IT&I-2021). Conference Proceedings, Kyiv, Ukraine, December 01-03, 2021. – CEUR Workshop Proceedings. – Vol. 3179. – P. 89-100. (Scopus)
24. Chornous G. Optimization of Marketing Decisions Based on Machine Learning: Case for Telecommunications. / Chornous G., Farenjuk Y. // Selected Papers of the IX International Scientific Conference "Information Technology and Implementation" (IT&I-2022). Conference Proceedings, Kyiv, Ukraine,

- November 30 - December 02, 2022. – CEUR Workshop Proceedings. – Vol. 3347. – P. 112-124. (Scopus)
25. Chornous G. Marketing Mix Modeling of Traffic to the Store Under the Covid-19 Crisis. / Chornous G., Farenjuk Y. // KnowledgeE Social Sciences. Economies of the Balkan and Eastern European Countries (EBEEC). – 2023. – Vol.8. – P. 412–434. (Закордонне видання, ОАЕ)
26. Chornous G. Principles of effective planning of advertising activity in the 21st century. / Farenjuk Y., Chornous G. // Economy of Ukraine. – 2023. – No. 2. – P. 55-72. (Фахове видання)
27. Chornous G. Optimization of Media Strategy via Marketing Mix Modeling in Retailing. / Farenjuk Y., Chornous G. // Ekonomica (Economics). – 2023. – Vol. 102(1). – P. 6-25. (Scopus) (Закордонне видання, Литва)
28. Chornous G. A data science-based marketing decision support system for brand management. / Chornous G., Farenjuk Y., Giedraitis V.R., Ulvidienė E., Kharlamova G. // Innovative Marketing. – 2023. – Vol. 19(2). – P. 38-50. (Scopus, Web of Science)
29. Chornous G. Data Mining for Economists: Course book. / Chornous G., Farenjuk Y., Didenko I. // Kyiv: Publishing Lira-K. – 2023. – 290 p.
30. Charbuty B. Classification based on decision tree algorithm for machine learning. / Charbuty B., Abdulazeez A. – 2021. URL: <https://jastt.org/index.php/jasttpath/article/view/65>
31. Speiser J.L. A comparison of random forest variable selection methods for classification prediction modeling. / Speiser J.L., Miller M.E., Tooze J., Ip E. – 2019. <https://www.sciencedirect.com/science/article/abs/pii/S0957417419303574?via%3Dihub>
32. Chen S. A novel selective naïve Bayes algorithm. / Chen S., Webb G.I., Liu L., Ma X. – 2020. URL:

- <https://www.sciencedirect.com/science/article/abs/pii/S0950705119306185?via%3Dihub>
33. Greenacre M. Principal component analysis. / Greenacre M., Groenen P. J. F., Hastie T., Iodice d'Enza A., Markos A., Tuzhilina E. – 2022. URL: [https://www.researchgate.net/publication/366501387\\_Principal\\_component\\_analysis](https://www.researchgate.net/publication/366501387_Principal_component_analysis)
34. Wang Y.X. Nonnegative matrix factorization: A comprehensive review. / Wang Y.X., Zhang Y.J. – 2012. URL: <https://ieeexplore.ieee.org/document/6165290>
35. Isnain A.R. Implementation of K-Nearest Neighbor (K-NN) Algorithm for Public Sentiment Analysis of Online Learning. / Isnain A.R., Supriyanto J., Kharisma M.P. – 2021. URL: <https://jurnal.ugm.ac.id/ijccs/article/view/65176>
36. Abiodun O.I. Comprehensive review of artificial neural network applications to pattern recognition. / Abiodun O.I., Jantan A., Omolara A.E., Dada K.V., Umar A.M., Linus O.U., Arshad H., Kazaure A.A., Gana U., Kiru M.U. – 2019. URL: <https://ieeexplore.ieee.org/document/8859190/>
37. Kattenborz T. Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. / Kattenborz T., Leitloff J., Schiefer F., Hinz S. – 2021. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0924271620303488?via%3Dihub>
38. Sherstinsky A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. – 2020. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0167278919305974?via%3Dihub>
39. Arulkumaran K. Deep reinforcement learning: A brief survey. / Arulkumaran K., Deisenroth M.P., Brundage M., Bharath A.A. – 2017. URL: <https://ieeexplore.ieee.org/document/8103164/>

