

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
ІМЕНІ ТАРАСАШЕВЧЕНКА  
ЕКОНОМІЧНИЙ ФАКУЛЬТЕТ  
КАФЕДРА ЕКОНОМІЧНОЇ КІБЕРНЕТИКИ

КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА

**МОДЕЛЮВАННЯ ПЕРСОНАЛІЗОВАНОГО ЦІНОУТВОРЕННЯ НА  
ОСНОВІ ПІДХОДІВ DATA SCIENCE**

Студентки 2 курсу магістратури  
спеціальності 051 «Економіка»  
ОНП «Економічна кібернетика»  
денної форми навчання  
Горбунова Ярослава Миколаївна

**Науковий керівник:**  
д.е.н., професор  
Чорноус Галина Олександрівна

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

Студент \_\_\_\_\_  
(підпис)

Робота допущена до захисту в ЕК  
рішенням кафедри економічної кібернетики  
від 4 травня 2022 р., протокол № 13

Завідувач кафедри економічної кібернетики,  
доктор економічних наук, професор  
Ляшенко Олена Ігорівна

\_\_\_\_\_  
(підпис)

## **. РЕФЕРАТ**

Кваліфікаційна робота магістра містить: 53 ст., 15 рис., 4 табл., 80 джерел, 2 додатки

Ключові слова: персоналізоване ціноутворення, інформаційна економіка, ціноутворення, маркетплейс, Data Science, машинне навчання, CatBoost регресія

Об'єкт дослідження: персоналізоване ціноутворення в інформаційній економіці.

Мета дослідження: систематизація теоретично-методологічні напрацювання щодо ціноутворення, обґрунтування економіко-математичного інструментарію на основі Data Science для визначення факторів, що мають найбільше значення при формуванні ціни при персоналізованому ціноутворенні.

Методи дослідження: методи наукової абстракції, порівнянь, індукції та дедукції, аналізу та синтезу, економіко-математичні методи та інформаційні технології.

Наукова новизна, теоретична значимість дослідження: виявлення специфічних особливостей персоналізованого ціноутворення. Представлено рейтинг релевантних методів машинного навчання для підтримки формування раціональної персональної ціни. Запропоновано використовувати CatBoost регресію – метод, який раніше не використовувався в дослідженні і підтримці персоналізованого ціноутворення.

## **RESUME**

Taras Shevchenko National University of Kyiv,

Faculty of Economics, Department of Economic Cybernetics

Key words: personalized pricing, information economy, pricing, marketplace, Data Science, Machine Learning, CatBoost regression

The graduation research of student Modeling personalized pricing based on Data Science approaches

deals with personalized pricing

The work is interesting for pricing specialists, scientists.

Pages 53, table 4, bibliog. 80, append. 2

## Зміст

ВСТУП .....	4
РОЗДІЛ 1. ОСНОВИ ЦІНОУТВОРЕННЯ В ІНФОРМАЦІЙНІЙ ЕКОНОМІЦІ ..	8
1.1. Види та особливості ціноутворення в інформаційній економіці.....	8
1.2. Персоналізоване ціноутворення та практика його застосування .....	14
РОЗДІЛ 2. МЕТОДИ МОДЕЛЮВАННЯ ПЕРСОНАЛІЗОВАНОГО ЦІНОУТВОРЕННЯ.....	24
2.1 Підходи, моделі та інструменти Data Science для підтримки персоналізованого ціноутворення.....	24
2.2 Використання методів машинного навчання до персоналізованого ціноутворення.....	31
РОЗДІЛ 3. МОДЕЛЮВАННЯ ЦІНОУТВОРЕННЯ НА ПРИКЛАДІ МАРКЕТПЛЕЙСУ «AIRBNB» .....	38
3.1 Постановка задачі й методологія .....	38
3.1 Моделювання та інтерпретування результатів.....	45
ВИСНОВКИ.....	51
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ .....	54
ДОДАТКИ.....	62

