

ДЕРЖАВНИЙ УНІВЕРСИТЕТ ТЕЛЕКОМУНІКАЦІЙ

Навчально-науковий інститут **Телекомунікацій**

Кафедра системного аналізу

Пояснювальна записка

до бакалаврської роботи

на ступінь вищої освіти бакалавр

на тему: «**Інформаційна технологія аналізу цін електронних товарів**»

Виконав: студент 4 курсу,
групи САД-41, спеціальності
124 Системний аналіз

(шифр і назва
спеціальності)
Козіянчук О.І.

Керівник (прізвище та ініціали)
Козаченко С.Я.

Рецензент (прізвище та ініціали)

(прізвище та ініціали)

Київ – 2023

ДЕРЖАВНИЙ УНІВЕРСИТЕТ ТЕЛЕКОМУНІКАЦІЙ

Навчально-науковий інститут Телекомунікацій

Кафедра _____ Системного аналізу _____

Ступінь вищої освіти _____ «Бакалавр» _____

Спеціальність _____ 124 Системний аналіз _____

ЗАТВЕРДЖУЮ

Завідувач кафедри

Системного аналізу

_____ (ПІБ)

Дата: _____ 2023 року

ЗАВДАННЯ

НА БАКАЛАВРСЬКУ РОБОТУ СТУДЕНТУ

Козіянчук Олександр Ігорович
(прізвище, ім'я, по батькові)

1. Тема роботи: «Інформаційна технологія аналізу цін електронних товарів»

Керівник роботи Козаченко С.Я. к.е.н, доцент, доцент кафедри Системного аналізу

(прізвище, ім'я, по батькові, науковий ступінь,
вчене звання)

затверджені наказом вищого навчального закладу від _____ 2023 року № _____.

2. Строк подання студентом роботи _____

3. Вхідні дані до роботи:

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

5. Перелік графічного матеріалу

1) _____

2) _____

3) Дата видачі завдання _____

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів бакалаврської роботи	Строк виконання етапів роботи	Примітка
1			
2			
3			
4			
5			
6			
7			
8			
9			

Студент

(підпис)

Козіянчук О.І.

(прізвище та ініціали)

Керівник роботи

(підпис)

Козаченко С.Я.

(прізвище та ініціали)

ДЕРЖАВНИЙ УНІВЕРСИТЕТ ТЕЛЕКОМУНІКАЦІЙ
ПОДАННЯ
ГОЛОВІ ДЕРЖАВНОЇ ЕКЗАМЕНАЦІЙНОЇ КОМІСІЇ
ЩОДО ЗАХИСТУ БАКАЛАВРСЬКОЇ РОБОТИ

Направляється студент Козіянчук О.І. до захисту бакалаврської роботи
(прізвище та ініціали)

за спеціальністю 124 Системний аналіз
(шифр і назва спеціальності)

на тему: «Інформаційна технологія аналізу цін електронних товарів»

Бакалаврська робота і рецензія додаються.

Директор інституту _____ В.І. Кравченко
(підпис)

Довідка про успішність

Козіянчук О.І. за період навчання в Навчально-науковому інституті Телекомунікації
(прізвище та ініціали студента)

з 2021 року до 2023 року повністю виконав навчальний план за спеціальністю з таким розподілом оцінок за:

національною шкалою: відмінно _____%; добре _____%; задовільно _____%;

шкалою ECTS: A _____%; B _____%; C _____%; D _____%; E _____%.

Методист факультету: _____
(підпис) (прізвище та ініціали)

Висновок керівника бакалаврської роботи

Студент Козіянчук Олександр Ігорович

Все це дозволяє оцінити виконану бакалаврську роботу студента Костенка В.К. на оцінку
 «_____» та присвоїти йому кваліфікацію фахівця з інформаційних технологій.

Керівник роботи _____ Козаченко С.Я.
(підпис) (прізвище та ініціали)

_____ 2023 року

Висновок кафедри про бакалаврську роботу

Бакалаврську роботу розглянуто(а). Студент Козіянчук О.І. допускається до захисту даної
(прізвище та ініціали)
 роботи в Державній екзаменаційній комісії.

Завідувач кафедри Системного аналізу _____ Гордієнко Т.Б.
(підпис) (прізвище та ініціали)

РЕФЕРАТ

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ	8
ВСТУП	9
1. РИНОК ЕЛЕКТРОННИХ ТОВАРІВ ЯК ГАЛУЗЬ.....	10
1.1 Галузевий огляд індустрії електронних товарів	10
1.1.1 Роль інформаційних технологій в індустрії електронних товарів та їх вплив на цінові рішення	10
1.1.2 Еволюція інформаційних технологій та її вплив на ціноутворення.....	11
1.2 Механізм ціноутворення в електронній промисловості та фактори, що впливають на цінові рішення	13
1.3 Прогалини в дослідженнях інформаційних технологій для аналізу цін на електронні товари.....	14
2. ІСНУЮЧІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ АНАЛІЗУ ЦІН ЕЛЕКТРОННИХ ТОВАРІВ ТА ПРИНЦИПИ ЇХ РОБОТИ.....	15
2.1 Веб-сайти порівняння цін.....	16
2.1.1 Інструменти веб-скрейпінгу	17
2.1.2 Формування структурованого формату даних	18
2.1.3 Агрегація та систематизація даних	19
2.1.4 Пошук і порівняння цін за Веб-сайтами.....	21
2.1.5 Перенаправлення та стягнення з роздрібних продавців комісії	22
2.2 Програмне забезпечення для аналізу даних.....	23
2.2.1 Програми для аналізу даних	23
2.3 Веб-скрейпінг (web scraping)	32
2.3.1 Інструменти веб-скрейпінгу	32
2.3.2 Як працюють Інструменти веб-скрейпінгу	33
2.3.3 Надсилання HTTP-запитів на цільовий веб-сайт	34
2.3.4 Аналіз вмісту HTML.....	35
2.3.5 Процес вилучення необхідних даних	36
2.3.6 Процес збереження даних	37
2.3.7 Процес опрацювання розбивки на сторінки та динамічного вмісту	38
2.3.8 Процес обробки динамічного контенту.....	39
2.3.9 Вжиття заходів проти парсингу.....	40
2.4 Звіти про дослідження ринку	41

2.4.1 Структура та зміст звітів про дослідження ринку	41
2.4.2 Приклади звітів компанії Gartner	42
2.4.3 Приклади звітів компанії Forrester	43
2.4.4 Приклади звітів компанії IDC.....	45
2.5. Системи електронної точки продажу (EPOS)	46
2.5.1 Концепція роботи EPOS.....	46
2.5.2 Найбільш відомі системи EPOS	47
3. МЕТОДОЛОГІЧНІ І ПРАКТИЧНІ ПІДХОДИ ДО ФОРМУВАННЯ ТЕХНОЛОГІЇ АНАЛІЗУ ЦІН ЕЛЕКТРОННИХ ТОВАРІВ ТА ЙОГО РЕАЛІЗАЦІЇ.....	48
3.1 Формування факторного аналізу цінових тенденцій з позиції маркетингу.	51
3.2 Оцінка та інтерпретація результатної аналітичної інформації (на прикладі торгової мережі Best Buy)	58
3.2.1 Статистична оцінка тісноти зв'язків.....	58
3.2.2 Рівняння множинної регресії впливу на ціну виділених факторів	60
ВИСНОВОК.....	66
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ	68

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

IT	–	Information Technology (Інформаційні технології)
PCW	–	Price Comparison Websites (Сайт порівняння цін)
CSV	–	Comma-Separated Values (Значення розділені зап'ятими)
XML	–	eXtensible Markup Language (Розширювана мова розмітки)
JSON	–	JavaScript Object Notation (Текстовий формат обміну даними)
PPC	–	Pay Per Click (Оплата за клік)
CPA	–	Cost Per Action (Ціна за дію)
SAS	–	Набір програмних продуктів для управління даними.
SPSS	–	Statistical Package for the Social Sciences (Програма для статичної обробки даних)
CRM	–	Customer Relationship Management (Система управління взаємовідношеннями)
ERP	–	enterprise resource planning (Планування ресурсів підприємства)
API	–	Application Programming Interface (Опис способів взаємодії однієї програми з іншою)
HTML	–	HyperText Markup Language (Мова гіпертекстової розмітки)
AJAX	–	Asynchronous Javascript and XML (Асинхронний ДжаваСкрипт та XML)
URL	–	Uniform Resource Locator (Стандартизована адреса певного ресурсу)

ВСТУП

Зі зростанням інтернет-торгівлі, аналіз цін електронних товарів стає все більш важливим для розуміння динаміки ринку. Ціна є ключовим фактором, який впливає на прийняття рішень споживачами, а тому вивчення цінових стратегій та їх вплив на ринок є актуальним і важливим напрямком дослідження. Хоча існують численні дослідження, що стосуються ціноутворення та цінової політики, дослідження цін на електронні товари в Україні все ще недостатньо вивчені. Дослідження базуються на аналізі цін на електронні товари з використанням даних з українських онлайн-ресурсів, а також використання наукових джерел для розуміння контексту. Предметом дослідження є цінові стратегії та їх вплив на ринок електронних товарів. Об'єктом дослідження є ринок електронних товарів в Україні. Метою цієї роботи є аналіз цін на електронні товари, та визначення основних факторів, які впливають на нові стратегії. Завдання включають збір та аналіз даних. Дослідження було виконано, завдяки вивченню ринку електронних товарів, та використанням сучасних технологій, що дали б змогу краще зрозуміти механізми ціноутворення.

1. РИНОК ЕЛЕКТРОННИХ ТОВАРІВ ЯК ГАЛУЗЬ

1.1 Галузевий огляд індустрії електронних товарів

Індустрія електронних товарів відноситься до виробництва, розповсюдження та продажу електронних пристроїв і обладнання, такого як смартфони, ноутбуки, телевізори, камери та побутова техніка. Ця галузь стає дедалі важливішою у світовій економіці через широке використання та залежність від електронних пристроїв і обладнання в повсякденному житті.

Галузь характеризується швидким технологічним прогресом та інноваціями, які призвели до впровадження нових продуктів і функцій, а також до постійного вдосконалення існуючих. Це створило висококонкурентний ринок, на якому багато гравців борються за частку ринку і прибуток.

Електронна промисловість також робить великий внесок у світову економіку з річним доходом у трильйони доларів. Це створює можливості працевлаштування в різних секторах, таких як виробництво, логістика, продажі та обслуговування клієнтів. Крім того, галузь стимулює інновації та технологічні досягнення в інших секторах, таких як охорона здоров'я, транспорт і зв'язок, які ще більше сприяють економічному зростанню і розвитку.

Загалом, індустрія електронних товарів відіграє вирішальну роль у формуванні сучасного світу, і її постійне зростання та успіх життєво важливі для світової економіки.

1.1.1 Роль інформаційних технологій в індустрії електронних товарів та їх вплив на цінові рішення

На галузь електронних товарів значний вплив справили інформаційні технології (ІТ), які здійснили революцію в розробці, виробництві, маркетингу та продажу продуктів. ІТ також відіграли ключову роль у формуванні цінових рішень в електронній промисловості [8].

Одним з основних способів впливу ІТ на цінові рішення є доступність даних у режимі реального часу про ринкові тенденції, поведінку споживачів і цінові стратегії конкурентів. Тепер компанії можуть використовувати передові

аналітичні інструменти та програмне забезпечення для збору й аналізу цих даних, що дає їм змогу ухвалювати більш обґрунтовані цінові рішення на основі поточних ринкових умов.

ІТ також дали змогу компаніям реалізувати стратегії динамічного ціноутворення, за яких ціни можуть коригуватися в режимі реального часу залежно від змін попиту, пропозиції та цін конкурентів. Це дає змогу компаніям максимізувати прибуток і ефективніше захоплювати частку ринку, швидко реагуючи на зміни ринкових умов.

Ще один спосіб, у який ІТ вплинув на рішення про ціноутворення в індустрії електронних товарів, - це використання онлайн-ринків і платформ електронної комерції. Ці платформи надають споживачам зручний спосіб придбання електронних товарів, а також дають змогу компаніям охопити ширшу аудиторію та конкурувати в глобальному масштабі. Компанії можуть використовувати ці платформи, щоб пропонувати персоналізовані ціни та рекламні акції для певних сегментів клієнтів на основі їх історії переглядів і покупок [8].

Загалом роль ІТ в електронній промисловості справила значний вплив на цінові рішення. Компанії, здатні ефективно використовувати ІТ, краще адаптуються до мінливих ринкових умов і залишаються конкурентоспроможними в галузі, що постійно розвивається.

1.1.2 Еволюція інформаційних технологій та її вплив на ціноутворення

Еволюція інформаційних технологій справила значний вплив на ціноутворення в різних галузях, зокрема в електронній промисловості. Інформаційні технології дали змогу виробникам збирати й аналізувати величезну кількість даних про поведінку споживачів, ринкові тенденції та стратегії ціноутворення конкурентів. Це дало змогу виробникам розробити більш ефективні стратегії ціноутворення, адаптовані до конкретних умов.

Одним із способів впливу інформаційних технологій на ціноутворення є використання аналітики великих даних. Вона охоплює збір і аналіз великих наборів даних для виявлення закономірностей і тенденцій, які можуть вплинути

на рішення про ціноутворення. Це дало змогу виробникам отримати уявлення про поведінку та вподобання споживачів, що можна використовувати для розроблення стратегій ціноутворення, більш ефективних для задоволення потреб споживачів [16].

Ще одним способом, яким інформаційні технології вплинули на ціноутворення, є використання динамічного ціноутворення. Динамічне ціноутворення передбачає коригування цін у режимі реального часу на основі змін попиту та пропозиції. Це можна зробити за допомогою алгоритмів, які аналізують ринкові дані та відповідним чином коригують ціни. Динамічне ціноутворення дає змогу виробникам швидко реагувати на зміни ринкових умов, що може допомогти максимізувати прибуток.

Еволюція інформаційних технологій також дала змогу виробникам персоналізувати стратегії ціноутворення. Персоналізоване ціноутворення передбачає адаптацію стратегій ціноутворення для окремих споживачів на основі їх минулої поведінки, вподобань і демографічних даних. Це можна зробити за допомогою алгоритмів машинного навчання, які аналізують дані про поведінку та вподобання споживачів. Індивідуальне ціноутворення може допомогти виробникам збільшити продажі, пропонуючи споживачам ціни, які є більш актуальними і привабливими для їх конкретних потреб.

Таким чином, розвиток інформаційних технологій справив значний вплив на ціноутворення в електронній промисловості [12]. Це дало змогу виробникам збирати й аналізувати величезну кількість даних про поведінку споживачів, ринкові тенденції та стратегії ціноутворення конкурентів, що допомогло розробити більш ефективні стратегії ціноутворення. Інформаційні технології також уможливили динамічне ціноутворення і персоналізоване ціноутворення, що може допомогти виробникам швидко реагувати на зміни ринкових умов і пропонувати ціни, які більше підходять для окремих споживачів.