40. Creswell A. Generative adversarial networks: an overview. / Creswell A., White T., Dumoulin V., Arulkumaran K., Sengupta B., Bharath A.A. – 2018. URL: <https://ieeexplore.ieee.org/document/8253599/>
41. Devlin J. Bert: Pre-training of deep bidirectional transformers for language understanding. / Devlin J., Chang M.W., Lee K., Toutanova K. – 2018. URL: <https://aclanthology.org/N19-1423/>
42. Alparthi S. Bidirectional Encoder Representations from Transformers (BERT): A sentiment analysis odyssey. / Alparthi S., Mishra M. – 2020. URL: <https://arxiv.org/pdf/2007.01127>
43. Jelodar H. Latent Dirichlet Allocation (LDA) and Topic modeling: models, applications, a survey. / Jelodar H., Wang Y., Yuan C., Feng X., Jiang X., Li Y., Zhao L. – 2017. URL: [https://www.researchgate.net/publication/321069759\\_Latent\\_Dirichlet\\_Allocation\\_LDA\\_and\\_Topic\\_modeling\\_models\\_applications\\_a\\_survey](https://www.researchgate.net/publication/321069759_Latent_Dirichlet_Allocation_LDA_and_Topic_modeling_models_applications_a_survey)
44. Geetha L. Vader: A Comprehensive Guide to Sentiment Analysis in Python. – 2023. URL: <https://medium.com/@rslavanyageetha/vader-a-comprehensive-guide-to-sentiment-analysis-in-python-c4f1868b0d2e>
45. Liu Y. RoBERTa: A Robustly Optimized BERT Pretraining Approach. / Liu Y., Ott M., Goyal N., Du J., Joshi M., Chen D., Levy O., Lewis M., Zettlemoyer L., Stoyanov V. – 2019. <https://arxiv.org/pdf/1907.11692>
46. Могиль С. Використання технологій Data Science та Artificial Intelligence для оптимізації маркетингових стратегій. / Г.Чорноус, С.Могиль // Матеріали Міжнародної науково-практичної конференції Data Science та інформаційно-аналітичні системи: застосування в економіці та фінансах. – 2024. – С. 69-72.

## ДОДАТКИ

Додаток А

## Зведена таблиця описової статистики та якості початкових даних

Field	Valid	Valid, %	Mismatched	Mismatched, %	Missing	Missing, %	Mean	Std. Deviation	Min	25%	50%	75%	Max	Unique	Most Common, %
id	119000	100	0	0	0	0	59700	34400	0	29800	59700	89500	119000		
author_id	119000	100	10	0	0	0									
rating	119000	100	0	0	0	0	4,35	1,11	1	4	5	5	5		
is_recommended	110104	93	0	0	8896	7	0,86	0,35	0	1	1	1	1		
helpfulness	67300	56	0	0	52	44	0,76	0,32	0	0,63	0,91	1	1		
total_feedback_count	119000	100	0	0	0	0	4,9	15,4	0	0	1	5	1135		
total_neg_feedback_count	119000	100	0	0	0	0	1,12	5,96	0	0	1	1	1094		
total_pos_feedback_count	119000	100	0	0	0	0	3,78	12,2	0	0	1	3	743		
submission_time	119000	100	0	0	0	0			29Aug08				21Mar23		
review_text	119000	100	0	0	89	0								115970	0
review_title	83800	70	0	0	35600	30								53900	1
skin_tone	104000	87	0	0	15200	13								14	26
eye_color	105000	88	0	0	14700	12								6	48
skin_type	112000	94	0	0	7605	6								4	52
hair_color	98800	83	0	0	20600	17								7	36
product_id	119000	100	0	0	0	0								499	0
product_name	119000	100	0	0	0	0								498	0
brand_name	119000	100	0	0	0	0								107	5
price_usd	119000	100	0	0	0	0	53,7	43,2	3	30	44	68	410		

Джерело: сформовано автором на основі відомостей про БД "Sephora Skincare Reviews" з ресурсу Kaggle

## Вміст файлу main.py

```

import base64
import pandas as pd
from flask import Flask, request, render_template
from data_processing import get_common_words_plot, get_most_frequent_common_words,
topic_modeling, get_distribution_of_customer_rating_plot,
get_common_words_around_target_word, get_avarage_weekly_sentiment_score_plot,
extract_bi_gram, get_clean_text,
get_sentiment_scores, get_average_weekly_rating_plot, get_reviews_per_week_plot,
get_pairplot, get_reviews_by_sentiment

def b64encode(data):
    """
    Custom Jinja2 filter for base64 encoding.
    """
    return base64.b64encode(data).decode('utf-8')

app = Flask(__name__)
app.jinja_env.filters['b64encode'] = b64encode

@app.route('/', methods=['GET', 'POST'])

def upload_file():
    if request.method == 'POST':
        file = request.files['file']

        if file:
            df = pd.read_csv(file)

            df['submission_time'] = pd.to_datetime(df['submission_time'])
            df = df.sort_values(by='submission_time', ascending=False)
            df = df.set_index('submission_time')
            df['review_text'] = df['review_text'].fillna('').astype(str)
            df_2023 = df.loc['2023']

            qty_plot = get_reviews_per_week_plot(df_2023)
            weekly_sentiment_score_plot =
get_avarage_weekly_sentiment_score_plot(df_2023)
            average_rating_plot = get_average_weekly_rating_plot(df_2023)
            distribution_of_customer_rating_plot =
get_distribution_of_customer_rating_plot(df_2023)

            cleaned_text = get_clean_text(df_2023)
            most_frequent_common_words = get_most_frequent_common_words(cleaned_text, 3)
            print(most_frequent_common_words)
            common_words_plot = get_common_words_plot(cleaned_text)
            common_words_1 = get_common_words_around_target_word(df_2023,
most_frequent_common_words[0]['word'])
            common_words_2 = get_common_words_around_target_word(df_2023,
most_frequent_common_words[1]['word'])
            common_words_3 = get_common_words_around_target_word(df_2023,
most_frequent_common_words[2]['word'])
            bi_dict = extract_bi_gram(cleaned_text)
            topec_modeling_results = topic_modeling(cleaned_text)