## ВСТУП

**Актуальність теми дослідження.** Швидкий розвиток технологій, а також ріст інформаційного сектору економіки сприяє пошуку різних методів для збільшення прибутків фірми або зменшення її витрат. Поява аналітики великих даних сприяла створенню форми цінової дискримінації на основі профілів споживачів та їх поведінки в Інтернеті - персоналізованого ціноутворення. Персоналізоване ціноутворення набуває популярності, особливо в електронній комерції, адже про поведінку споживачів в Інтернеті можна зібрати дані, за допомогою яких онлайн-магазини можуть запропонувати найкращу ціну для покупця і продавця.

Покупці стають усе більш вимогливими на сучасному ринку. Тому продавцям важливо пропонувати персоналізовані ціни, щоб завоювати серця покупців і залишатися конкурентоспроможними. Оскільки обчислити всі дані, необхідні для визначення оптимальних цін, неможливо, тому роздрібні компанії використовують підходи Data Science, зокрема алгоритми машинного навчання для обчислення та пропозиції оптимальних цін для якомога більшої кількості сегментів клієнтів. Проте окрім роздрібних продавців існують такі специфічні суб'єкти електронної комерції як маркетплейси. Персоналізоване ціноутворення при цьому має бути вигідним як для покупців, так і для продавців, і для посередника – маркетплейсу. Персональні ціни розширюють доступ до ринку як продавців, так і для споживачів із низькою готовністю платити й це приносить додаткові доходи для компанії, які раніше вона не отримувала через фіксовані ціни. Це підтверджує актуальність теми дослідження, представленої у цій кваліфікаційній роботі.

Питання ціноутворення в інформаційній економіці опрацьовані у наукових дослідженнях М. Біхлер, Х. Брукс, В. Елемагребі, Е. Загайнова, С. Кузнецов, М. Кумар, Я. Кхан, І. Нікітина, К. Равікумар, В. Рейнард.

Моделювання персоналізованого ціноутворення розглядали М. Дж. Аван, А. Аккісті, В. Дж. Аллендер, А. Ахуджа, Й. Бауер, Д. Берцимас, Х. Р. Варіан, Ф. Ву, А. Ельмахтуб, М. Д. Іслам, Н. Каллус, Й. Ляуконите, В. Місик, Ш. Насер,

Ч. Нью, Т. Дж. Річардс, Я. Сатсангі, Ш. Тан, Р. Філіпс, Г. Чан, Ч. Чжен, А. Чжоу, З. Чжоу, Н. Шукла, Д. Яннак. Наукові надбання цих вчених стали джерелами для проведених в роботі досліджень.

**Метою** випускної кваліфікаційної роботи є систематизація теоретично-методологічні напрацювання щодо ціноутворення в інформаційній економіці, визначення можливостей й перспективи персоналізованого ціноутворення, обґрунтування економіко-математичного інструментарію на основі Data Science для визначення факторів, що мають найбільше значення при формуванні ціни при персоналізованому ціноутворенні.

Досягнення поставленої мети забезпечено вирішенням таких завдань:

- узагальнити теоретичні підходи щодо ціноутворення загалом та в інформаційній економіці зокрема;
- визначити сутність та особливості персоналізованого ціноутворення;
- дослідити сучасну практику впровадження персоналізованого ціноутворення в компаніях;
- вивчити економіко-математичні підходи, моделі та інструменти Data Science, які застосовуються при персоналізованому ціноутворенні;
- проаналізувати переваги і недоліки алгоритмів машинного навчання, що можуть застосовуватися при персоналізованому ціноутворенні;
- визначити ключові фактори, що впливають на ціну, при бронюванні житла на платформі «Airbnb»;
- проаналізувати вплив персональної інформації про господаря на ціноутворення при бронюванні житла на платформі «Airbnb»;
- показати значимість динамічних та персональних характеристик при формуванні ціни;
- запропонувати набір релевантних алгоритмів машинного навчання для дослідження факторів впливу на ціну, що підтримують формування раціональної персональної ціни.