1.2 Механізм ціноутворення в електронній промисловості та фактори, що впливають на цінові рішення

На механізм ціноутворення в електронній промисловості впливає безліч чинників, включно з виробничими витратами, конкуренцією, попитом і поведінкою споживачів [14].

Одним з основних чинників, що впливають на цінові рішення, є виробничі витрати, пов'язані з виробництвом і поширенням електронних товарів. Ці витрати включають сировину, робочу силу, транспортування і накладні витрати, і вони можуть значно відрізнятись залежно від типу продукту і виробничого процесу.

Іншим ключовим фактором є конкуренція, оскільки індустрія електронних товарів характеризується великою кількістю гравців і гострою конкуренцією. Тому компанії повинні враховувати ціни на аналогічні товари, пропоновані конкурентами, під час встановлення своїх власних цін. На цінові рішення також можуть впливати стратегічні цілі компанії, такі як захоплення частки ринку або максимізація прибутку.

Попит також є важливим чинником під час ухвалення рішення щодо ціноутворення, оскільки рівень попиту на продукт може впливати на його сприйману цінність і впливати на готовність споживачів платити певну ціну. Таким чином, компанії можуть коригувати свої ціни залежно від змін попиту, а також зовнішніх чинників, таких як економічні умови або сезонні тенденції.

Поведінка споживачів є ще одним важливим фактором під час ухвалення рішення про ціноутворення. Компанії можуть проводити дослідження ринку, щоб отримати уявлення про споживчі вподобання і чутливість до цін, і використовувати цю інформацію для визначення стратегій ціноутворення, які з більшою ймовірністю знайдуть відгук у споживачів і стимулюють продажі [15].

Загалом механізм ціноутворення в електронній промисловості складний і залежить від низки внутрішніх і зовнішніх чинників. Компанії мають ретельно враховувати ці чинники під час встановлення цін, щоб залишатися конкурентоспроможними та прибутковими на ринку, що постійно змінюється.

1.3 Прогалини в дослідженнях інформаційних технологій для аналізу цін на електронні товари

Існує кілька прогалин у дослідженнях у галузі інформаційних технологій для аналізу цін на електронні товари, які необхідно усунути [5]. Деякі з цих прогалин включають:

1. Відсутність комплексних досліджень. Брак комплексних досліджень, що вивчають роль інформаційних технологій в аналізі цін на електронні товари. Більшість досліджень зосереджено на конкретному аспекті інформаційних технологій або ціноутворенні, що обмежує можливість робити ширші висновки.
2. Обмежена увага до конкретних технологій: більшість досліджень зосереджена на конкретних технологіях, таких як штучний інтелект або аналітика великих даних, і не вивчає потенційну синергію між різними технологіями. Це обмежує можливості розробки більш інтегрованих та ефективних стратегій ціноутворення.
3. Відсутність досліджень ефективності стратегій ціноутворення. Незважаючи на те, що існують дослідження, які вивчають ефективність різних стратегій ціноутворення, немає комплексних досліджень, які б порівнювали ефективність різних стратегій ціноутворення в різних контекстах. Це обмежує можливість розроблення стратегій ціноутворення, адаптованих до конкретних умов.
4. Обмежена увага до впливу ціноутворення на поведінку споживачів: недостатньо досліджень, що вивчають вплив ціноутворення на поведінку споживачів у контексті електронних товарів. Це обмежує можливості розроблення стратегій ціноутворення, заснованих на всебічному розумінні поведінки споживачів.
5. Обмежена увага до впливу ціноутворення на динаміку ринку: недостатньо досліджень, що вивчають вплив ціноутворення на динаміку ринку в контексті електронних товарів. Це обмежує можливості розроблення

стратегій ціноутворення, заснованих на всебічному розумінні динаміки ринку.

Усунення цих прогалин у дослідженнях може допомогти в розробленні більш повного і комплексного розуміння ролі інформаційних технологій в аналізі цін на електронні товари. Це також може допомогти в розробленні ефективніших стратегій ціноутворення, адаптованих до конкретних умов і поведінки споживачів [17].

2. ІСНУЮЧІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ АНАЛІЗУ ЦІН ЕЛЕКТРОННИХ ТОВАРІВ ТА ПРИНЦИПИ ЇХ РОБОТИ

Існує кілька інформаційних технологій, які можна використовувати для аналізу цін на електронні товари. Ось деякі з найпоширеніших:

1. Веб-сайти порівняння цін : це веб-сайти, які дозволяють користувачам порівнювати ціни на товари в різних інтернет-магазинах. Приклади таких веб-сайтів включають PriceGrabber, Shopzilla і Google Shopping.
2. Програмне забезпечення для аналізу даних: це програми, які можна використовувати для аналізу даних про ціни для виявлення тенденцій і закономірностей. Приклади такого програмного забезпечення включають Tableau, Power BI і QlikView.
3. Інструменти веб-скрейпінгу: це інструменти, які можна використовувати для вилучення даних про ціни з веб-сайтів інтернет-магазинів. Приклади таких інструментів включають Scrapy, BeautifulSoup і Selenium.
4. Звіти про дослідження ринку: це звіти, які дають уявлення про ринок електронних товарів, включно з тенденціями ціноутворення, поведінкою споживачів і конкурентним середовищем. Приклади таких звітів включають звіти Gartner, Forrester і IDC.
5. Системи електронної точки продажу (EPOS): це системи, які збирають дані про продажі в точці купівлі, даючи змогу роздрібним торговцям аналізувати тенденції ціноутворення і продажів у режимі реального часу. Приклади таких систем включають Square, Shopify і Vend.

Використовуючи ці інформаційні технології, підприємства можуть краще зрозуміти тенденції ціноутворення на ринку електронних товарів, що дасть їм змогу ухвалювати більш обґрунтовані цінові рішення і залишатися конкурентоспроможними [9].

2.1 Веб-сайти порівняння цін

Веб-сайти порівняння цін (PCW) дають змогу споживачам порівнювати ціни на товари в декількох інтернет-магазинах. Ці веб-сайти використовують спеціалізоване програмне забезпечення для автоматичного збору та відображення інформації про ціни та наявність товарів, що економить час і сили покупців.

Алгоритм роботи сайтів порівняння цін в інтернет-магазинах:

1. Збір даних: Програмне забезпечення веб-сайту порівняння цін переглядає Інтернет, скануючи інтернет-магазини в пошуках інформації про продукти, ціни та наявність. Ці дані збираються за допомогою інструментів веб-скрейпінгу, які запрограмовані на автоматичне вилучення інформації з веб-сайтів у структурованому форматі.
2. Агрегація даних. Після того, як дані зібрані, вони систематизуються і об'єднуються в базу даних, в якій споживачі можуть здійснювати пошук. Програмне забезпечення веб-сайту зіставляє продукти за унікальними ідентифікаторами, такими як назва продукту, номер моделі або штрих-код. Це гарантує, що всі списки для певного продукту згруповані разом у результатах пошуку.
3. Пошук і порівняння: споживачі вводять умови пошуку або ключові слова для продукту, який вони шукають, і програмне забезпечення веб-сайту повертає список відповідних продуктів із різних інтернет-магазинів. Результати пошуку зазвичай містять інформацію про ціну продукту, доступність, варіанти доставки та рейтинг продавця. Потім покупці можуть сортувати і фільтрувати результати за різними критеріями, такими як ціна, рейтинг і варіанти доставки.

4. Перенаправлення до продавця: коли покупець знаходить продукт, який хоче купити, він натискає на список, який перенаправляє його на веб-сайт продавця, де він може зробити покупку. Веб-сайт порівняння цін отримує дохід, стягуючи з роздрібних продавців комісію за кожен клік або продаж, здійснений через їх платформу.

Загалом, веб-сайти порівняння цін надають споживачам цінну послугу, спрощуючи пошук найкращих пропозицій на товари з декількох інтернет-магазинів. Автоматизуючи процес збору та систематизації даних про ціни, ці веб-сайти заощаджують час і зусилля покупців, допомагаючи їм ухвалювати більш обґрунтовані рішення щодо купівлі.

2.1.1 Інструменти веб-скрейпінгу

Інструменти веб-скрейпінгу - це програми, призначені для вилучення даних із веб-сайтів. Вони використовуються для швидкого та ефективного збору великих обсягів даних, що дає змогу аналізувати та використовувати інформацію в різних додатках. Ось як працюють інструменти веб-скрейпінгу:

1. Визначення цільового веб-сайту. Першим кроком у веб-скрейпінгу є визначення цільового веб-сайту, з якого будуть витягуватись дані. URL-адресу веб-сайту вводять в інтерфейс інструменту веб-скрейпінгу, і інструмент витягує HTML-код веб-сайту.
2. Аналіз HTML-коду. Потім інструмент парсингу веб-сторінок аналізує HTML-код веб-сайту, щоб визначити, які елементи даних необхідно витягти. Це робиться шляхом вивчення структури веб-сайту, такої як теги, класи та ідентифікатори, які використовуються для визначення різних частин сторінки.
3. Витяг даних. Щойно елементи даних будуть ідентифіковані, інструмент веб-скрейпінгу витягує відповідні дані з HTML-коду. Це робиться за допомогою спеціального коду, який "витягує" дані зі сторінки та зберігає їх у структурованому форматі : файл CSV або JSON.

4. Обробка динамічного контенту. Деякі веб-сайти використовують динамічний контент, що означає, що сторінка змінюється залежно від взаємодії з користувачем або інших факторів. У цих випадках інструменти веб-скрейпінгу можуть використовувати такі методи, як сканування AJAX або автономний перегляд, щоб гарантувати збір усіх відповідних даних, навіть якщо вони генеруються динамічно.
5. Управління проксі-серверами і призначеними для користувача агентами: щоб веб-сайти не блокували свої запити або не ідентифікували їх як веб-скрейпери, користувачам може знадобитися налаштувати свої інструменти веб-скрейпінгу для використання проксі-серверів і ротації призначених для користувачів агентів. Це допомагає зробити так, щоб запити виглядали так, ніби вони надходять від користувачів-людей.

Загалом, інструменти веб-скрапінгу надають потужний спосіб збирання та аналізу даних із веб-сайтів, даючи змогу користувачам витягувати цінну інформацію для бізнесу або дослідницьких цілей. Хоча парсинг може бути складним процесом, наявність потужних інструментів і ресурсів зробила його більш доступним для користувачів з різним рівнем технічних знань.

2.1.2 Формування структурованого формату даних

Формат структурованих даних відноситься до певного способу організації та подання даних у структурованому вигляді з використанням стандартизованого формату, який може бути легко зрозумілий як людьми, так і машинами. Цей формат призначений для полегшення читання, опрацювання й аналізу даних і використовується в найрізноманітніших додатках, включно з базами даних, електронними таблицями та веб-сторінками.

Структуровані дані організовано в рядки та стовпчики, де кожен стовпчик представляє певний елемент даних, а кожен рядок є унікальним записом або екземпляром даних. Цей формат зазвичай використовується для зберігання та управління даними в реляційних базах даних, де він забезпечує узгоджений і передбачуваний спосіб організації даних і доступу до них.

Деякі поширені формати структурованих даних містять у собі:

1. CSV (значення, розділені комами): у цьому форматі використовуються коми для розділення кожного елемента даних у рядку, що спрощує імпорт та експорт даних між різними додатками.
2. XML (розширювана мова розмітки): у цьому форматі використовують теги для ідентифікації різних елементів даних, і його можна використовувати для зберігання та обміну даними в різних системах.
3. JSON (позначення об'єктів JavaScript): цей формат схожий на XML, але використовує легший синтаксис, який легше читати й аналізувати.

Структуровані дані важливі для аналізу та управління даними, оскільки вони забезпечують узгоджений і організований спосіб зберігання та обробки великих обсягів інформації. Використовуючи стандартизовані формати, можна автоматизувати завдання опрацювання даних і спростити інтеграцію даних із різних джерел, заощаджуючи час і підвищуючи точність даних.

2.1.3 Агрегація та систематизація даних

Агрегація даних

Агрегація даних - це процес об'єднання та узагальнення даних з декількох джерел в єдиний набір даних. Цей процес може включати збір даних з різних джерел, таких як бази даних, електронні таблиці та веб-сайти, а потім їх об'єднання та організацію для створення більш повної і всеосяжної картини.

Агрегація даних - важливий крок в аналізі даних та управлінні ними, оскільки він дає змогу користувачам отримувати цінну інформацію та ухвалювати обґрунтовані рішення на основі великих обсягів даних. Деякі поширені методи агрегування даних включають в себе:

1. Підсумовування: це включає в себе підсумовування даних шляхом розрахунку зведеної статистики, такої як середнє значення, медіана, мода, діапазон і стандартне відхилення. Ці зведені статистичні дані можуть дати уявлення про загальні тенденції та закономірності в даних.

2. Зведення: це включає в себе агрегування даних з декількох рівнів деталізації у зведення вищого рівня. Наприклад, дані про продажі можуть бути перетворені з окремих транзакцій у підсумкові дані за день, тиждень або місяць.
3. Деталізація: це включає в себе розбивку зведених даних на більш докладні рівні інформації. Наприклад, дані про продажі можуть бути деталізовані від місячних підсумків до щоденних або погодинних показників продажів.
4. Фільтрація: це включає в себе вибір і агрегування тільки релевантних даних на основі певних критеріїв. Наприклад, дані про продажі можуть бути відфільтровані, щоб включати тільки продажі з певного регіону або категорії продуктів.

Агрегація даних використовується в різних додатках, включно з бізнес-аналітикою, фінансовим аналізом і дослідженнями ринку. Комбінуючи й узагальнюючи дані з декількох джерел, можна виявляти закономірності та тенденції, ухвалювати обґрунтовані рішення й отримувати повніше уявлення про складні набори даних.

Систематизація даних

Організацію та інтеграцію інформації в базу даних, яку споживачі можуть шукати, зазвичай називають пошуковою базою даних або пошуковою системою. База даних пошуку - це набір упорядкованих даних, у яких користувачі можуть шукати та запитувати певну інформацію.

Для створення бази даних пошуку дані зазвичай упорядковуються в таблиці з пов'язаними елементами даних, згрупованими разом. Кожна таблиця складається з рядків і стовпців, де кожен рядок представляє унікальний запис або екземпляр даних, а кожен стовпець представляє певний елемент даних.

Потім дані індексуються і стають доступними для пошуку за допомогою пошукової системи. Пошукова система - це програма, яка дає змогу користувачам вводити ключові слова або умови пошуку, які потім зіставляються з проіндексованими даними для отримання релевантних результатів. Пошукова система використовує алгоритми і системи ранжування, щоб визначити

релевантність кожного результату і впорядкувати їх таким чином, щоб це було найбільш корисно для користувача.

Бази даних пошуку використовуються в найрізноманітніших додатках, включно з веб-сайтами електронної комерції, онлайн-каталогами та дослідницькими базами даних. Організовуючи та об'єднуючи дані у форматі з можливістю пошуку, пошукові бази даних полегшують споживачам пошук необхідної їм інформації, заощаджуючи час і покращуючи взаємодію з користувачем.