```

```

results_df = get_sentiment_scores(df_2023)
plot_bytes = get_pairplot(results_df)

pos_reviews_roberta = get_reviews_by_sentiment(results_df, 1, 'roberta',
'pos')
pos_reviews_vader = get_reviews_by_sentiment(results_df, 1, 'vader', 'pos')
neg_reviews_roberta = get_reviews_by_sentiment(results_df, 5, 'roberta',
'neg')
neg_reviews_vader = get_reviews_by_sentiment(results_df, 5, 'vader', 'neg')

return render_template('results.html',
    plot=plot_bytes,
    bi_dict=bi_dict,
    topics_results=topec_modeling_results,
    most_frequent_common_words=most_frequent_common_words,
    common_words_around_1_result=common_words_1,
    common_words_around_2_result=common_words_2,
    common_words_around_3_result=common_words_3,
    weekly_sentiment_score_plot=weekly_sentiment_score_plot,
    distribution_of_customer_rating_plot=distribution_of_customer_rating_plot,
    common_words_plot=common_words_plot,
    average_rating_plot=average_rating_plot,
    pos_reviews_roberta=pos_reviews_roberta,
    pos_reviews_vader=pos_reviews_vader,
    neg_reviews_roberta=neg_reviews_roberta,
    neg_reviews_vader=neg_reviews_vader,
    qty_plot=qty_plot)

return render_template('upload.html')

if __name__ == '__main__':
    app.run(debug=True)

```

*Джерело: створено автором у середовищі Python*

## Вміст файлу data\_processing.py

```
import string
import nltk
import re
import io
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

from nltk.util import ngrams
from nltk.corpus import stopwords
from nltk.tokenize import sent_tokenize
from nltk.stem import WordNetLemmatizer
from nltk.sentiment import SentimentIntensityAnalyzer
from nltk.tokenize import regexp_tokenize

from sklearn.feature_extraction.text import CountVectorizer
from sklearn.decomposition import LatentDirichletAllocation

from transformers import AutoTokenizer, AutoModelForSequenceClassification
from scipy.special import softmax
from wordcloud import WordCloud
from gensim.models import Word2Vec
from collections import Counter

nltk.download('vader_lexicon')
nltk.download('punkt')
nltk.download('stopwords')
nltk.download('wordnet')
nltk.download('omw-1.4')

plt.style.use('ggplot')

def preprocess_text(text):
    """
    Performs text preprocessing, including tokenization, punctuation removal,
    stopword removal and lemmatization.

    Parameters:
    text (str): The text to preprocess.

    Returns:
    str: The preprocessed text as a string.
    """
    words = regexp_tokenize(text, pattern=r'\b\w+\b')
```

```

tagged_words = nltk.pos_tag(words)
processed_words = [word for word, pos in tagged_words if pos != 'DT']
tokens = [w for w in processed_words if w.lower() not in string.punctuation]
stop_words = stopwords.words('english')
tokens = [token for token in tokens if token not in stop_words]
tokens = [word for word in tokens if len(word)>=3]

lemmatizer = WordNetLemmatizer()
tokens = [lemmatizer.lemmatize(word) for word in tokens]

preprocessed_text = ' '.join(tokens)
return preprocessed_text

def get_clean_text(dataframe):
    """
    Cleans and preprocesses text in a DataFrame by replacing HTML entities,
    removing punctuation, stopwords and lemmatizing words.

    Parameters:
    dataframe (pandas.DataFrame): DataFrame containing text data to process.

    Returns:
    numpy.ndarray: Array of processed texts.
    """

    dataframe['clean_review'] = [str(i).replace("&", '').replace("'", '') for i in
dataframe['review_text']]
    dataframe['clean_review'] = dataframe['clean_review'].apply(preprocess_text)
    dataframe['clean_review'] = dataframe['clean_review'].str.lower()

    return dataframe['clean_review'].values

def save_plot_to_bytes_obj(plt):
    """
    Saves a plot to a bytes object in PNG format and returns its byte value.

    Parameters:
    plt (matplotlib.pyplot): The plot to save.

    Returns:
    bytes: The byte value of the saved plot in PNG format.
    """

    buf = io.BytesIO()
    plt.savefig(buf, format='png')
    buf.seek(0)
    plt.close()

    return buf.getvalue()