**Предметом** кваліфікаційної роботи є економіко-математичні методи та моделі для підтримки персоналізованого ціноутворення.

**Об'єктом** кваліфікаційної роботи є персоналізоване ціноутворення в інформаційній економіці.

Теоретичною та методологічною основою кваліфікаційної роботи стали фундаментальні положення сучасної економічної теорії, а також новітні наукові досягнення вітчизняних та зарубіжних науковців щодо ціноутворення.

Інформаційною базою при написанні роботи виступали офіційні законодавчі й нормативно-правові документи України, монографії, наукові статті вітчизняних та зарубіжних вчених, фахові періодичні видання. Інформаційною базою для реалізації моделей стали дані з онлайн-платформи для перегляду та бронювання житла Airbnb [1].

**Методи дослідження.** Вирішення поставлених у роботі завдань здійснювались із залученням низки загальнонаукових та спеціальних методів наукових досліджень. Методи наукової абстракції, порівнянь, індукції та дедукції було використано при вивченні підходів до визначення категорій ціноутворення, визначенні сутності та особливостей персоналізованого ціноутворення, класифікації моделей ціноутворення. Методи аналізу та синтезу дали змогу визначити особливості моделювання персоналізованого ціноутворення; економіко-математичні методи та інформаційні технології, зокрема методи машинного навчання, дозволили проаналізувати фактори впливу на ціноутворення в Airbnb. Програмну підтримку реалізації математичного моделювання персоналізованого ціноутворення здійснено у середовищі Python.

**Наукова новизна** полягає в систематизації теоретичних напрацювань та практики персоналізованого ціноутворення, виявленні його специфічних особливостей, загроз і перспектив при використанні суб'єктами електронної комерції. Представлено рейтинг релевантних методів машинного навчання для підтримки формування раціональної персональної ціни. Запропоновано використовувати CatBoost регресію – метод, який раніше не використовувався в дослідженні і підтримці персоналізованого ціноутворення.

**Практичне значення** одержаних результатів полягає в обґрунтуванні науково-практичних рекомендацій щодо доцільності застосування та реалізації

персоналізованого ціноутворення на засадах сучасних алгоритмів машинного навчання на підприємствах електронної комерції, а саме, на маркетплейсах.

Основні положення кваліфікаційної роботи апробовані на двох міжнародних науково-практичних конференціях студентів, аспірантів та молодих вчених «Шевченківська весна: Економіка» (Київ, 2020, 2022). За результатами доповідей опубліковано тези за проблематикою дослідження: «Персоналізоване ціноутворення: економіко-математичний аналіз практики використання» (2022 р.) та «Динамічне ціноутворення: аналіз підходів і моделей» (2020 р.). Також результати роботи апробовано на VII Міжнародній науково-практичній конференції «Інформаційні технології та взаємодії» (IT&I'2020). За результатами доповіді опубліковано наукову статтю на тему «Modeling and forecasting dynamic factors of pricing in e-commerce», проіндексовану у МНМБ Scopus.

Структура дослідження. Робота структурно складається зі вступу, трьох розділів, висновків і списку використаних джерел зі 80 найменувань. Повний обсяг роботи становить 54 сторінки, містить 4 таблиці, 15 рисунків, усього 2 додатки.

## РОЗДІЛ 1. ОСНОВИ ЦІНОУТВОРЕННЯ В ІНФОРМАЦІЙНІЙ ЕКОНОМІЦІ

### 1.1. Види та особливості ціноутворення в інформаційній економіці

Стрімкий розвиток цифрових технологій та поява платформних рішень призводять до формування концепції інформаційної економіки.

Інформаційна економіка - це електронна економічна діяльність, де більша частина зайнятого населення бере участь у сфері інформаційних послуг, де основними ресурсами є знання та інформація [2].

Рівень розвитку інформаційної економіки характеризується ступенем її інформатизації. До інформатизації відносять використання автоматизованих систем, трансформація та реінжиніринг бізнес-процесів, збільшення об'ємів накопичення інформації.