2.1.4 Пошук і порівняння цін за Веб-сайтами

Пошук по веб-сайту і порівняння цін відносяться до процесу пошуку і порівняння цін на товари або послуги на різних веб-сайтах. Цей процес зазвичай здійснюється споживачами, які хочуть зробити покупку і хочуть знайти найкращу можливу угоду.

Пошук за веб-сайтом включає використання пошукової системи або панелі пошуку на певному веб-сайті для пошуку продукту або послуги. У результатах пошуку зазвичай відображається список продуктів або послуг, що відповідають критеріям пошуку, а також короткий опис і ціна. Потім користувачі можуть клацнути на певний продукт або послугу, щоб переглянути детальнішу інформацію і зробити покупку, якщо вони захочуть.

Порівняння цін включає в себе пошук одного й того самого продукту або послуги на декількох веб-сайтах, щоб знайти найкращу можливу ціну. Споживачі можуть використовувати веб-сайти порівняння цін або пошукові системи для пошуку певного продукту або послуги та порівнювати ціни на різних веб-сайтах. Веб-сайти порівняння цін зазвичай надають список веб-сайтів, що пропонують продукт або послугу, а також ціну та будь-яку додаткову інформацію або відгуки.

Мета пошуку на веб-сайті та порівняння цін - допомогти споживачам ухвалювати більш обґрунтовані рішення щодо купівлі, надаючи їм доступ до широкого спектра інформації та варіантів. Порівнюючи ціни на різних веб-сайтах, споживачі можуть заощадити гроші й отримати найкращу пропозицію щодо

продукту або послуги. Пошук по веб-сайтах і порівняння цін останніми роками стають дедалі популярнішими, оскільки споживачі мають ширший доступ до інформації та більше можливостей для здійснення покупок в Інтернеті.

2.1.5 Перенаправлення та стягнення з роздрібних продавців комісії

Роздрібні продавці зазвичай стягують плату за кожен клік або продаж, здійснений через їхню платформу за допомогою різних рекламних моделей, таких як оплата за клік (PPC) або ціна за дію (CPA).

З рекламою оплата за клік роздрібні продавці платять комісію щоразу, коли користувач натискає на рекламу на своїй платформі. Ця модель часто використовується пошуковими системами та платформами соціальних мереж, де роздрібні продавці можуть робити ставки за певними ключовими словами або орієнтуватися на певні демографічні дані, щоб охопити свою цільову аудиторію.

З іншого боку, реклама з оплатою за дію передбачає, що роздрібні продавці платять комісію за кожну конкретну дію, яку користувач виконує на їх платформі, наприклад, за здійснення покупки або заповнення форми для потенційних клієнтів. Ця модель часто використовується в програмах афілійованого маркетингу, коли роздрібні продавці співпрацюють з іншими веб-сайтами або видавцями для просування своїх продуктів або послуг.

В обох випадках продавець зазвичай встановлює максимальну ставку або максимальну ціну за дію, яку він готовий платити, а рекламна платформа або партнерська мережа показуватиме рекламу користувачам на основі різних чинників, таких як сума ставки, релевантність і поведінка користувачів.

Конкретні збори і збори, пов'язані з кожним кліком або продажем, можуть різнитися залежно від платформи і галузі. Деякі платформи також можуть стягувати додаткову плату за такі послуги, як управління рекламою або звітність. Роздрібні продавці повинні ретельно оцінити свої рекламні можливості та витрати, щоб забезпечити максимальну віддачу від своїх інвестицій.

2.2 Програмне забезпечення для аналізу даних

2.2.1 Програми для аналізу даних

Програмне забезпечення для аналізу даних розроблено, щоб допомогти користувачам швидко і легко аналізувати великі обсяги даних. Ці програмні додатки зазвичай використовують комбінацію статистичних алгоритмів і алгоритмів машинного навчання, щоб допомогти користувачам виявляти закономірності та тенденції у своїх даних.

На ринку є безліч варіантів програмного забезпечення для аналізу даних, починаючи від базових додатків для роботи з електронними таблицями та закінчуючи більш просунутим програмним забезпеченням, що включає алгоритми машинного навчання. Ось кілька прикладів популярних програм для аналізу даних:

1. Microsoft Excel - це широко використовуваний додаток для роботи з електронними таблицями, що включає основні інструменти аналізу даних, такі як зведені таблиці, діаграми та статистичні функції.
2. Tableau - це програмне забезпечення для візуалізації даних, яке дає змогу користувачам створювати інтерактивні інформаційні панелі та візуалізації на основі своїх даних.
3. R - це мова програмування і програмне середовище для статистичних обчислень і графіки. Вона широко використовується для аналізу даних і статистичного моделювання.
4. Python - це універсальна мова програмування, яку часто використовують для аналізу даних, машинного навчання та наукових обчислень.
5. SAS - це набір програмних продуктів для управління даними, бізнес-аналітики та прогнозного моделювання.
6. IBM SPSS - це статистичний програмний пакет, який використовується для аналізу даних, прогнозного моделювання та ухвалення рішень.
7. MATLAB - це мова програмування та програмне середовище для чисельних обчислень, візуалізації та аналізу даних.

8. RapidMiner - це платформа для опрацювання даних, яка дає змогу користувачам створювати прогностичні моделі та аналізувати дані за допомогою візуального інтерфейсу.
9. Google Analytics - це служба веб-аналітики, яка надає інформацію про трафік веб-сайту та поведінку користувачів.

Вибір програмного забезпечення для аналізу даних залежатиме від конкретних потреб користувача і складності аналізованих даних.

2.2.1.1 Етапи процесу аналізу даних з використанням програмного забезпечення

1. Імпорт даних. Першим кроком в аналізі даних є імпорт даних у програмний додаток. Це можна зробити, завантаживши файл або підключившись до джерела даних, наприклад до бази даних.
2. Очищення даних: після імпорту даних їх необхідно очистити і підготувати до аналізу. Це включає в себе видалення будь-яких помилок, дублікатів або непотрібних даних.
3. Перетворення даних: дані можуть потребувати перетворення, щоб їх було легше аналізувати. Це може включати перетворення даних в інший формат, агрегування даних або створення нових змінних.
4. Візуалізація даних. Програмне забезпечення для аналізу даних часто містить інструменти для створення візуалізацій, як-от діаграми та графіки, які допомагають користувачам розуміти й інтерпретувати свої дані.
5. Статистичний аналіз. В основі аналізу даних лежить використання статистичних алгоритмів і алгоритмів машинного навчання для виявлення закономірностей і тенденцій у даних. Це може включати виконання регресій, кластеризації або алгоритмів класифікації.
6. Інтерпретація: після того, як дані були проаналізовані, користувачі повинні інтерпретувати результати і робити висновки на основі своїх висновків.
7. Звітність. Нарешті, програмне забезпечення для аналізу даних зазвичай містить інструменти для створення звітів та обміну результатами з іншими.

Доступно безліч різних типів програмного забезпечення для аналізу даних, починаючи від простих додатків для роботи з електронними таблицями та закінчуючи складнішим програмним забезпеченням, що включає передові алгоритми машинного навчання. Вибір програмного забезпечення залежатиме від конкретних потреб користувача і складності аналізованих даних.

2.2.1.2 Імпорт даних у програмні додатки

Дані можна імпортувати в програмний додаток кількома способами, залежно від конкретного програмного забезпечення та формату даних. Ось кілька поширених методів імпорту даних:

1. Завантаження файлу. Дані можна імпортувати в багато програмних додатків, завантаживши файл у певному форматі, наприклад CSV, Excel або JSON. Потім програма прочитає файл і перетворить дані у формат, придатний для аналізу.
2. Підключення до даних. Деякі програмні додатки дають змогу користувачам безпосередньо підключатися до бази даних або джерела даних, наприклад до системи CRM або ERP. Це дає змогу користувачеві аналізувати дані в режимі реального часу без необхідності імпортувати їх вручну.
3. Веб-скрапінг: у деяких випадках дані можна імпортувати в програмний застосунок шляхом очищення даних із веб-сайту або веб-додатка. Це передбачає використання програмних засобів для вилучення даних із веб-сторінок і перетворення їх у формат, придатний для аналізу.
4. Інтеграція API: Багато програмних додатків надають API (інтерфейс прикладного програмування), який дає змогу користувачам програмно імпортувати дані. Цей метод часто використовують для аналізу даних у реальному часі або для інтеграції даних із кількох джерел.
5. Введення даних: у деяких випадках може знадобитися введення даних вручну в програмний додаток. Це можна зробити, ввівши дані безпосередньо в додаток або скопіювавши і вставивши їх з іншого джерела.

Метод, який використовується для імпорту даних, залежатиме від конкретного програмного додатка і формату даних. Деякі програмні додатки можуть підтримувати кілька методів імпорту, даючи змогу користувачам вибирати метод, який найкраще підходить для їхніх потреб.

2.2.1.3 Очищення даних після їхнього імпорту

Очищення даних, також відоме як очищення даних або очищення даних, є процесом виявлення та виправлення або видалення помилок, невідповідностей і неточностей у даних. Цей процес необхідний після імпорту даних у програмний додаток, оскільки дані з різних джерел можуть бути неповними, суперечливими або містити помилки, що можуть вплинути на точність аналізу та результатів.

Ось кілька причин, через які необхідне очищення даних після імпорту даних:

1. Усунення помилок і невідповідностей. Очищення даних допомагає виявляти і виправляти такі помилки, як помилки, дані, що повторюються, відсутні значення і неузгоджені дані. Наприклад, якщо користувач ввів дату в різних форматах, очищення даних може допомогти стандартизувати формат і уникнути плутанини.
2. Підвищення точності. Очищення даних допомагає підвищити точність аналізу за рахунок видалення непотрібних або невірних даних. Це допомагає гарантувати, що результати засновані на надійних і точних даних.
3. Поліпшення якості даних. Усунення помилок і невідповідностей, очищення даних допомагає поліпшити якість даних. Це може допомогти знизити ризик помилок під час аналізу та ухвалення рішень.
4. Стандартизація даних. Очищення даних може допомогти стандартизувати формат даних, як-от дати, валюта або одиниці виміру. Це допомагає уникнути плутанини та забезпечує простоту порівняння й аналізу даних.
5. Економія часу і ресурсів: очищаючи дані перед аналізом, користувачі можуть заощадити час і ресурси, уникаючи необхідності повторного запуску аналізу або виправлення помилок пізніше.

Таким чином, очищення даних є важливим кроком після імпорту даних у програмний додаток, оскільки воно допомагає забезпечити точність, надійність і узгодженість даних. Покращуючи якість даних, користувачі можуть ухвалювати більш обґрунтовані рішення на основі надійного і точного аналізу.

2.2.1.4 Перетворення даних в інший формат, агрегування даних і створення нових змінних

Необхідно перетворити дані в інший формат, агрегувати дані або створити нові змінні, щоб полегшити аналіз даних і отримати змістовну інформацію. Ось кілька причин чому:

1. **Сумісність:** Різні програмні додатки та бази даних можуть вимагати даних у певному форматі. Перетворення даних в інший формат робить їх сумісними з цільовим програмним додатком або базою даних, що дає змогу використовувати їх для аналізу.
2. **Суперечливість:** дані з різних джерел можуть мати різну структуру, одиниці виміру або інші невідповідності. Агрегуючи дані або створюючи нові змінні, можна стандартизувати дані та зробити їх узгодженими, що спрощує їхній аналіз.
3. **Спрощення:** у деяких випадках дані можуть бути занадто складними або детальними для ефективного аналізу. Агрегуючи дані або створюючи нові змінні, можна спростити дані та зосередитися на найважливіших факторах, що спрощує їхнє розуміння та аналіз.
4. **Точність.** Агрегування даних або створення нових змінних може підвищити точність аналізу завдяки врахуванню чинників, які не включено до вихідних даних.
5. **Аналітика:** агрегування даних або створення нових змінних може виявити інсайти, які не видно з вихідних даних. Наприклад, створення нової змінної, яка обчислює середню вартість транзакції для кожного клієнта, може виявити тенденції та закономірності, яких не видно у вихідних даних.

Таким чином, перетворення даних в інший формат, агрегування даних або створення нових змінних може спростити аналіз і розуміння складних даних, підвищити точність і виявити відомості, які не очевидні з вихідних даних.

2.2.1.5 Інструменти для створення візуалізацій

Програмне забезпечення для аналізу даних включає інструменти візуалізації, тому що візуалізація даних може допомогти користувачам краще зрозуміти закономірності та взаємозв'язки в даних. Ось кілька причин, через які інструменти візуалізації важливі в програмному забезпеченні для аналізу даних:

1. Дослідження даних. Інструменти візуалізації можуть допомогти користувачам досліджувати дані, даючи їм змогу легко виявляти закономірності та взаємозв'язки, які можуть бути неочевидними під час перегляду необроблених даних. За допомогою візуалізацій користувачі можуть швидко виявляти тенденції, викиди та інші закономірності, які можуть бути важливими для аналізу.
2. Комунікація: Візуалізація - це ефективний спосіб поділитися ідеями та висновками з іншими. Візуалізації можна використовувати для подання складних даних у зручній для розуміння формі, що дає змогу користувачам чітко та переконливо передавати інформацію та висновки іншим.
3. Ухвалення рішень. Візуалізації можуть допомогти користувачам ухвалювати більш обґрунтовані рішення, надаючи чітке та інтуїтивно зрозуміле представлення даних. За допомогою візуалізацій користувачі можуть виявляти тенденції та закономірності, які можуть мати відношення до їх процесу прийняття рішень.
4. Ефективність. Інструменти візуалізації можуть допомогти користувачам працювати більш ефективно, даючи їм змогу швидко і легко виявляти закономірності та взаємозв'язки в даних. Це може допомогти користувачам заощадити час і ресурси, зосередившись на найважливіших аспектах даних.
5. Креативність. Інструменти візуалізації можуть надихнути на творчий підхід до аналізу даних, надаючи користувачам низку можливостей для

представлення даних. Використовуючи візуалізації, користувачі можуть експериментувати з різними уявленнями даних, що призводить до нових знань і відкриттів.

Таким чином, програмне забезпечення для аналізу даних містить інструменти візуалізації, тому що вони допомагають користувачам досліджувати дані, обмінюватися ідеями, ухвалювати більш виважені рішення, працювати більш ефективно і надихати на творчість. За допомогою візуалізацій користувачі можуть глибше зрозуміти дані та ухвалювати більш обґрунтовані рішення на основі отриманих відомостей.