```

```

def get_common_words_around_target_word(dataframe, target_word):
    """
    Finds the most common words appearing around a given word
    in the texts of a DataFrame.

    Parameters:
    dataframe (pandas.DataFrame): DataFrame containing text data.
    target_word (str): The word around which to find the most common words.

    Returns:
    list: A list of dictionaries containing predicted words and their probabilities.
    """

    custom_stop_words = set(stopwords.words('english') + ['super', 'duper',
'product'])
    good_token_clean = []

    all_reviews = ' '.join(dataframe['review_text'])
    sentences = sent_tokenize(all_reviews)

    for sentence in sentences:
        eng_word = re.findall(r'[A-Za-z-]+', sentence)
        good_token_clean.append([i.lower() for i in eng_word if i.lower() not in
custom_stop_words])

    model = Word2Vec(sentences=good_token_clean, window=10, min_count=1, workers=4,
sg=0)
    predictions = model.predict_output_word([target_word], topn=10)

    result = [{'word': target_word, 'predicted_word': prediction[0], 'probability':
prediction[1]} for prediction in predictions]

    return result

def topic_modeling(reviews_text):
    """
    Performs topic modeling on a list of texts using LDA.

    Parameters:
    reviews_text (list): A list of text documents.

    Returns:
    list: A list of dictionaries containing topics and their associated graphs.
    """

    class LemmaCountVectorizer(CountVectorizer):
        def build_analyzer(self):
            """
            Builds an analyzer that lemmatizes tokens.
            """

```

```

        analyzer = super(LemmaCountVectorizer, self).build_analyzer()
        lemmatizer = WordNetLemmatizer()

        return lambda doc: (lemmatizer.lemmatize(w) for w in analyzer(doc))

def extract_topics(model, feature_names, n_top_words):
    """
    Extracts the top words for each topic from the LDA model.

    Parameters:
    model (LatentDirichletAllocation): The trained LDA model.
    feature_names (array): The feature names from the vectorizer.
    n_top_words (int): The number of top words to extract for each topic.

    Returns:
    list: A list of lists containing the top words for each topic.
    """

    topics = []

    for index, topic in enumerate(model.components_):
        topic_words = [feature_names[i] for i in topic.argsort()[::-n_top_words - 1 :-
1]]
        topics.append(topic_words)

    return topics

def get_topic_words_plot(lda, feature_names, topic_index):
    """
    Generates a plot for the words in a given topic.

    Parameters:
    model (LatentDirichletAllocation): The trained LDA model.
    feature_names (array): The feature names from the vectorizer.
    topic_index (int): The index of the topic to plot.

    Returns:
    object: The plot object.
    """

    topic = lda.components_[topic_index]
    top_words = [feature_names[i] for i in topic.argsort()[::-1000 - 1 :-1]]
    firstcloud = WordCloud(background_color='black', width=3000,
height=2000).generate(" ".join(top_words))
    plt.imshow(firstcloud)
    plt.axis('off')

    return save_plot_to_bytes_obj(plt)

```

```

preprocessed_word_list = [preprocess_text(text) for text in reviews_text]

vectorizer = LemmaCountVectorizer(max_df=0.95, min_df=2, stop_words='english',
decode_error='ignore')
tf_matrix = vectorizer.fit_transform(preprocessed_word_list)

lda = LatentDirichletAllocation(n_components=5, max_iter=5, learning_method =
'online', learning_offset = 50., random_state = 0)
lda.fit(tf_matrix)

feature_names = vectorizer.get_feature_names_out()
topics = extract_topics(lda, feature_names, n_top_words=40)

results = []
for i in range(5):
    graph = get_topic_words_plot(lda, feature_names, i)
    results.append({'topic': topics[i], 'graph': graph})

return results

def get_average_weekly_sentiment_score_plot(dataframe):
    """
    Plots the average weekly sentiment score of the review texts in a DataFrame.

    Parameters:
    dataframe (pandas.DataFrame): DataFrame containing the review text data.

    Returns:
    bytes: The byte value of the saved plot in PNG format.
    """

    sid = SentimentIntensityAnalyzer()
    sentiment_scores = dataframe['review_text'].apply(sid.polarity_scores)
    sentiment = sentiment_scores.apply(lambda x: x['compound'])

    weekly_sentiment = sentiment.resample('W').mean()

    plt.figure(figsize=(10, 5))
    plt.plot(weekly_sentiment, marker='o')
    plt.xlabel('Date')
    plt.ylabel('Sentiment_score')
    plt.title('Average weekly sentiment score')
    plt.ylim(-1,1.5)

    return save_plot_to_bytes_obj(plt)

def extract_bi_gram(reviews_text):
    """
    Generates a frequency table of the most common bi-grams in the given texts.