Основою для інформаційної економіки є цифрові технології (комп'ютери, телекомунікаційні пристрої) та інфраструктура (Інтернет та телекомунікаційні мережі). На основі цифрових технологій виробляються ключові продукти та послуги, зокрема цифрові платформи, мобільні додатки та платіжні послуги. А набір цифрових продуктів та послуг використовується в е-комерції, е-бізнесі, спільній економіці, гіг-економіці [3].

Електронна комерція – сфера інформаційної економіки, що охоплює бізнес-процеси пов'язані з проведенням торгових (замовлення та доставка товарів та послуг) і фінансових операцій (проведення платежів) з використанням інформаційно-комунікаційних технологій [2].

Інтерес до ринку електронної комерції збільшується з кожним роком. Для цього існують об'єктивні причини: зростання чисельності інтернет-аудиторії, висока частка онлайн-продажів у загальному обсязі продажів, розвиток соціальних і мобільних мереж. Підтвердженням цього є дані, подані на рис. 1.1: обсяги продажів у світі в електронній комерції щороку зростають. Як показує графік, значний стрибок був між 2019-2020 роком, у період початку пандемії Covid-19. Через карантинні обмеження більшість магазинів почала працювати онлайн.

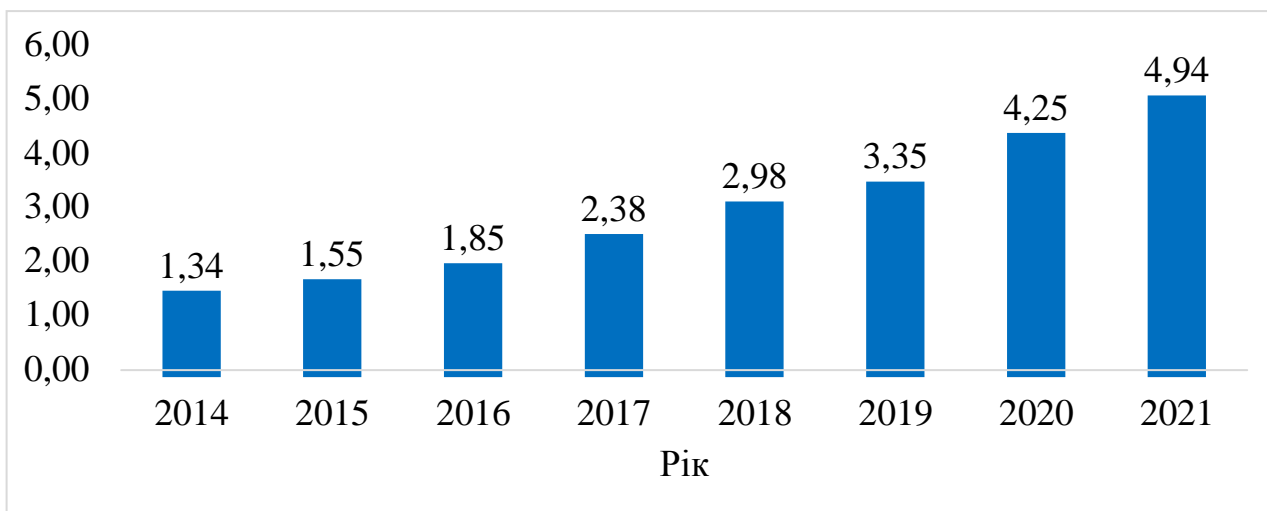


Рис. 1.1. Роздрібні продажі електронної комерції по всьому світу (млрд. дол. США)

Джерело: створено автором на основі [4]

Основними факторами впливу на інформаційну економіку є дослідження та розробки в галузі ІКТ; співробітники; розвиток сектору ІКТ; цифрові технології в бізнесі; інформаційна безпека [5].