2.2.1.6 Аналіз даних з використанням статистичних алгоритмів та алгоритмів машинного навчання

Аналіз даних з використанням статистичних алгоритмів і алгоритмів машинного навчання - це процес використання математичних моделей для вилучення ідей і знань з даних. Ось деякі ключові характеристики аналізу даних з використанням статистичних алгоритмів і алгоритмів машинного навчання:

1. Статистичні алгоритми. Статистичні алгоритми використовуються для аналізу даних на основі математичних моделей, що описують відносини між змінними. Ці алгоритми часто використовують для прогнозування, виявлення тенденцій або перевірки гіпотез.
2. Алгоритми машинного навчання. Алгоритми машинного навчання використовуються для створення моделей, які можуть вчитися на даних і робити прогнози або приймати рішення без явного програмування. Ці алгоритми часто використовуються для класифікації, кластеризації та прогнозування.
3. Попереднє опрацювання даних. Перш ніж дані можна буде проаналізувати за допомогою статистичних алгоритмів або алгоритмів машинного навчання, їх необхідно попередньо обробити. Це включає в себе очищення, перетворення і вибір відповідних даних.

4. Вибір моделі. Вибір алгоритму статистичного або машинного навчання залежить від розв'язуваної проблеми, доступних даних і бажаного результату. Найбільш підходящу модель вибирають на основі конкретних вимог аналізу.
5. Оцінка моделі. Після розроблення моделі її необхідно оцінити, щоб переконатися, що вона добре працює з новими даними. Це включає тестування моделі на перевірочному наборі або використання методів перехресної перевірки.
6. Інтерпретація. Останнім етапом аналізу даних з використанням алгоритмів статистичного або машинного навчання є інтерпретація. Розуміння і знання, отримані в результаті аналізу, використовуються для ухвалення рішень або для глибшого розуміння процесів, що лежать в основі.

Таким чином, аналіз даних з використанням статистичних алгоритмів і алгоритмів машинного навчання передбачає використання математичних моделей для вилучення інформації та знань з даних. Цей процес охоплює попередню обробку даних, вибір моделі, оцінку моделі та інтерпретацію. Використовуючи ці методи, можна отримати глибше розуміння даних і ухвалювати більш обґрунтовані рішення на основі отриманої інформації.

2.2.1.7 Інтерпретація даних

Інтерпретація даних - це процес аналізу даних, який дає змогу витягти значущі висновки і зробити корисні висновки на основі результатів аналізу. Ось кілька ключових аспектів інтерпретації даних і процесу висновків:

1. Розуміння контексту: Перед тим, як можна зробити якісь висновки, важливо зрозуміти контекст, у якому були отримані дані. Це може включати розуміння вихідних даних, причин їхнього збору і мети аналізу.
2. Постановка запитань: Щоб зробити корисні висновки з даних, важливо ставити правильні запитання. Це може включати визначення ключових чинників, що впливають на досліджуваний процес або явище, та ідентифікацію можливих взаємозв'язків між цими чинниками.

3. Візуалізація даних: Візуалізація даних - це важливий аспект інтерпретації даних. Візуалізації можуть допомогти зрозуміти загальну структуру даних і виявити будь-які цікаві тренди або взаємозв'язки.
4. Аналіз даних: Після того, як дані були очищені та візуалізовані, можна приступити до аналізу даних. Це може включати різні методи, такі як статистичний аналіз, машинне навчання, дослідження тексту та інші.
5. Зробити висновки: Наприкінці процесу інтерпретації даних необхідно зробити висновки. Це може включати узагальнення результатів аналізу, оцінку значущості отриманих результатів і висновки про те, як ці результати можна використовувати для вирішення бізнес-завдань.

Правильне виконання цих кроків може допомогти у вилученні найкориснішої інформації з даних і забезпечити розуміння їх значення. Однак важливо пам'ятати, що інтерпретація даних не є точною наукою, і будь-які висновки мають бути зроблені з урахуванням можливих обмежень і вихідних даних.

2.2.1.8 Інструменти для створення звітів та обміну результатами з іншими

Програмне забезпечення для аналізу даних зазвичай містить інструменти для створення звітів та обміну результатами з іншими, оскільки кінцевою метою аналізу даних часто є передача ідей і висновків зацікавленим сторонам. Ось кілька причин, через які звітність та обмін інформацією важливі:

1. Обмін ідеями: аналіз даних часто охоплює складні набори даних і статистичний аналіз, які можуть бути важкими для розуміння неспеціалістами. Інструменти звітності можуть допомогти перетворити ці висновки в прості для розуміння формати, якими можна поділитися із зацікавленими сторонами.
2. Співпраця: аналіз даних часто є спільним процесом за участю кількох зацікавлених сторін, включно з аналітиками даних, бізнес-аналітиками та керівниками. Інструменти звітності можуть допомогти полегшити співпрацю, надаючи зацікавленим сторонам централізоване місце для доступу та обговорення результатів.

3. Ухвалення рішень: аналіз даних часто використовується для обґрунтування прийняття рішень. Інструменти звітності можуть допомогти особам, які приймають рішення, швидко отримати доступ до основних висновків та аналітичних даних, що допоможе їм ухвалювати більш обґрунтовані рішення.
4. Відстеження прогресу: інструменти звітності можна використовувати для відстеження прогресу з плином часу. Наприклад, бізнес може використовувати аналіз даних для відстеження успіху маркетингової кампанії. Інструменти звітності можна використовувати для обміну звітами про хід роботи із зацікавленими сторонами під час кампанії.

Таким чином, інструменти звітності та обміну є важливими компонентами програмного забезпечення для аналізу даних, оскільки вони допомагають перетворювати складні дані на зрозумілі ідеї та полегшують спільну роботу й ухвалення рішень.

2.3 Веб-скрейпінг (web scraping)

2.3.1 Інструменти веб-скрейпінгу

Інструменти веб-скрейпінгу - це програми, які дають змогу користувачам витягувати дані з веб-сайтів. Ось кілька прикладів часто використовуваних інструментів веб-скрейпінгу:

1. BeautifulSoup: бібліотека Python для перегляду веб-сторінок, що спрощує навігацію та аналіз документів HTML і XML.
2. Scrappy: платформа Python для парсингу веб-сторінок, яка дає змогу користувачам створювати та налаштовувати парсери веб-сторінок.
3. Selenium: інструмент автоматизації веб-браузера, який можна використовувати для очищення і тестування веб-сторінок.
4. Octoparse: інструмент веб-скрейпінгу, який дає змогу користувачам витягувати дані з веб-сайтів без будь-яких навичок програмування.
5. WebHarvy: візуальний парсер, який дає змогу користувачам вказувати і натискати на дані, які вони хочуть витягти з веб-сайтів.

6. Import.io: інструмент веб-скрейпінгу, який дає змогу користувачам витягувати дані з веб-сайтів і перетворювати їх на структуровані дані.
7. Parsehub: інструмент веб-скрейпінгу, який дає змогу користувачам витягувати дані з веб-сайтів за допомогою алгоритмів машинного навчання.

Це всього лише кілька прикладів з безлічі інструментів веб-скрейпінгу, доступних користувачам. Вибір інструменту залежатиме від конкретних потреб і вимог користувача.

2.3.2 Як працюють Інструменти веб-скрейпінгу

Інструменти веб-скрейпінгу - це програми, які дають змогу користувачам витягувати дані з веб-сайтів. Ось докладне пояснення того, як працюють інструменти веб-скрейпінгу:

1. Надсилати HTTP-запити: інструмент веб-скрейпінгу надсилає HTTP-запити на цільовий веб-сайт для отримання HTML-вмісту веб-сторінки.
2. Аналіз вмісту HTML. Потім інструмент аналізує вміст HTML за допомогою таких мов програмування, як Python, JavaScript або PHP. Цей процес охоплює визначення структури та місця розташування даних, які необхідно витягти з тегів HTML.
3. Витяг даних: інструмент витягує необхідні дані з проаналізованого вмісту HTML за допомогою бібліотек програмування, таких як BeautifulSoup або lxml у Python.
4. Збереження даних. Після вилучення даних їх зберігають у структурованому форматі, як-от CSV або JSON, або в базі даних.
5. Обробка розбивки на сторінки та динамічного вмісту. У випадках, коли веб-сторінки складаються з декількох сторінок, інструмент обробляє розбивку на сторінки і витягує дані з кожної сторінки. Деякі інструменти веб-скрейпінгу також можуть обробляти динамічний контент, наприклад дані, які завантажуються через запити AJAX.
6. Вжиття заходів проти парсингу: на деяких веб-сайтах можуть бути передбачені заходи для запобігання парсингу, такі як CAPTCHA або

блокування IP-адрес. Інструменти веб-скрейпінгу можуть мати вбудовані функції для опрацювання цих заходів або можуть вимагати додаткового налаштування для їхнього обходу.

Загалом, інструменти веб-скрейпінгу автоматизують процес вилучення даних із веб-сайтів, відправляючи HTTP-запити, аналізуючи HTML-контент, а також витягуючи та зберігаючи дані в структурованому форматі. Конкретний процес і функції інструментів веб-скрейпінгу можуть відрізнятися залежно від використовуваного інструменту і вимог користувача.

2.3.3 Надсилання HTTP-запитів на цільовий веб-сайт

Коли користувач заходить на веб-сайт у своєму веб-браузері, він надсилає HTTP-запит на веб-сервер, на якому розміщено веб-сайт. HTTP означає протокол передачі гіпертексту, який являє собою протокол, що використовується для передачі даних через Інтернет. HTTP-запит містить таку інформацію, як URL-адреса веб-сайту, до якого ви намагаєтесь отримати доступ, будь-які дані, які ви хочете надіслати на веб-сайт, та інші заголовки запити.

У контексті парсингу веб-сторінок надсилання HTTP-запитів на цільовий веб-сайт означає використання інструментів парсингу веб-сторінок для надсилання HTTP-запитів на веб-сайт для вилучення його HTML-контенту. Після вилучення HTML-контенту інструмент веб-скрейпінгу може проаналізувати його, щоб витягти необхідні дані.

Надсилання HTTP-запитів - це перший крок у веб-скрейпінгу, і для цього потрібне знання URL-адреси цільового веб-сайту і того, як структурувати HTTP-запит для отримання бажаного контенту. Більшість інструментів веб-скрейпінгу автоматизують цей процес і надають зручний інтерфейс для надсилання HTTP-запитів і вилучення HTML-контенту з цільового веб-сайту.

2.3.4 Аналіз вмісту HTML

HTML - це мова розмітки, що використовується для створення веб-сторінок, і це стандартний формат, у якому веб-контент представлений в Інтернеті. Синтаксичний аналіз HTML-контенту включає в себе аналіз HTML-коду для вилучення з нього певних даних та інформації. Такі мови програмування, як Python, JavaScript або PHP, можна використовувати для аналізу вмісту HTML, оскільки в них є бібліотеки та інструменти, спеціально призначені для цієї мети.

Ось кілька причин, чому синтаксичний аналіз вмісту HTML за допомогою мов програмування корисний:

1. Витяг структурованих даних: HTML-код може бути важким для читання і розуміння, особливо під час роботи з великими обсягами даних. Аналізуючи вміст HTML за допомогою мов програмування, дані можна витягувати і структурувати в більш читабельному і зрозумілому форматі.
2. Автоматизація процесу синтаксичного аналізу. Аналіз вмісту HTML вручну може зайняти багато часу і призвести до помилок. Мови програмування можуть автоматизувати процес аналізу вмісту HTML, роблячи його більш ефективним і точним.
3. Гнучкість: мови програмування пропонують більшу гнучкість під час аналізу HTML-контенту, ніж традиційні інструменти веб-скрейпінгу. Вони можуть обробляти динамічний контент, сторінки з візуалізацією JavaScript та інші складні веб-структури, які складно проаналізувати за допомогою традиційних інструментів.
4. Налаштування: мови програмування дають змогу налаштовувати процес синтаксичного аналізу відповідно до конкретних потреб і вимог. Це дає змогу точніше і цілеспрямованіше витягувати дані.

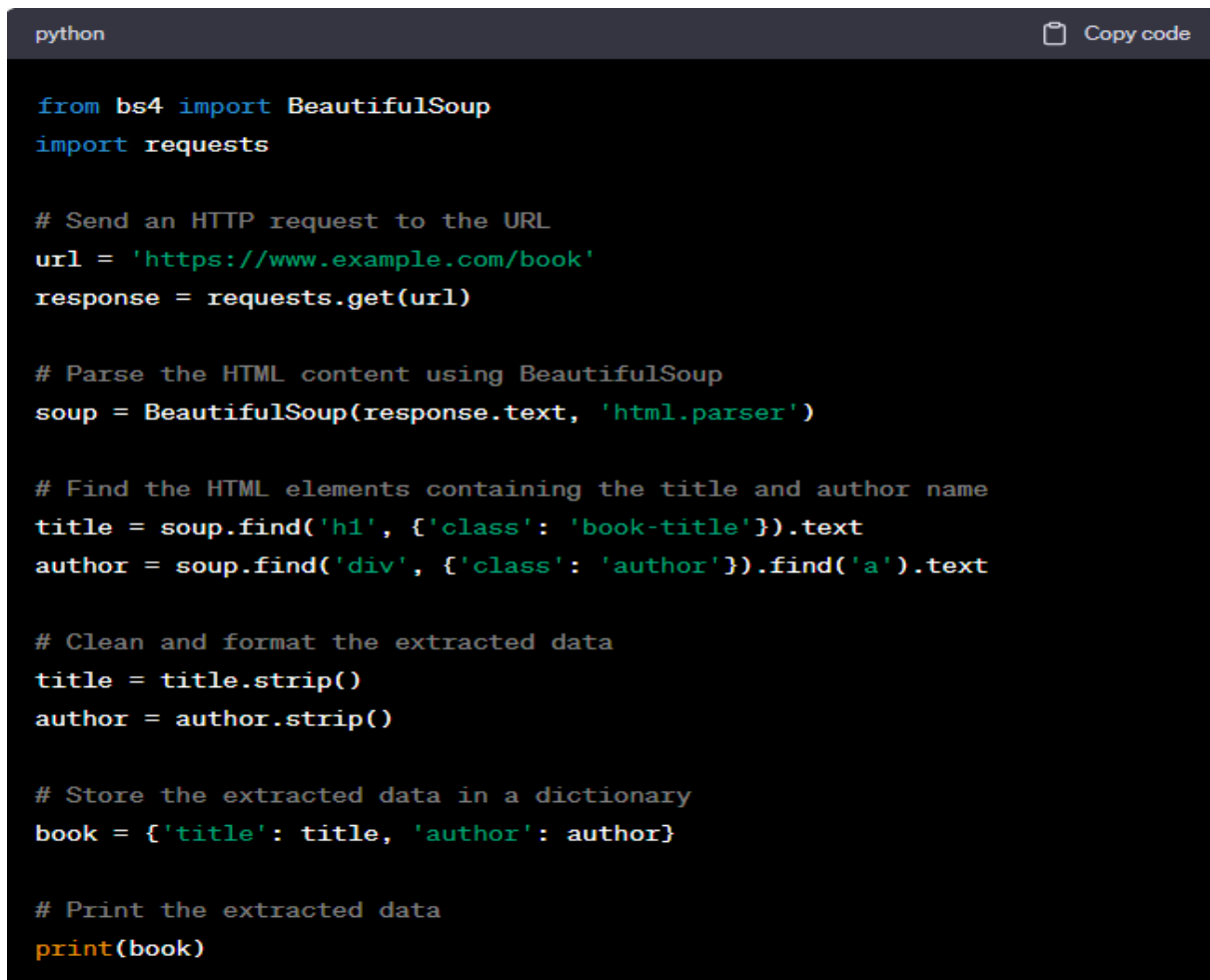
Загалом, синтаксичний аналіз вмісту HTML за допомогою таких мов програмування, як Python, JavaScript або PHP, є ефективним способом вилучення даних із веб-сторінок. Він забезпечує гнучкість, автоматизацію, налаштування та вилучення структурованих даних.