```

```

Parameters:
texts (list): A list of text documents.

Returns:
list: A list of dictionaries containing bi-grams and their frequencies.
"""

def get_ngrams(text, n):
    """
    Generates n-grams from the input text.

    Parameters:
    text (str): The text to generate n-grams from.
    n (int): The number of words in each n-gram.

    Returns:
    list: A list of n-grams.
    """

    n_grams = ngrams(nltk.word_tokenize(text),n)
    return [' '.join(grams) for grams in n_grams]

def remove_punctuations(text):
    """
    Removes punctuation from the input text.

    Parameters:
    text (str): The text to remove punctuation from.

    Returns:
    str: The text without punctuation.
    """

    words = nltk.word_tokenize(text)
    punctuation_removed = [w for w in words if w.lower() not in string.punctuation]
    return " ".join(punctuation_removed)

preprocessed_texts = [preprocess_text(text) for text in reviews_text]
cleaned_text = remove_punctuations(str(preprocessed_texts))

bigrams = get_ngrams(cleaned_text, 2)
bigrams_count = Counter(bigrams)

bigram_freq = pd.DataFrame.from_dict(bigrams_count, orient='index',
columns=['frequency'])
bigram_freq = bigram_freq.rename_axis('words').reset_index()
bigram_freq = bigram_freq.sort_values(by='frequency', ascending=False).head(20)

return bigram_freq.to_dict(orient='records')

```

```

def get_most_frequent_common_words(reviews_text, number_of_words=3):
    """
    Generates a word cloud of the most frequent words in the given texts and returns
    the top 3 most frequent words.

    Parameters:
    reviews_text (list): A list of text documents.
    number_of_words (number): A number of most frequent common words to return.

    Returns:
    tuple: A tuple containing the top most frequent words and their frequencies.
    """
    word_frequencies = Counter(str(reviews_text).split())
    top_words = word_frequencies.most_common(number_of_words)

    return [{'word': word, 'frequency': freq} for word, freq in top_words]

def get_common_words_plot(reviews_text):
    """
    Generates and plots a word cloud of the most frequent words in the given texts.

    Parameters:
    texts (list): A list of text documents.

    Returns:
    bytes: The byte value of the saved word cloud plot in PNG format.
    """
    wordcloud = WordCloud().generate(str(reviews_text))

    plt.imshow(wordcloud.recolor(random_state=2017))
    plt.title('Most frequent words')
    plt.axis("off")

    return save_plot_to_bytes_obj(plt)

def read_data(file_path, rows=100):
    """
    Reads data from a CSV file and returns the specified number of rows as a
    DataFrame.

    Parameters:
    file_path (str): The path to the CSV file.
    rows (int, optional): The number of rows to read from the file. Default is 100.

    Returns:
    pandas.DataFrame: The DataFrame containing the read data.
    """
    dataframe = pd.read_csv(file_path)

```

```

dataframe = dataframe.head(rows)

return dataframe

def get_sentiment_scores(dataframe):
    """
    Calculate sentiment scores using VADER and RoBERTa models for each review in the
    DataFrame.

    Parameters:
    dataframe (pandas.DataFrame): The DataFrame containing review text and author ID.

    Returns:
    pandas.DataFrame: The DataFrame with the calculated sentiment scores merged with
    the original data.
    """

    sia = SentimentIntensityAnalyzer()
    MODEL = "cardiffnlp/twitter-roberta-base-sentiment"
    tokenizer = AutoTokenizer.from_pretrained(MODEL)
    model = AutoModelForSequenceClassification.from_pretrained(MODEL)

    sentiment_results = {}

    for i, row in dataframe.iterrows():
        try:
            text = row['review_text']
            author_id = row['author_id']

            vader_scores = sia.polarity_scores(text)
            vader_result_rename = {f"vader_{key}": value for key, value in
            vader_scores.items()}

            roberta_scores = get_polarity_scores_roberta(tokenizer, model, text)
            combined_scores = {**vader_result_rename, **roberta_scores}

            sentiment_results[author_id] = combined_scores
        except RuntimeError:
            print(f'Error processing sentiment for author_id {author_id}')

    sentiment_df =
pd.DataFrame(sentiment_results).T.reset_index().rename(columns={'index':
'author_id'})

    results_df = sentiment_df.merge(dataframe, how='left')
    return results_df

def get_polarity_scores_roberta(tokenizer, model, text):
    """
    Calculate sentiment scores using a pre-trained RoBERTa model.

```

```

Parameters:
tokenizer (transformers.AutoTokenizer): The tokenizer for the RoBERTa model.
model (transformers.AutoModelForSequenceClassification): The pre-trained RoBERTa
model.
text (str): The text to analyze.

Returns:
dict: A dictionary with the sentiment scores from the RoBERTa model.
"""

encoded_text = tokenizer(text, return_tensors='pt')
output = model(**encoded_text)
scores = output[0][0].detach().numpy()
scores = softmax(scores)
scores_dict = {
    'roberta_neg': scores[0],
    'roberta_neu': scores[1],
    'roberta_pos': scores[2]
}

return scores_dict

def get_distribution_of_customer_rating_plot(dataframe):
    """
    Plot distribution of customer rating using Seaborn and return the plot as a bytes
    object.

    Parameters:
    dataframe (pandas.DataFrame): The DataFrame containing customer ratings.

    Returns:
    bytes: The byte value of the saved bar plot in PNG format.
    """

    rating_counts = dataframe['rating'].value_counts()
    labels = rating_counts.index.tolist()

    plt.pie(rating_counts, labels=labels, autopct='%1.1f%%', startangle=140)
    plt.axis('equal')
    plt.title('Distribution of customer rating')

    plt.legend(bbox_to_anchor=(1, 0.5), loc="center right", title="Rating",
title_fontsize="13")

    return save_plot_to_bytes_obj(plt)

def get_average_weekly_rating_plot(dataframe):
    """
    Plot the average weekly customer rating and return the plot as a bytes object.

```

```

Parameters:
    dataframe (pandas.DataFrame): The DataFrame containing customer ratings with a
    datetime index.

Returns:
    bytes: The byte value of the saved line plot in PNG format.
    """

plt.figure(figsize=(10, 5))
plt.plot(dataframe['rating'].resample('W').mean(), marker='o')
plt.xlabel('Date')
plt.ylabel('Rating')
plt.title('Average weekly customer rating')
plt.ylim(0, 6)

return save_plot_to_bytes_obj(plt)

def get_reviews_per_week_plot(dataframe):
    """
    Plot the number of reviews per week.

    Args:
    - df (DataFrame): DataFrame containing the data with a 'date' column.

    Returns:
    - bytes: The plot as a bytes object.
    """

    plt.figure(figsize=(10, 5))
    plt.plot(dataframe.resample('W')['review_text'].count(), marker='o')
    plt.xlabel('Date')
    plt.ylabel('Number of reviews')
    plt.title('Number of reviews per week')
    plt.ylim(0, dataframe['review_text'].resample('W').count().max() + 1)

    return save_plot_to_bytes_obj(plt)

def get_pairplot(results_df):
    """
    Plot a pairplot using Seaborn and return the plot as a bytes object.

    Parameters:
    results_df (pandas.DataFrame): DataFrame containing the sentiment scores and
    ratings.

    Returns:
    bytes: The byte value of the saved pair plot in PNG format.
    """

```

```

sns.pairplot(data=results_df,
             vars=['vader_neg', 'vader_neu', 'vader_pos',
                  'roberta_neg', 'roberta_neu', 'roberta_pos'],
             hue='rating',
             palette='tab10')

return save_plot_to_bytes_obj(plt)

def get_reviews_by_sentiment(results_df, rating, model_type, sentiment_type,
                             top_n=10):
    """
    Get the top N reviews based on sentiment score and rating.

    Parameters:
    results_df (pandas.DataFrame): DataFrame containing the reviews, ratings, and
    sentiment scores.
    rating (int): The rating to filter reviews by.
    model_type (str): The sentiment analysis model to use ('vader' or 'roberta').
    sentiment_type (str): The sentiment type to sort by ('neg', 'neu', or 'pos').
    top_n (int, optional): The number of top reviews to return. Default is 10.

    Returns:
    list: A list of the top N reviews based on the specified sentiment score and
    rating.
    """

    if model_type == 'vader':
        sort_column = f'vader_{sentiment_type}'
    elif model_type == 'roberta':
        sort_column = f'roberta_{sentiment_type}'
    else:
        raise ValueError("model_type must be 'vader' or 'roberta'")

    reviews = results_df.query(f'rating == {rating}') \
        .sort_values(sort_column, ascending=False)['review_text'].head(top_n).values

    return reviews

```

*Джерело: створено автором у середовищі Python*

**Вміст файлу темплейту upload.html**

```
<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <title>Upload CSV File</title>
  <style>
    body {
      font-family: Arial, sans-serif;
      line-height: 1.6;
      margin: 0;
      padding: 0;
      background-color: #f4f4f4;
      color: #333;
      display: flex;
      justify-content: center;
      align-items: center;
      height: 100vh;
    }

    .container {
      width: 80%;
      max-width: 500px;
      padding: 20px;
      background-color: #fff;
      box-shadow: 0 0 10px rgba(0, 0, 0, 0.1);
      border-radius: 10px;
      text-align: center;
    }

    h1 {
      color: #333;
      margin-bottom: 20px;
    }

    form {
      display: flex;
      flex-direction: column;
      align-items: center;
    }

    input[type="file"] {
      margin-bottom: 20px;
      padding: 10px;
      border: 1px solid #ddd;
      border-radius: 5px;
    }
  </style>
</head>
<body>
  <h1>Upload CSV File</h1>
  <form>
    <input type="file" />
  </form>
</body>
</html>
```

```
        background-color: #fff;
        cursor: pointer;
    }

    input[type="file"]:hover {
        border-color: #333;
    }

    input[type="submit"] {
        padding: 10px 20px;
        border: none;
        border-radius: 5px;
        background-color: #333;
        color: #fff;
        cursor: pointer;
        font-size: 16px;
    }

    input[type="submit"]:hover {
        background-color: #555;
    }
</style>
</head>

<body>
  <div class="container">
    <h1>Upload CSV File</h1>
    <form method="post" enctype="multipart/form-data">
      <input type="file" name="file" accept=".csv">
      <input type="submit" value="Upload">
    </form>
  </div>
</body>
</html>
```