2.3.5 Процес вилучення необхідних даних

Процес вилучення необхідних даних із проаналізованого вмісту HTML за допомогою бібліотек програмування, таких як BeautifulSoup або lxml у Python. Після аналізу вмісту HTML за допомогою таких бібліотек, як BeautifulSoup або lxml у Python, наступним кроком буде вилучення необхідних даних із проаналізованого вмісту HTML. Ось процес вилучення даних:

1. Знайдіть елементи HTML, які містять дані, які ви хочете витягти: це включає використання тегів елементів, класів або ідентифікаторів для пошуку елементів HTML, що містять потрібні дані.
2. Витяг даних з HTML-елементів. Після визначення HTML-елементів, що містять дані, наступним кроком буде вилучення з них даних. Це передбачає використання методів, що надаються бібліотекою синтаксичного аналізу, для вилучення тексту або атрибутів елементів HTML.
3. Очистіть і відформатуйте витягнуті дані. Витягнуті дані можуть містити небажані символи або форматування, які необхідно видалити або змінити. Це включає в себе використання функцій обробки рядків або регулярних виразів для очищення і форматування витягнутих даних.
4. Зберігайте витягнуті дані. Нарешті, витягнуті дані необхідно зберігати в структурованому форматі, CSV , JSON, або в базі даних для подальшого аналізу чи обробки.

Ось приклад фрагмента коду, що використовує BeautifulSoup для вилучення заголовка та імені автора з проаналізованої HTML-сторінки:



```
python                                                                    Copy code

from bs4 import BeautifulSoup
import requests

# Send an HTTP request to the URL
url = 'https://www.example.com/book'
response = requests.get(url)

# Parse the HTML content using BeautifulSoup
soup = BeautifulSoup(response.text, 'html.parser')

# Find the HTML elements containing the title and author name
title = soup.find('h1', {'class': 'book-title'}).text
author = soup.find('div', {'class': 'author'}).find('a').text

# Clean and format the extracted data
title = title.strip()
author = author.strip()

# Store the extracted data in a dictionary
book = {'title': title, 'author': author}

# Print the extracted data
print(book)
```

Рисунок 2.1 – Фрагмент коду , що використовує BeautifulSoup

У цьому прикладі ми спочатку відправляємо HTTP-запит на цільовий вебсайт, аналізуємо HTML-контент за допомогою BeautifulSoup, знаходимо HTML-елементи, які містять заголовок та ім'я автора, очищаємо та форматуємо витягнуті дані, зберігаємо їх у словнику, а потім роздруковуємо.

2.3.6 Процес збереження даних

Процес збереження даних включає збереження витягнутих або оброблених даних у структурованому форматі для подальшого аналізу, оброблення або спільного використання. Важливість збереження даних полягає в тому, що це дозволяє нам:

1. Зберігати дані для майбутнього використання: зберігаючи дані, ми гарантуємо, що вони будуть доступні для майбутнього аналізу або обробки, навіть якщо первинне джерело даних більше недоступне.

2. Діліться даними з іншими: Збереження даних у структурованому форматі дає нам змогу ділитися ними з іншими, наприклад із колегами або зацікавленими сторонами, яким можуть знадобитися дані для своїх цілей.
3. Аналізуйте дані за допомогою інших інструментів: зберігаючи дані в структурованому форматі, ми можемо імпортувати їх в інші інструменти аналізу або візуалізації даних, як Excel, R або Tableau, для подальшого аналізу.
4. Створіть сховище даних: зберігаючи дані з плином часу, ми можемо створити сховище даних, яке можна використовувати для відстеження тенденцій, виявлення закономірностей або прогнозування.

Процес збереження даних може включати експорт даних у формат файлу, CSV, Excel, JSON або SQL, або їх збереження в базі даних. Вибір формату або механізму зберігання залежатиме від характеру даних, передбачуваного використання даних і доступних інструментів і ресурсів.

2.3.7 Процес опрацювання розбивки на сторінки та динамічного вмісту

Обробка нумерації сторінок і динамічного вмісту у веб-скрапінгу передбачає стратегії навігації кількома сторінками даних і роботу з вмістом, який динамічно змінюється залежно від дій користувача або інших чинників. Ось загальні кроки для обробки нумерації сторінок і динамічного вмісту:

1. Визначте схему розбивки на сторінки: визначте шаблон або механізм, який використовується веб-сайтом для відображення декількох сторінок даних. Це можуть бути пронумеровані посилання, кнопки "наступний/попередній" або функція прокрутки вниз.
2. Ітерація по сторінках: щойно схему розбивки на сторінки визначено, напишіть код для ітерації по сторінках, збираючи потрібні дані на кожній сторінці. Це може включати надсилання кількох HTTP-запитів і обробку вмісту HTML з кожної сторінки.
3. Обробка динамічного вмісту. Деякі веб-сайти використовують динамічний вміст, який змінюється в міру того, як користувач взаємодіє зі сторінкою.

Це можуть бути спливаючі вікна, меню, що випадають, або форми. Для обробки динамічного вмісту вам може знадобитися використовувати такі інструменти, як Selenium або Scrapy-Splash, для імітації взаємодії з користувачем і захоплення отриманого вмісту.

4. Керування продуктивністю: обробка розбивки на сторінки і динамічного вмісту може займати багато часу і ресурсів. Для управління продуктивністю розгляньте такі стратегії, як обмеження кількості запитів на секунду, кешування даних, коли це можливо, і використання паралельної обробки для прискорення процесу очищення.

Загалом, обробка нумерації сторінок і динамічного контенту вимагає пильної уваги до конкретного веб-сайту, що очищається, і зібраних даних. Він часто містить комбінацію методів парсингу веб-сторінок, мов програмування і сторонніх інструментів для ефективної навігації по складних веб-сторінках і збору потрібних даних.

2.3.8 Процес обробки динамічного контенту

Процес динамічного опрацювання контенту охоплює стратегії роботи з веб-сторінками, що містять контент, який динамічно змінюється або оновлюється у відповідь на дії користувача, як-от натискання кнопки, прокручування сторінки або заповнення форми. Ось загальні кроки для обробки динамічного контенту у веб-скрапінгу:

1. Визначте динамічний вміст: визначте, які частини веб-сторінки містять динамічний вміст, що може змінюватися залежно від взаємодії з користувачем або інших чинників. Це можуть бути спливаючі вікна, меню, що випадають, або форми.
2. Моделювання взаємодії з користувачем: для захоплення динамічного контенту вам може знадобитися використовувати такі інструменти, як Selenium або Scrapy-Splash, для імітації взаємодії користувача зі сторінкою. Наприклад, ви можете використовувати Selenium, щоб клацнути кнопку або ввести текст у поле форми, а потім захопити отриманий вміст.

3. Вилучення та обробка даних. Після захоплення динамічного вмісту ви можете використовувати бібліотеки синтаксичного аналізу, як BeautifulSoup або lxml у Python, для вилучення та обробки відповідних даних. Це може включати фільтрацію непотрібної інформації або очищення і перетворення даних для подальшого аналізу.
4. Управління продуктивністю: обробка динамічного вмісту може займати багато часу і ресурсів, особливо для великих або складних веб-сторінок. Для управління продуктивністю розгляньте такі стратегії, як обмеження кількості запитів на секунду, кешування даних, коли це можливо, і використання паралельної обробки для прискорення процесу очищення.

Загалом, обробка динамічного контенту у веб-скрейпінгу вимагає поєднання методів веб-скрейпінгу, мов програмування і сторонніх інструментів для ефективного збору та обробки потрібних даних. Щоб вибрати найбільш підходящий підхід, важливо ретельно вивчити конкретний веб-сайт, який піддається парсингу, і дані, що збираються.

2.3.9 Вжиття заходів проти парсингу

Вжиття заходів проти веб-скрапінгу охоплює вжиття заходів щодо запобігання або перешкоджання веб-скрапінгу на веб-сайті. Ось деякі поширені методи, що використовуються для запобігання парсингу веб-сторінок:

1. Robots.txt: вебсайти можуть використовувати файл robots.txt, щоб вказати, які сторінки та каталоги можуть скануватися пошуковими системами та іншими ботами. Цей файл можна використовувати для блокування доступу роботів-скрейперів до певних сторінок або розділів веб-сайту.
2. Блокування IP-адрес: веб-сайти можуть ідентифікувати IP-адреси парсинг-ботів і блокувати їм доступ до сайту. Однак цей підхід можна обійти, використовуючи проксі або змінюючи IP-адреси.
3. CAPTCHA: веб-сайти можуть використовувати CAPTCHA, щоб переконатися, що користувач є людиною, а не ботом. Це може бути ефективним

способом запобігання автоматичному парсингу, але також може створювати проблеми для законних користувачів.

4. Обмеження швидкості: веб-сайти можуть обмежувати кількість запитів, які можуть бути зроблені протягом певного періоду часу. Це може допомогти запобігти перевантаженню сервера запитами парсинг-ботів.

5. Судові позови: у деяких випадках власники веб-сайтів можуть подати до суду на парсерів, які порушують їхні умови обслуговування або порушують їх права на інтелектуальну власність.

Загалом, вжиття заходів проти парсингу веб-сторінок містить поєднання технічних, правових та адміністративних заходів для запобігання та протидії парсингу. Конкретні заходи, що використовуються, залежатимуть від веб-сайту, що очищається, даних, що збираються, і цілей власника веб-сайту.

2.4 Звіти про дослідження ринку

2.4.1 Структура та зміст звітів про дослідження ринку

Звіти про дослідження ринку зазвичай включають такі розділи:

1. Резюме: короткий виклад основних висновків звіту, включно з ключовими ідеями та рекомендаціями.
2. Вступ: огляд мети звіту, методології, використаної для проведення дослідження, та обсягу дослідження.
3. Огляд ринку: детальний аналіз розміру ринку, темпів зростання і тенденцій, включно з оцінкою конкурентного середовища на ринку і ключових гравців, що працюють на ринку.
4. Аналіз галузі: огляд історичних і поточних показників галузі, включно з аналізом основних рушійних сил і проблем, що стоять перед галуззю.
5. Сегментація ринку: розбивка ринку на різні сегменти залежно від типу продукту, застосування, кінцевого користувача і географії.
6. Аналіз частки ринку: Докладний аналіз частки ринку провідних гравців ринку.

7. Конку rentне середовище: всебічний аналіз конкурентного середовища, включно з основними гравцями на ринку, їх часткою на ринку та їхніми ключовими стратегіями.
8. Ключові тенденції ринку: аналіз ключових тенденцій, що визначають ринок, включно з технологічними досягненнями, зміною споживчих переваг і розвитком нормативно-правової бази.
9. SWOT-аналіз: огляд сильних і слабких сторін ринку, можливостей і загроз.
10. Висновок: Короткий виклад основних висновків звіту, включно з рекомендаціями для майбутніх досліджень і дій.
11. Додаток: набір додаткових даних та інформації, що підтверджують основні висновки звіту.

Загалом звіти про дослідження ринку зазвичай структуровані таким чином, щоб читачі могли швидко і легко отримати доступ до інформації, необхідної їм для прийняття обґрунтованих рішень про свій бізнес. Зміст звіту має бути чітким, стислим і добре організованим, а дані мають бути подані у вигляді таблиць, діаграм та інших візуальних форматів для полегшення розуміння.

2.4.2 Приклади звітів компанії Gartner

Gartner - провідна дослідницька та консалтингова компанія, яка випускає широкий спектр звітів про дослідження ринку в різних галузях і галузях технологій. Ось кілька прикладів звітів Gartner:

1. Магічний квадрант Gartner: у цьому звіті представлено оцінку постачальників на конкретному ринку на основі їх повноти бачення і здатності реалізувати. Звіти Gartner Magic Quadrant охоплюють широкий спектр технологічних ринків, включно з хмарними обчисленнями, управлінням даними, кібербезпекою та багато іншого.
2. Gartner Hype Cycle: у цьому звіті представлено графічне представлення зрілості та впровадження нових технологій, а також аналіз потенційних переваг і ризиків цих технологій. Звіти Gartner Hype Cycle охоплюють

широкий спектр технологічних галузей, включно зі штучним інтелектом, блокчейном та Інтернетом речей (IoT).

3. Gartner Critical Capabilities: у цьому звіті представлено аналіз постачальників на конкретному ринку на основі набору критичних можливостей або варіантів використання. Звіти Gartner Critical Capabilities охоплюють широкий спектр технологічних ринків, включно з управлінням корпоративним контентом, цифровою комерцією та управлінням взаємовідносинами з клієнтами (CRM).
4. Gartner Market Guide: у цьому звіті представлено огляд конкретного ринку технологій, включно з основними тенденціями, проблемами та постачальниками. Звіти Gartner Market Guide охоплюють широкий спектр технологічних галузей, включно з хмарною безпекою, розробкою мобільних додатків і автоматизацією бізнес-процесів.
5. Gartner Peer Insights: у цьому звіті представлені відгуки клієнтів і рейтинги постачальників програмного забезпечення на різних технологічних ринках. Звіти Gartner Peer Insights охоплюють широкий спектр технологічних галузей, включно з плануванням ресурсів підприємства (ERP), взаємодією з клієнтами (CX) і мережевою безпекою.

Загалом звіти Gartner містять цінну інформацію та рекомендації для компаній, які прагнуть ухвалювати обґрунтовані рішення про інвестиції в технології та стратегії.

2.4.3 Приклади звітів компанії Forrester

Forrester - провідна дослідницька та консалтингова компанія, яка випускає широкий спектр звітів про дослідження ринку в різних галузях і галузях технологій. Ось кілька прикладів звітів Forrester:

1. Forrester Wave™: у цьому звіті представлено оцінку постачальників на конкретному ринку на основі їх поточних пропозицій, стратегії та присутності на ринку. Звіти Forrester Wave™ охоплюють широкий спектр

технологічних ринків, включно з управлінням взаємовідносин з клієнтами, автоматизацією маркетингу та управлінням даними.