*Джерело: створено автором у середовищі Python*

## Вміст файлу темплейту results.html

```
<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <title>Analysis Results</title>
  <style>
    body {
      font-family: Arial, sans-serif;
      line-height: 1.6;
      margin: 0;
      padding: 0;
      background-color: #f4f4f4;
      color: #333;
    }

    .container {
      width: 80%;
      margin: 20px auto;
      padding: 20px;
    }

    h1, h2, h3 {
      color: #333;
      margin-bottom: 10px;
    }

    .section {
      margin-bottom: 20px;
    }

    img {
      max-width: 100%;
      height: auto;
      display: block;
      margin-bottom: 20px;
    }

    table {
      width: 100%;
      border-collapse: collapse;
      margin-bottom: 20px;
    }

    table th, table td {
      padding: 10px;
```

```
border: 1px solid #ddd;
text-align: left;
}

table th {
  text-transform: capitalize;
  text-align: center;
}

table tr:hover {
  background-color: #333;
  color: #fff;
}

.reviews {
  display: flex;
  flex-wrap: wrap;
  justify-content: space-between;
}

.reviews {
  width: 100%;
  padding: 0;
  list-style: none;
}

.reviews li {
  background: #f9f9f9;
  margin-bottom: 10px;
  padding: 10px;
  width: 30%;
  border-radius: 5px;
  box-shadow: 1px 4px 4px rgba(0, 0, 0, 0.05);
}

@media (max-width: 768px) {
  .reviews {
    width: 100%;
  }
}

.multi-column-table {
  width: 100%;
  margin-bottom: 20px;
  border-collapse: collapse;
}

.multi-column-table td, .multi-column-table th {
  padding: 10px;
  border: 1px solid #ddd;
}
```

```

}

.multi-column-table td:nth-child(odd) {
  background-color: #f9f9f9;
}

.multi-column-table td:nth-child(even) {
  background-color: #fff;
}

.multi-column-table th {
  background-color: #f4f4f4;
  color: #333;
}

.multi-column-table tr:hover td {
  background-color: #333;
  color: #fff;
}
</style>
</head>

<body>
  <div class="container">
    <h1>Analysis Results</h1>
    <h2>Graph</h2>
    

    <h3>Number of reviews per week</h3>
    

    <h3>Distribution of customer rating</h3>
    

    <h3>Average weekky customer rating</h3>
    

    <h3>Average weekly sentiment score</h3>
    

    <h3>The most common and high-frequency words in the review data</h3>
    

    <h2>Common words around <span>{{ most_frequent_common_words|length }}</span>
most frequent words</h2>
    <table class="multi-column-table">

```

```

<thead>
  <tr>
    <th colspan="2">{{ common_words_around_1_result[0]['word'] }}</th>
    <th colspan="2">{{ common_words_around_2_result[0]['word'] }}</th>
    <th colspan="2">{{ common_words_around_3_result[0]['word'] }}</th>
  </tr>
  <tr>
    <th>Word</th>
    <th>Frequency</th>
    <th>Word</th>
    <th>Frequency</th>
    <th>Word</th>
    <th>Frequency</th>
  </tr>
</thead>
<tbody>
  {% for i in range(0, common_words_around_1_result|length) %}
    <tr>
      <td>{{ common_words_around_1_result[i].predicted_word }}</td>
      <td>{{ common_words_around_1_result[i].probability }}</td>
      <td>{{ common_words_around_2_result[i].predicted_word }}</td>
      <td>{{ common_words_around_2_result[i].probability }}</td>
      <td>{{ common_words_around_3_result[i].predicted_word }}</td>
      <td>{{ common_words_around_3_result[i].probability }}</td>
    </tr>
  {% endfor %}
</tbody>
</table>

<h2>Top 20 Bigram Frequency</h2>
<table class="multi-column-table">
  <tr>
    <th>Bigram</th>
    <th>Frequency</th>
  </tr>
  {% for row in bi_dict %}
    <tr>
      <td>{{ row.words }}</td>
      <td>{{ row.frequency }}</td>
    </tr>
  {% endfor %}
</table>

<h1>Topic modeling results</h1>
{% for result in topics_results %}
  <h2>Topic #{{ loop.index }}</h2>
  <p>{{ result.topic | join(', ') }}</p>
  
  {% endfor %}

```

```
<h2>Positive reviews</h2>
<h3>Roberta</h3>
<ul class="reviews">
  {% for review in pos_reviews_roberta %}
    <li>{{ review }}</li>
  {% endfor %}
</ul>

<h3>Vader</h3>
<ul class="reviews">
  {% for review in pos_reviews_vader %}
    <li>{{ review }}</li>
  {% endfor %}
</ul>

<h2>Negative reviews</h2>
<h3>Roberta</h3>
<ul class="reviews">
  {% for review in neg_reviews_roberta %}
    <li>{{ review }}</li>
  {% endfor %}
</ul>

<h3>Vader</h3>
<ul class="reviews">
  {% for review in neg_reviews_vader %}
    <li>{{ review }}</li>
  {% endfor %}
</ul>
</div>
</body>
</html>
```