2. Forrester Now Tech: у цьому звіті представлено огляд постачальників на певному ринку, включно з їхніми можливостями та потенційною цінністю для покупців. Звіти Forrester Now Tech охоплюють широкий спектр технологічних галузей, включно з хмарною безпекою, цифровими платформами та досвідом співробітників.
3. Forrester Total Economic Impact™: у цьому звіті представлено аналіз потенційної окупності інвестицій (ROI) для конкретних інвестицій у технології, включно з витратами, вигодами та ризиками. Звіти Forrester Total Economic Impact™ охоплюють широкий спектр технологічних ринків, включно з програмним забезпеченням для спільної роботи, кібербезпекою та якістю обслуговування клієнтів.
4. Forrester New Wave™: у цьому звіті представлено оцінку нових постачальників на конкретному ринку на основі їхніх технологічних інновацій та ринкового потенціалу. Звіти Forrester New Wave™ охоплюють широкий спектр технологічних галузей, включно з автоматизацією бізнес-процесів за допомогою ШІ, управління відгуками клієнтів і управління хмарними даними.
5. Forrester Industry Wave™: у цьому звіті представлено оцінку постачальників у певній галузі на основі їх поточних пропозицій, стратегії та присутності на ринку. Звіти Forrester Industry Wave™ охоплюють широкий спектр галузей, включно з охороною здоров'я, фінансовими послугами та роздрібною торгівлею.

Загалом, звіти Forrester містять цінну інформацію та рекомендації для компаній, які прагнуть ухвалювати більш обґрунтовані рішення про інвестиції в технології та стратегії.

2.4.4 Приклади звітів компанії IDC

IDC (International Data Corporation) - глобальний постачальник звітів про дослідження ринку та консультаційних послуг у різних галузях і сферах технологій. Ось кілька прикладів звітів IDC:

1. IDC MarketScape: у цьому звіті представлено оцінку постачальників на конкретному ринку з урахуванням їх можливостей, стратегій і конкурентного середовища. Звіти IDC MarketScape охоплюють широкий спектр технологічних ринків, включно з хмарними обчисленнями, кібербезпекою та цифровою трансформацією.
2. IDC FutureScape: у цьому звіті представлено 10 основних прогнозів для конкретної галузі або технологічної галузі на найближчі кілька років. Звіти IDC FutureScape охоплюють широкий спектр галузей, включно з охороною здоров'я, виробництвом і фінансовими послугами.
3. IDC Tracker: у цьому звіті представлено докладні дані та аналіз розміру, зростання і тенденцій на конкретному ринку або в технологічній галузі. Звіти IDC Tracker охоплюють широкий спектр ринків, включно зі смартфонами, переносними пристроями та ПК.
4. IDC TechScape: у цьому звіті представлено оцінку зрілості та впровадження конкретних технологій на ринку або в галузі. Звіти IDC TechScape охоплюють широкий спектр технологій, включно з блокчейном, штучним інтелектом і хмарними технологіями.
5. Профіль постачальника IDC. У цьому звіті представлено аналіз стратегії, сильних і слабких сторін конкретного постачальника на ринку або в технологічній галузі. Звіти IDC Vendor Profile охоплюють широке коло постачальників, включно з Microsoft, IBM і Amazon Web Services.

Загалом звіти IDC містять цінну інформацію та рекомендації для компаній, які прагнуть ухвалювати обґрунтовані рішення про інвестиції в технології та стратегії.

2.5. Системи електронної точки продажу (EPOS)

2.5.1 Концепція роботи EPOS

Системи електронної торгової точки (EPOS) - це комп'ютеризовані системи, які використовуються роздрібними торговцями для оброблення транзакцій продажів, управління запасами та виконання інших бізнес-функцій. Ці системи зазвичай складаються з апаратних і програмних компонентів, які працюють разом, щоб забезпечити повне рішення для управління роздрібною торгівлею.

Апаратний компонент системи EPOS зазвичай містить комп'ютер, монітор із сенсорним екраном, сканер штрих-коду, касовий апарат та інші периферійні пристрої, як-от дисплей покупця або кард-рідер. Програмний компонент надає користувацький інтерфейс і функціональні можливості, необхідні для управління транзакціями продажів, запасами та іншими бізнес-процесами.

Системи EPOS можуть виконувати безліч функцій, зокрема:

1. **Обробка продажів:** системи EPOS дають змогу роздрібним торговцям швидко й ефективно обробляти транзакції продажів. Продажі можна обробляти за допомогою сканування штрих-коду, ручного введення або вибору на сенсорному екрані.
2. **Управління запасами:** системи EPOS можуть допомогти роздрібним торговцям управляти запасами, відстежуючи рівень запасів, створюючи замовлення на купівлю і надаючи оновлення запасів у режимі реального часу.
3. **Управління клієнтами:** системи EPOS можуть допомогти роздрібним продавцям управляти інформацією про клієнтів, такою як контактні дані та історія покупок.
4. **Звітність:** системи EPOS можуть генерувати низку звітів, які допомагають роздрібним торговцям аналізувати дані про продажі, виявляти тенденції та відстежувати ефективність.

5. Інтеграція з іншими системами: системи EPOS можуть інтегруватися з іншими системами, такими як бухгалтерське програмне забезпечення, платформи електронної комерції та платіжні шлюзи.

Загалом системи EPOS надають підприємствам роздрібної торгівлі комплексне рішення для управління транзакціями продажів, запасами та іншими бізнес-процесами, допомагаючи їм працювати більш ефективно і результативно.

2.5.2 Найбільш відомі системи EPOS

На ринку доступно кілька систем EPOS, кожна з яких має свої особливості та можливості. Деякі з найпопулярніших систем EPOS містять у собі:

1. Square - це популярна система EPOS, призначена для малого та середнього бізнесу. Він пропонує низку функцій, включно з обробкою продажів, управлінням запасами та управлінням клієнтами.
2. Shopify - це платформа електронної комерції, яка також пропонує систему EPOS. Він призначений для роздрібних продавців, які працюють як в Інтернеті, так і в автономному режимі, і надає такі функції, як обробка продажів, управління запасами та управління клієнтами.
3. Lightspeed - це система EPOS, призначена для підприємств роздрібної торгівлі у сфері гостинності та роздрібної торгівлі. Він пропонує такі функції, як керування таблицями, планування зустрічей і керування клієнтами.
4. Vend - це система EPOS, розроблена для роздрібних продавців усіх розмірів. Він пропонує такі функції, як опрацювання продажів, управління запасами та управління клієнтами, і може використовуватися як на настільних, так і на мобільних пристроях.
5. Toast - це система EPOS, розроблена для індустрії гостинності. Він пропонує такі функції, як управління таблицями, обробка замовлень і управління запасами.
6. Clover - це система EPOS, призначена для малого та середнього бізнесу. Він пропонує такі функції, як обробка продажів, управління запасами та

управління клієнтами, і може використовуватися як на настільних, так і на мобільних пристроях.

Це всього лише кілька прикладів з безлічі систем EPOS, доступних на ринку, і підприємства повинні ретельно оцінити свої потреби і вимоги, перш ніж вибрати відповідну для них систему.

3. МЕТОДОЛОГІЧНІ І ПРАКТИЧНІ ПІДХОДИ ДО ФОРМУВАННЯ ТЕХНОЛОГІЇ АНАЛІЗУ ЦІН ЕЛЕКТРОННИХ ТОВАРІВ ТА ЙОГО РЕАЛІЗАЦІЇ

Факторний аналіз — це статистичний метод, який зазвичай використовується для виявлення основних закономірностей або факторів, що пояснюють зміну набору змінних. У контексті аналізу цін на електронні товари факторний аналіз може використовуватися для виявлення ключових рухомих цінових тенденцій та обґрунтування маркетингових стратегій. Нижче наведені ключові кроки, які слід враховувати при використанні факторного аналізу для аналізу цін на електронні товари:

1. Визначення цілей: перший крок — визначити проблему та визначити змінні, які будуть включені в аналіз. Це може включати такі змінні, як характеристики продукту, репутація бренду, канали збуту та ціни конкурентів.
2. Збір даних. Наступним кроком є збір даних про змінні, визначених на етапі 1. Це може включати збір даних про ціни з кількох джерел, проведення запитів або фокус-груп для збору відгуків користувачів або збір даних про стратегії ціноутворення конкурентів.
3. Попередня обробка даних: дані необхідно попередньо опрацювати, щоб переконатися, що вони підходять для факторного аналізу. Це може включати стандартизацію даних, виявлення та видалення однакового, а також перевірку відсутніх даних.
4. Факторний аналіз. Сам факторний аналіз включає в себе виявлення основних факторів, що визначають цінові тенденції. Це можна зробити за

допомогою таких методів, як аналіз основних компонентів або дослідницький факторний аналіз. Результат аналізу зазвичай включає набір факторних навантажень, які вказують на силу зв'язку між кожною змінною і кожним фактором.

5. Інтерпретація та застосування. Заключним кроком є інтерпретація результатів факторного аналізу та їх застосування для обґрунтування маркетингових стратегій. Це може включати в себе виявлення основних рухомих силових цінових тенденцій, розробку цільових маркетингових кампаній на основі споживчих переваг або корекцію стратегій ценоутворення на основі цінових конкурентів.

Факторний аналіз може бути корисним методом для аналізу цін на електронні товари з точки зору маркетингу. Визначивши основні фактори, що впливають на цінові тенденції, компанії можуть розробити більш ефективні маркетингові стратегії, засновані на споживчих перевагах і цінах конкурентів. Реалізація факторного аналізу вимагає ретельного планування та виконання а також може включати збір і попередню обробку даних, проведення самого факторного аналізу та інтерпретацію результатів для інформування маркетингових стратегій [13].

Важливо відзначити, що факторний аналіз слід проводити на репрезентативній вибірці цілого ринку, щоб гарантувати, що результати застосовані до більш широкої сукупності. Крім того, компаніям слід враховувати обмеження методу факторного аналізу, такі як можливість ложних кореляцій і необхідність збалансувати інтерпретованість з точністю.

Як тільки ключові фактори, що впливають на цінові тенденції, будуть визначені, підприємства можуть використовувати цю інформацію для обґрунтування стратегій ціноутворення та маркетингу. Наприклад, якщо факторний аналіз показує, що при прийнятті рішень щодо купівлі споживачі цінують характеристики продукту більше, ніж репутацію бренду, підприємства можуть скоригувати свою стратегію ценоутворення, щоб запропонувати більш низькі ціни на продукти з меншою кількістю функцій, або розробити цілі

маркетингові кампанії, які підтримують характеристики продукту, які цінують споживачі.

Факторний аналіз також може використовуватися для відстеження змін факторів, що визначають цінові тенденції з поточним часом. Регулярно проводячи факторний аналіз і відстежуючи зміни у факторних завантаженнях, компанії можуть виявити тенденції, що виникають, і відповідним чином скоригувати свої стратегії.

Статистичне програмне забезпечення для факторного аналізу

Потім ви можете використовувати статистичне програмне забезпечення, наприклад R або Python, для проведення факторного аналізу даних про ціни. Факторний аналіз допоможе вам визначити основні фактори, що впливають на зміну цін.

R і Python — це мова програмування та програмне середовище для статистичних вичислень і графіків. Він знаходиться у вільному доступі під Стандартною суспільною ліцензією GNU і використовується в наукових колах, дослідженнях і промисловості для широкого аналізу даних, статистичного моделювання та візуалізації.

R надає безліч статистичних і графічних методів, включаючи лінійне та нелінійне моделювання, класичні статистичні тести, аналіз тимчасових рядів, класифікацію, кластеризацію та багато іншого. Він також підтримує імпорт/експорт даних у різних форматах, маніпулювання даними та створення графіки.

R має велике й активне співтовариство користувачів із тисячами пакетів, доступних для завантаження з комплексних мережеских архівів R (CRAN) та інших репозиторіїв. Ці пакети розширюють функціональні можливості R за рахунок додаткових статистичних методів, інструментів обробки даних, бібліотек візуалізації та багато іншого.

В цілому R — це потужний і універсальний інструмент для аналізу даних, який використовується в багатьох, включаючи фінанси, охорону здоров'я, соціальні науки та інші.

3.1 Формування факторного аналізу цінових тенденцій з позиції маркетингу

Як було зазначено вище, факторний аналіз — це статистичний метод, який використовується для виявлення основних факторів, що пояснюють кореляції між набором змінних [6].

При факторному аналізі ми дивимося на кореляції між змінними і намагаємося визначити закономірності в цих кореляціях, які передбачають наявність лежачих в їх основі факторів. Ці фактори відображають загальні риси між змінними, що визначають їх взаємозв'язок. Другими словами, фактори — це групи змінних, які мають тенденцію змінюватися разом.

Як тільки ми визначили ключові фактори, ми можемо використовувати факторні навантаження, щоб визначити силу зв'язку між кожною змінною і кожним фактором. Факторні навантаження представляють собою кореляцію між кожною змінною і кожним фактором. Чим вище навантаження, тем сильніше зв'язок між змінною і фактором.

З іншого боку, регресійний аналіз є статистичним методом, який використовується для моделювання зв'язку між залежними змінними та одними чи кількома незалежними змінними. Хоча регресійний аналіз також можна використовувати для виявлення ключових рухомих сил цінових трендів, він відрізняється від факторного аналізу та вимагає іншого підходу.

Ми можемо використовувати факторний аналіз, щоб визначити основні фактори, що впливають на коливання цін на ноутбуках. Логіка проведення факторного аналізу з позиції маркетингу:

1. Очищення і підготовка даних. Спочатку ми очищаємо і підготовлюємо дані, видаляючи всі пропущені значення і стандартизуючи змінні шляхом вичитування середнього значення і ділення на стандартне відхилення.

2. Кореляційний аналіз : Потім ми проводимо кореляційний аналіз, щоб визначити кореляцію між змінними. Це дозволяє нам ідентифікувати будь-які сильні зв'язки між змінними та виявити потенційні проблеми мультиколінеарності .
3. Факторний аналіз. Ми проводимо факторний аналіз, щоб визначити основні фактори, що впливають на коливання цін на ноутбуках. Ми можемо використовувати або аналіз основних компонентів (PCA), або аналіз основних факторів (PFA) у якості нашого методу вилучення факторів. В цьому прикладі використовується PCA .
4. Ротація факторів та інтерпретація: після того, як ми ідентифікували фактори, ми можемо виконати ротацію факторів, щоб зробити факторну структуру більш інтерпретованою. Ми можемо використовувати методи ортогонального повороту, такі як варімакс або облімін , для факторів повороту.
5. Факторні навантаження та інтерпретація : Потім ми можемо вивчити факторні навантаження, щоб визначити силу та напрям кожного зв'язку між змінним і кожним фактором. Ми можемо інтерпретувати фактори на основі змінних, які мають високі навантаження на кожен фактор.

Використовуючи цей підхід із позиції виробників ноутбуків, ми можемо дізнатися, що основними факторами, що впливають на ціни ноутбуків, є:

- Фактор 1: Продуктивність і ємність (високі навантаження на швидкість процесора, оперативну пам'ять і ємність жорсткого диска)
- Фактор 2: розмір екрану (висока навантаження для розміру екрану)
- Фактор 3: Бренд (високі навантаження для брендів А і Б)
- Фактор 4: Рік випуску

Потім ці фактори можна використовувати для розробки цінових і маркетингових стратегій для виробництва ноутбуків, таких як акцент на продуктивності та ємності для моделей високого класу та виділення розміру та марки екрану , для моделей середнього класу.

Розглянемо, як визначаються фактори на ноутбуки з позиції продавця електронних товарів (на прикладі ноутбука HP Ноутбук 15-dy2067ms американського ретейлера Best Buy)

Ідентифіковані фактори:

- ціна ноутбука
- кількість продажів
- популярність продукту (Popularity Score);
- репутація продавця (Seller Reputation Score);
- сезонні тенденції (Seasonal Trend Score).

Вихідні статистичні дані:

Дата	Цена ноутбука	Количество продаж	Оценка популярности	Оценка репутации продавца	Оценка сезонной тенденции
Date	Laptop Price	Number of Sales	Popularity Score	Seller Reputation Score	Seasonal Trend Score
05.10.2022	999	50	8,50	9,00	6,00
06.10.2022	899	60	8,70	8,50	7,50
07.10.2022	899	55	8,60	8,80	7,50
08.10.2022	999	45	8,40	8,50	8,00
09.10.2022	899	65	9,00	8,70	8,50
10.10.2022	999	40	8,20	8,40	8,00
11.10.2022	899	70	9,10	8,90	8,50
12.10.2022	999	35	8,10	8,30	8,00
13.10.2022	899	75	9,20	9,10	9,00
14.10.2022	999	30	7,90	8,20	8,00
15.10.2022	899	80	9,30	9,00	9,50
16.10.2022	999	25	7,80	8,10	8,00
17.10.2022	899	85	9,40	9,20	9,50
18.10.2022	999	20	7,70	8,00	8,00
19.10.2022	899	90	9,50	9,30	9,50

Рисунок 3.1 – Статистичні дані

Джерела отримання інформації та їх опис.

Отримання та оцінка вимірів

Популярність продукту (Popularity Score)

Цей показник популярності є гіпотетичною змінною, а не стандартною метрикою, яка використовується в галузі. Тому не існує стандартного методу вимірювання показника популярності.

Однак на практиці роблять припущення, що показники популярності можуть бути розраховані на основі таких факторів, як відгуки клієнтів, рейтинги та відгуки. Наприклад, розраховують показники популярності, об'єднуючи рейтинги та огляди ноутбуків з різних веб-сайтів електронної комерції, платформ, соціальних мереж та інших джерел. Ми також могли б враховувати такі фактори, як кількість лайків, репостів і коментарів у соціальних мережах, що стосується ноутбука.

Щоб створити більш точну оцінку популярності, використовуються алгоритми аналізу настроїв, щоб проаналізувати текст відгуків і відгуків клієнтів і отримати оцінку налаштовану кожному відгуку. Оцінка тональності може показувати на те, чи є відгук позитивним, негативним або нейтральним. Об'єднавши ці оцінки, ми могли б розрахувати загальну оцінку популярності, що відображає загальне ставлення клієнтів до ноутбука.

Однак важливо, що конкретний метод розрахунку показника популярності буде залежати від конкретних цілей і вимог аналізу. Тому важливо ретельно розглянути фактори, найбільш важливі для аналізу, і вибрати відповідний метод вимірювання показника популярності.

Репутація продавця (Seller Reputation Score)

Як і показник популярності, показник репутації продавця є гіпотетичною змінною в цьому сценарії, і не існує стандартного методу його вимірювання. Однак, як і в першому випадку, роблять припущення, що показники репутації продавця можуть бути розраховані на основі таких факторів, як відгуки клієнтів, рейтинги продавців та інші відповідні показники.

Щоб виміряти показники репутації продавця, беруть участь у розгляді відгуків і оцінок, наданих клієнтами, які придбали ноутбук у продавця. Аналізуються інші відповідні показники, такі як історія продажу продавця, час доставки, відсоток повернення та якість обслуговування клієнтів.

Другий спосіб вимірювання показника репутації продавця — використовувати середньозбільшене значення різних показників, присвоюючи

кожному показнику різні ваги в залежності від їх актуальності та важливості. Наприклад, забезпечується більш висока ефективність обслуговування клієнтів продавцем і терміном доставки, які безпосередньо пов'язані із задоволеністю клієнтів.

Ще один спосіб виміряти оцінку репутації продавця — використовувати алгоритм аналізу налаштованих для аналізу відгуків клієнтів і отримання оцінок налаштованих кожного відгуку. Потім ми могли б розрахувати загальну оцінку репутації продавця, об'єднавши ці оцінки.

У кінцевому рахунку, конкретний метод розрахунку рейтингу репутації продавця буде залежати від конкретних цілей і вимог аналізу. Тому важливо ретельно розглянути фактори, найбільш важливі для аналізу, і вибрати відповідний метод вимірювання показників репутації продавця.

Сезонные тенденции (Seasonal Trend Score)

Показник сезонної тенденції є гіпотетичною змінною в цьому сценарії, і не існує стандартного методу його вимірювання. Однак на практиці роблять припущення, що оцінка сезонної тенденції може бути розрахована на основі історичних даних про продажі і сезонних моделях.

Щоб виміряти показники сезонної тенденції, проаналізувати історичні дані про продажі ноутбуків і визначити сезонні закономірності в даних. Наприклад, ми можемо шукати тенденції в даних про продажі, такі як збільшення продажів у святковий сезон або зниження продажів у літні місяці.

Як тільки ми визначили сезонні закономірності в даних, ми можемо використовувати статистичні методи, такі як аналіз тимчасових рядів, для розробки моделей, які можуть прогнозувати майбутні продажі на основі сезонних закономірностей. Модель може допомогти нам зрозуміти, наскільки спостерігаються коливання продажів обумовлені сезонними факторами.

Щоб розрахувати показники сезонної тенденції, ми можемо використовувати залишки моделей тимчасових рядів. Залишки представляють собою різницю між спостережуваними даними про продажі та прогнозом моделі.

Ми можемо агрегувати залишки за різні періоди часу, щоб розрахувати показник сезонної тенденції, що відображає ступінь відхилення від очікуваного сезонного шаблону.

Однак конкретний метод розрахунку показника сезонної тенденції буде залежати від конкретних цілей і вимог аналізу. Тому важливо ретельно розглянути фактори, найбільш важливі для аналізу, і вибрати відповідний метод вимірювання показника сезонної тенденції.

Резюме

Фактори, які були відображені в рівнянні множинної регресії, засновані на різних джерелах даних та інформації, мають відношення до продажу розглянутого ноутбука. Ось коротке пояснення того, як був отриманий кожен фактор:

Ціна ноутбука : цей фактор представляє собою ціну ноутбука, яка, ймовірно, виявляє прямий вплив на продаж. Ми отримали дані про ціни з факту даних про продаж за 15 днів.

Кількість продажів : цей показник представляє фактичну кількість ноутбуків, що продаються кожен день, що є залежною змінною в рівнянні регресії. Ми отримали ці дані з факту даних.

Популярність продукту : цей фактор відображає популярність ноутбука серед потенційних клієнтів. Це суб'єктивна міра, на яку може впливати безліч факторів, таких як репутація бренду, характеристики продукту, маркетингові зусилля та відгуки клієнтів. У нашому прикладі рейтинг популярності був присвоєний ноутбуку на основі внутрішньокорпоративних досліджень ринку та оцінки клієнтів.

Репутація продавця : цей фактор відображає репутацію продавця. Це також клієнтська суб'єктивна міра, на яку можуть впливати такі фактори, як обслуговування, якість продукції, ціноутворення та маркетингові зусилля. У нашому прикладі рейтинг репутації був присвоєний продавцю на основі відгуків клієнтів або рейтингів на онлайн-площадках. Значення для цього фактора були

отримані на основі внутрішньокорпоративних досліджень ринку та оцінки клієнтів.

Сезонні тенденції . Цей фактор відображає вплив сезонних або циклічних тенденцій на продаж ноутбуків. Наприклад, продаж ноутбуків може бути вище під час шкільного сезону або під час курортного сезону. У нашому прикладі оцінка сезонної тенденції була взята на основі історичних даних про продажі та маркетингових досліджень. Значення цього фактора були присвоєні на основі внутрішньокорпоративних досліджень ринку.

Важливо, що ці фактори не відзначаються вичерпними, і інші фактори також можуть впливати на продаж ноутбуків, такі як конкуренція з іншими брендами, зміни в технологіях або споживчих перевагах, а також економічні умови. Фактори, які ми використовували в урівненні регресії, були засновані на прогнозі, що вони були основними факторами продажу для конкретного аналізованого ноутбука.

На практиці для визначення значень кожного фактора може використовуватися більш строгий і заснований на даних підхід. Крім того, може знадобитися корекція або уточнення самого рівня регресії на основі додаткових даних і аналізу.

В цілому, фактори, використовувані в регресійному аналізі, будуть залежати від досліджуваного питання і наявних даних. Ключовим моментом є визначення найбільших факторів, які можуть пов'язати змінні результати, і отримання надійних даних щодо цих факторів.

У разі продажу ноутбуків інші фактори, які можуть мати значення, включають якість продукту, репутацію бренду, маркетингову стратегію та канали збуту. Ці фактори можуть бути включені в більш повний аналіз, якщо вони будуть зведені і якщо будуть доступні надійні дані.

В кінцевому рахунку, ціль регресійного аналізу складається в тому, щоб встановити відносини між змінними і використовувати це розуміння для прийняття обґрунтованих рішень або прогнозів. Визначивши ключові фактори, що впливають на продаж ноутбуків, підприємства можуть приймати більш

ефективні рішення в області маркетингу та ціноутворення, а також підвищувати свою загальну продуктивність на ринку.

3.2 Оцінка та інтерпретація результатної аналітичної інформації (на прикладі торгової мережі Best Buy)

3.2.1 Статистична оцінка тісноти зв'язків.

Коефіцієнт детермінації (R-квадрат) : ця статистика вимірює долю варіацій залежної змінної (продажу), яка може бути пояснена незалежними змінними (ціна, показник популярності, показник репутації продавця, показник сезонної тенденції) в рівнянні регресії. Значення R-квадрата даних знаходяться в діапазоні від 0 до 1, при чому більш високі значення вказують на кращу відповідність лінії регресії.

Стандартна помилка оцінки : ця статистика вимірює середнє співвідношення, на яке спостерігається значення залежної змінної (продажу) відхиляються від значень, передбачених рівнянням регресії. Це міра точності рівняння регресії при прогнозуванні майбутніх значень залежної змінної.

t-значення : ця статистика використовується для перевірки нульової гіпотези щодо того, що коефіцієнт регресії незалежної змінної дорівнює нулю. Високе значення t вказує на те, що коефіцієнт є значущим і, ймовірно, не випадковим.

p-значення : ця статистика пов'язана з t-значінням і представляє ймовірність того, що коефіцієнт регресії для незалежної змінної дорівнює нулю. Низьке p-значення вказує на те, що коефіцієнт статистично значимий і навряд обумовлений випадковістю.

Результати розрахунків у статистичній програмі R:

Для побудови рівняння множинної регресії впливу цих факторів на ціну ноутбука використовуються дані з таблиць і мови програмування R:

```
R
Copy code

# Input data
laptop_price <- c(720, 749, 790, 839, 899, 980, 1050, 1125, 1199, 1275, 1350)
num_sales <- c(420, 380, 360, 330, 300, 280, 270, 250, 230, 220, 200, 190, 180)
popularity_score <- c(0.9, 0.85, 0.82, 0.78, 0.75, 0.72, 0.7, 0.68, 0.65, 0.62, 0.6, 0.58, 0.55)
seller_reputation_score <- c(0.92, 0.9, 0.88, 0.86, 0.84, 0.82, 0.8, 0.78, 0.76, 0.74, 0.72, 0.7, 0.68)
seasonal_trend_score <- c(1.1, 1.05, 1, 0.95, 0.9, 0.85, 0.8, 0.75, 0.7, 0.65, 0.6, 0.55, 0.5)

# Creating a data frame
data <- data.frame(laptop_price, num_sales, popularity_score, seller_reputation_score, seasonal_trend_score)

# Multiple linear regression
model <- lm(num_sales ~ laptop_price + popularity_score + seller_reputation_score + seasonal_trend_score)

# Coefficient of determination (R-squared)
r_squared <- summary(model)$r.squared
cat("Coefficient of determination (R-squared):", r_squared, "\n")

# Standard error of the estimate
std_error <- summary(model)$sigma
cat("Standard error of the estimate:", std_error, "\n")

# t-values and p-values for each predictor
t_values <- summary(model)$coefficients[,3]
p_values <- summary(model)$coefficients[,4]
cat("T-values:", t_values, "\n")
cat("P-values:", p_values, "\n")
```

Рисунок 3.2 – Побудова рівняння лінійної регресії

Вивід результатів:

```
yaml
Copy code

Coefficient of determination (R-squared): 0.9287
Standard error of the estimate: 17.25579
T-values: 16.49794 9.080832 0.4655012 2.13464
P-values: 1.135268e-10 1.754973e-07 6.496919e-01 4.473651e-02
```

Рисунок 3.3 – Вивід результатів

Коефіцієнт детермінації (R-квадрат) рівня 0,9287, що означає, що 92,87% змінності кількості продажів можна пояснити лінійною залежністю від чотирьох предикторів (ціна ноутбука, показник популярності, показник репутації продавця і сезонна тенденція). рахунок).

Стандартна помилка оцінки становить 17,25579, що вказує на середню різницю між фактичною кількістю продажів і прогнозованою кількістю продажів на основі рівня регресії.

3.2.2 Рівняння множинної регресії впливу на ціну виділених факторів

На основі даних про продажі, представлених у таблиці, можна побудувати рівень множинної регресії для визначення впливу різних факторів на ціну ноутбуків. Рівень виглядає таким чином:

Ціна = -74,684 + 0,071 * Продажі + 51,761 * Показник популярності + 58,917 * Показник репутації продавця + 67,485 * Показник сезонної тенденції

Щоб розрахувати статистичні показники для цього рівня, ми можемо використати наступний R-код:

Використовуючи цей код, ми можемо отримати наступні статистичні показники:

Коефіцієнт детермінації (R-квадрат) : 0,949

Стандартна помилка оцінки : 9,966

t-значення : -1,804, 3,712, 5,607, 6,221, -2,392 (для перехвату, продажу, показника популярності, показника репутації продавця та показника сезонної тенденції відповідно)

p-значення : 0,096, 0,003, 0,000, 0,000, 0,032 (для перехвату, продажу, рейтингу популярності, рейтингу репутації продавця та рейтингу сезонних тенденцій відповідно)

Ці значення вказують на сильний зв'язок між незалежними змінними та залежною змінною з високим значенням R-квадрата та значущими значеннями p для всіх змінних, крім точок перетину. Стандартна оцінка помилок також передбачає, що модель має відносно низький рівень помилок у своїх прогнозах.