*Джерело: створено автором у середовищі Python*

**Київський національний університет імені Тараса Шевченка**

**Економічний факультет**

**Кафедра економічної кібернетики**

**ЗАВДАННЯ**

**на кваліфікаційну роботу бакалавра**

студентки 4 курсу спеціальності 051 «Економіка», ОПП «Економічна кібернетика»

Могиль Сніжани Сергіївни

1. Тема роботи: підтримка маркетингових рішень в електронній комерції на основі технологій штучного інтелекту.
2. Термін завершення роботи: 02.06.2024.
3. Попередній захист роботи: 03.06.2024.
4. Об'єкт дослідження: маркетингові рішення в сфері електронної комерції.
5. Предмет дослідження: технології штучного інтелекту, які використовуються для підтримки та оптимізації маркетингових рішень в контексті електронної комерції.
6. Мета дослідження: розробка теоретичних і практичних засад для підтримки прийняття маркетингових рішень в електронній комерції за допомогою технологій штучного інтелекту на основі аналізу сучасних методів та інструментів ШІ, оцінення їхньої ефективності в аспектах маркетингової діяльності, а також інтеграції ШІ у маркетингові стратегії компаній для підвищення їхньої конкурентоспроможності.
7. Завдання дослідження:
  - 7.1. Визначити ключові аспекти функціонування е-комерції та е-маркетингу в цифровому середовищі.
  - 7.2. Проаналізувати компоненти маркетингового міксу в контексті електронної комерції.

7.3. Запропонувати концепцію впровадження технологій ШІ для оптимізації маркетингових стратегій.

7.4. Дослідити вплив технологій штучного інтелекту на різні аспекти маркетингової діяльності, такі як таргетинг, сегментація ринку, рекомендаційні системи, аналіз споживчої поведінки та інші.

7.5. Визначити та оцінити можливості штучного інтелекту та машинного навчання для проведення маркетингових досліджень.

7.6. Створити програмне забезпечення для аналізу користувацьких відгуків з використанням методів сентимент-аналізу та тематичного моделювання.

7.7. Провести оцінку результатів аналізу фідбеку користувачів, надати рекомендації з імплементації програми та визначити перспективи її використання для вдосконалення маркетингових стратегій у електронній комерції.

Науковий керівник: доктор економічних наук, професор Черноус Галина Олександрівна

Студент: Могиль Сніжана Сергіївна

Затверджено на засіданні кафедри економічної кібернетики  
протокол № 4 від 22.11.2023 р.

### Календарний план виконання кваліфікаційної роботи бакалавра

№	Етапи роботи	Терміни виконання	Відмітка керівника про виконання
1	Вибір теми кваліфікаційної роботи бакалавра.	До 22.11.2023	Виконано
2	Розробка та затвердження завдання кваліфікаційної роботи бакалавра.	До 22.11.2023	Виконано
3	Визначення понять електронної комерції та електронного маркетингу, їх особливостей. Написання підрозділу 1.1.	До 7.03.2024	Виконано
4	Аналіз ключових компонентів маркетингової стратегії бізнесу, що функціонує онлайн. Написання підрозділу 1.2.	До 14.03.2024	Виконано
5	Дослідження ролі технологій ШІ у підтримці маркетингових рішень на платформах електронної комерції. Написання підрозділу 1.3.	До 21.03.2024	Виконано
6	Огляд концепції впровадження технологій ШІ для оптимізації маркетингових стратегій. . Написання підрозділу 2.1.	До 11.04.2024	Виконано
7	Дослідження методів ШІ та машинного навчання для	До 18.04.2024	Виконано

	підтримки концепції. Написання підрозділів 2.2, 2.3.		
8	Аналіз бізнес-процесів та постановка цілей дослідження, вибір інструментів реалізації програмного забезпечення. Написання підрозділу 3.1.	До 25.04.2024	Виконано
9	Опис етапів отримання результатів аналізу відгуків користувачів з використанням запропонованого інструменту. Написання розділу 3.2.	До 25.04.2024	Виконано
10	Реалізація програмного забезпечення у середовищі Python.	До 16.05.2024	Виконано
11	Реалізація запропонованого сценарію на конкретному датасеті. Написання підрозділу 3.3.	До 16.05.2024	Виконано
12	Опис упровадження інструменту аналізу відгуків та визначення напрямків удосконалення програмного засобу. Написання підрозділу 3.4.	До 23.05.2024	Виконано
13	Оформлення вступу та висновків, джерел, додатків. Форматування файлу роботи.	До 30.05.2024	Виконано

14	Подання роботи для перевірки на плагіат.	До 31.05.2024	Виконано
15	Подання роботи до попереднього захисту.	До 03.06.2024	Виконано

Науковий керівник: доктор економічних наук, професор

Чорноус Галина Олександрівна

Студент: Могиль Сніжана Сергіївна