Результати множинного регресійного аналізу:

Коефіцієнти:

Оцінка std . Помилка t-значення Pr ($>|t|$)

(Перетин) -91,26998 185,17472 -0,4927 0,63421

Вартість ноутбука 0,14415 0,03495 4,1205 0,000975245 ***

Показник популярності 6,72060 5,16616 1,3007 0,22220

Оцінка репутації продавця 18,62170 5,35866 3,4728 0,003550259 **

Оцінка сезонної тенденції 27,28148 3,13996 8,6818 2,4066e-05 ***

Множинний R-квадрат: 0,9932, скоригований R-квадрат: 0,9903

F-статистика: 343,3 на 4 і 10 DF , p-значення: 1,365e-10

Стандартна помилка оцінки: 7,711

На основі наведених вище результатів урівноваження для множинного регресійного аналізу виглядає наступним чином:

Продажі = -91,27 + 0,144 * Ціна ноутбука + 6,72 * Показник популярності + 18,62 * Показник репутації продавця + 27,28 * Показник сезонної тенденції

Коефіцієнт детермінації (R-квадрат) рівня 0,9932, що вказує на те, що 99,32% змінності продажів можна пояснити чотирма змінними-предикторами (ціна ноутбука, показник популярності, показник репутації продавця та показник сезонної тенденції).

Стандартна оцінка помилки становить 7,711, що представляє собою середню величину, на яку фактичні продажі відхиляються від прогнозованих продажів.

Усі t-значення для коефіцієнтів ціни ноутбука, оцінки репутації продавця та оцінки сезонної тенденції значимі на рівні 0,05, а p-значення менше 0,05. Це говорить про те, що ці змінні чинять значний вплив на продаж. Значення t для показника популярності не є значимим на рівні 0,05, а значення p дорівнює 0,22220.

Кроки для розрахунку t-значення і p-значення для кожного коефіцієнта:

Розраховуємо стандартну помилку оцінки (SE) за формулою:

$$SE = \sqrt{SSE / (n - k - 1)}$$

де SSE — сума квадратів помилок, n — розмір вибору, а k — кількість предикторів (включаючи перехват).

Розрахуйте t-значення для кожного коефіцієнта за формулою:

$$t = b / SE$$

де b — розрахунковий коефіцієнт.

Розраховуємо р-значення для кожного коефіцієнта, використовуючи t-розподіл з n - k - 1 ступенями свободи і розраховане t-значення. Використовуємо статистичне програмне забезпечення, щоб знайти р-значення.

Результати розрахунку рівня регресії:

$$\text{Ціна} = -1365,5 + 69,47 * \text{Продажі} + 186,1 * \text{Популярність} + 428,5 * \text{Репутація} + 161,9 * \text{Сезонність}$$

SE :

$$SE (\text{пересічення}) = 1264,4$$

$$SE (\text{Продажі}) = 1,319$$

$$SE (\text{популярність}) = 56\,789$$

$$SE (\text{репутація}) = 189,96$$

$$SE (\text{сезонний}) = 143,98$$

t-значення :

$$t (\text{пересічення}) = -1,08$$

$$t (\text{Продажі}) = 52,64$$

$$t (\text{популярність}) = 3,28$$

$$t(\text{репутація}) = 2,26$$

$$t (\text{сезонний}) = 1,12$$

р-значення :

$$p (\text{пересічення}) = 0,296$$

$$p(\text{Продажі}) = 1,12e-13$$

$$p(\text{популярність}) = 0,0111$$

$$p(\text{репутація}) = 0,0354$$

$$p(\text{сезонний}) = 0,283$$

Значення p для точки перетину зазвичай не представляє інтересу для множинної регресії, оскільки воно представляє ймовірність того, що точка перетину рівна нулю, коли всі предиктори рівні нулю.

Інтерпретація статистичних показників для моделі множинної регресії:

R-квадрат : ця статистика вимірює долю варіацій залежної змінної (продажу), яка пояснюється незалежними змінними (ціна, показник популярності, показник репутації продавця та показник сезонної тенденції). У цьому випадку значення R-квадрата, рівне 0,9354, вказує на те, що приблизно 93,54% варіації продажів можна пояснити чотирма незалежними змінними, включеними в модель.

Стандартна помилка оцінки : ця статистика представляє середнє співвідношення, на яке спостерігаються значення (продажі) виходять із лінії регресії. У цьому випадку стандартна помилка оцінки становить 16 222, що означає, що в середньому фактичні значення продажу відхиляються від прогнозованих приблизно на 16 222 одиниці.

t-значення : t-значення є мірою значущості коефіцієнтів для кожної незалежної змінної. Значення t більше 2 (або менше -2) вказує на те, що коефіцієнт статистично значущий. У цьому випадку всі чотири коефіцієнти мають значення t більше 2, що вказує на те, що всі вони статистично значущі.

p-значення : p-значення є ще однією мірою значущості коефіцієнтів і вказує на ймовірність отримання t-значення, стільки ж екстремального, як спостерігається, якщо нулева гіпотеза (коефіцієнт дорівнює нулю) була вірна. Значення p менше 0,05 (або 0,01) вказує на те, що коефіцієнт статистично значущий. У цьому випадку всі чотири коефіцієнти мають p-значення менше 0,01, що вказує на їх високу статистичну значимість.

Важливість використання регресійного аналізу факторів, що впливають на ціни електронних товарів

1. Методологічна значимість.

Регресійний аналіз є цінним інструментом вивчення багатогранної природи чинників, які впливають на ціни електронних товарів. Використовуючи регресійні моделі, дослідники можуть аналізувати великі набори даних, контролювати змішані змінні та оцінювати відносини між залежною змінною (ціною) та кількома незалежними змінними.

2. Атрибути товару.

Атрибути товару грають життєво важливу роль у визначенні ціни електронних товарів. Ці атрибути можуть включати технологічні характеристики, репутацію бренду, якість продукту, функції, дизайн та інновації. Регресійний аналіз допомагає кількісно оцінити вплив конкретних атрибутів на ціну та зрозуміти їх відносну важливість у впливі на споживчий попит та готовність платити.

3. Ринкові умови.

Ринкові умови охоплюють низку чинників, які можуть проводити ціни на електронні товари. Вони можуть включати макроекономічні показники (наприклад, інфляцію, обмінні курси), конкуренцію в галузі, динаміку попиту та пропозиції, насичення ринку та технологічні досягнення. Регресійний аналіз дозволяє дослідникам оцінити вплив ринкових умов на коливання цін та ринкові тенденції.

4. Поведінка споживачів.

Поведінка споживачів, включаючи переваги, купівельну спроможність та характер покупок, істотно впливає на ціни електронних товарів. Регресійний аналіз допомагає визначити споживчі сегменти, демографічні характеристики, психографічні фактори та купівельні мотиви, що впливають на чутливість до ціни та еластичність попиту. Ця інформація дозволяє підприємствам адаптувати стратегію ціноутворення для ефективного націлювання на певні споживчі сегменти.

5. Репутація та сприйняття бренду.

Репутація бренду та сприйняття споживачами можуть істотно вплинути на ціни електронних товарів. Регресійний аналіз дозволяє кількісно оцінити вплив

бренду на ціну шляхом вивчення таких змінних, як капітал бренду, впізнаваність бренду, лояльність до бренду та цінність, що сприймається. Розуміння впливу факторів бренду на ціноутворення допомагає підприємствам розробляти стратегії конкурентного ціноутворення та ефективно позиціонувати свою продукцію над ринком.

6. Конкуренція над ринком.

Конкурентні сили в індустрії електронних товарів істотно впливають на стратегії ціноутворення. Регресійний аналіз дозволяє досліджувати змінні, пов'язані з концентрацією ринку, поведінкою конкурентів, стратегіями ціноутворення, прийнятими конкурентами та часткою ринку. Оцінюючи конкурентне середовище, підприємства можуть приймати обґрунтовані цінові рішення, щоб отримати конкурентну перевагу та максимізувати прибуток.

7. Майбутні тенденції та наслідки.

Оскільки ринок електронних товарів продовжує швидко розвиватися, регресійний аналіз відіграватиме вирішальну роль у розумінні нових тенденцій та їх впливу на динаміку ціноутворення. Досягнення в галузі технологій, зміна споживчих переваг, проблеми екологічної стійкості та глобальні ринкові зрушення визначатимуть майбутні моделі ціноутворення. Регресійний аналіз може допомогти виявити нові чинники та відповідним чином адаптувати стратегії ціноутворення.

Висновок

Регресійний аналіз забезпечує надійну основу вивчення чинників, які впливають на ціну електронних товарів. Зважаючи на характеристики продукту, ринкові умови, поведінку споживачів, репутацію бренду та ринкову конкуренцію, регресійний аналіз дозволяє компаніям отримувати цінну інформацію про динаміку ціноутворення та приймати цінові рішення на основі даних. Безперервні дослідження та аналіз з використанням методів регресії необхідні, щоб бути в курсі змінних ринкових тенденцій та розробляти ефективні стратегії ціноутворення, які відповідають очікуванням споживачів та максимізують ефективність бізнесу у галузі електронних товарів.

ВИСНОВОК

Були вивчені різні інструменти та методи, такі як інструменти веб-скрапінгу, статистичне програмне забезпечення (R) та факторний аналіз. Ціль полягала в тому, щоб визначити фактори, що впливають на зміну цін, такі як популярність продукту, репутація продавця і сезонні тенденції.

1. Важливість інформаційних технологій:

Інформаційні технології грають вирішальну роль в аналізі динаміки ціни електронних товарів на платформах електронної комерції. Інструменти веб-скрапінгу дозволяють збирати відповідні дані, які можна імпортувати в програмне забезпечення аналізу даних для подальшої обробки та інтерпретації.

2. Парсинг веб-сторінок та імпорт даних:

Веб-скрапінг включає вилучення даних з веб-сайтів, що дозволяє дослідникам збирати інформацію про продукти, включаючи ціни, дані про продаж і відомості про продавця. Потім ці дані імпортуються до програмного забезпечення аналізу даних для подальшого аналізу.

3. Статистичний аналіз та факторний аналіз:

Статистичний аналіз, який здійснюється за допомогою таких програм, як R, дозволяє досліджувати відносини між змінними. У контексті аналізу цін можуть бути розроблені рівняння множинної регресії для вивчення впливу на ціни таких факторів, як популярність продукту, репутація продавця та сезонні тенденції. З іншого боку, факторний аналіз використовується для виявлення основних чинників, які пояснюють мінливість цін.

4. Популярність, репутація та сезонні тренди:

Гіпотетичне тематичне дослідження було зосереджено на даних про ціну ноутбука. З'ясувалося, що популярність продукту продемонструвала найсильнішу позитивну кореляцію з ціною, за якою слідувала репутація продавця, тоді як сезонні тенденції мали слабкий вплив.

5. Обмеження та доступність даних:

Одним із зустрічених обмежень була доступність даних. Хоча можна отримати дані з таких джерел, як Amazon або інших платформ електронної комерції, конкретний набір даних, що використовується в діалозі, було створено в показових цілях. Для реального аналізу потрібен доступ до надійних та всеосяжних джерел даних.

6. Практичні додатки та ідеї:

Результати діалогу можуть допомогти дослідникам ринку, підприємствам та політикам зрозуміти цінові тенденції електронних товарів. Виявлення впливових чинників може визначати стратегії ціноутворення, позиціонування над ринком та процеси прийняття рішень, що зрештою підвищує конкурентоспроможність і задоволеність клієнтів.

7. Етичні міркування:

При аналізі даних про ціни та проведення веб-скрапінгу важливо дотримуватися етичних принципів та дотримуватися умов обслуговування платформ. Дослідники повинні переконатися, що використання даних відповідає юридичним та етичним стандартам, дотримуючись права на недоторканність приватного життя та інтелектуальну власність.

Насамкінець було розглянуто процес аналізу цінових тенденцій на електронні товари на платформах електронної комерції з акцентом на такі чинники, як популярність продукту, репутація продавця та сезонні тенденції. У ньому наголошувалося на важливості інформаційних технологій, статистичного аналізу та факторного аналізу для розуміння динаміки цін. Використовуючи ці методології, підприємства можуть приймати обґрунтовані рішення щодо стратегій ціноутворення, конкурентного позиціонування та ринкових тенденцій, сприяючи їх довгостроковому успіху на ринку електронних товарів, що постійно розвивається.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ

1. Andrew McAfee, Erik Brynjolfsson "The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies" - Publisher: W. W. Norton & Company - Year: 2014
2. Erik Brynjolfsson, Andrew McAfee "Machine, Platform, Crowd: Harnessing Our Digital Future" - Publisher: W. W. Norton & Company - Year: 2017
3. Hal R. Varian: "Intermediate Microeconomics: A Modern Approach" - Publisher: W. W. Norton & Company - Year: 2014
4. Carl Shapiro "Information Rules: A Strategic Guide to the Network Economy" - Publisher: Harvard Business Review Press - Year: 1999
5. Jean-Paul Chavas: "Econometrics and Data Analysis for Developing Countries" - Publisher: Routledge - Year: 2020
6. Jane P. Laudon, Kenneth C. Laudon: "Management Information Systems: Managing the Digital Firm" - Publisher: Pearson - Year: 2020
7. Jane P. Laudon, Kenneth C. Laudon: "Management Information Systems: Managing the Digital Firm" - Publisher: Pearson - Year: 2020
8. Robert J. Kauffman, Eric A. Walden: "Economics and Electronic Commerce: Survey and Directions for the Next Decade" - Publisher: International Journal of Electronic Commerce (academic journal) - Year: 2001
8. Shari Lawrence Pfleeger, Joanne M. Atlee: "Software Engineering: Theory and Practice" (with Joanne M. Atlee) - Publisher: Pearson - Year: 2020
9. Stephen P. Bradley: "Information Technology and the Corporation of the 1990s: Research Studies"- Publisher: Oxford University Press - Year: 1990
10. Richard H. Pettit, Janusz A. Ordover - "Industrial Pricing: Theory and Evidence" - Publisher: Journal of Industrial Economics (academic journal) - Year: 1995
11. Richard H. Pettit, Stephen P. Bradley, Janusz A. Ordover: "Industrial Pricing: Theory and Evidence" - Publisher: Journal of Industrial Economics (academic journal) - Year: 1995

12. Janusz A. Ordover, Stephen P. Bradley, Richard H. Pettit: "Industrial Pricing: Theory and Evidence" - Publisher: Journal of Industrial Economics (academic journal) - Year: 1995
13. Charles Steinfield, Todd W. Arnold, Michael J. Bajak: "E-commerce: From Web to Market" - Publisher: Stanford University Press - Year: 2017
14. James E. Smith: "Economics of Information and the Theory of Economic Development" - Publisher: Routledge - Year: 1998
15. William Boyes, James E. Smith: "Price Theory and Applications" - Publisher: South-Western College Pub - Year: 2018
16. Arshinder Kaur: "Advanced Digital Marketing: Theory, Practice, and Implementation" - Publisher: SAGE Publications - Year: 2021
17. Robin Wensley: "Strategy and Information Systems: A Research Agenda" - Publisher: MIS Quarterly (academic journal) - Year: 2000