Проникнення Інтернету є важливим фактором розвитку електронної комерції. На рис. 1.2 представлено динаміку долі населення, які користуються Інтернетом у світі та Україні для порівняння.

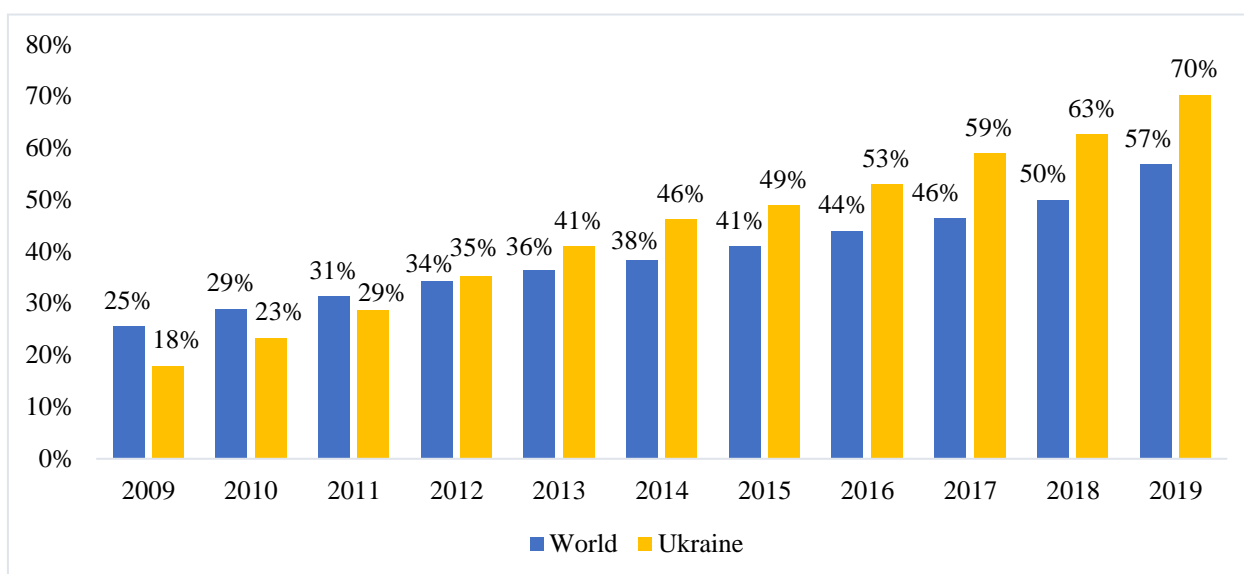


Рис. 1.2. Особи, які користуються Інтернетом (% населення)

Джерело: створено автором на основі [6]

У 2019 році відсоток населення, який користується інтернетом, у світі становив 57%, що є не достатньо високим, але з урахуванням країн, що розвивається, цей показник є достатньо непоганим. В Україні ж 70% населення використовує інтернет, що свідчить про високий рівень покриття країни.

Ціноутворення, як економічна категорія, пройшло довгий шлях у своєму історичному розвитку та набуло нових рис в інформаційній економіці.

Розглянемо основні етапи становлення ціноутворення та що було здобуто до сьогодні. Під ціноутворенням розуміються процес встановлення цін на товари та послуги, що пропонується на ринку. Загалом виділяють три основні етапи становлення теорії ціни[7]:

1. Теорія вартості: важлива сфера утворення вартості - суспільне виробництво. Основні автори: Аристотель, У. Петті, А. Сміт, Д. Рікардо, Дж. С. Мілль, К. Маркс, Ф. Енгельс, Ж. Б. Сей.

2. Теорія цінності (маржиналістична теорія): в основу цінності покладена споживча концепція ціни; ціна визначається ступенем корисного ефекту від споживання товару чи блага. Основні автори: К. Менгер, Е. Бем-Баверк, Ф. Візер, Е. Енгель, Курно, Е. Я. Брегель, П. Самуельсон.

3. Теорія ринкової ціни: в результаті взаємодії попиту і пропозиції утворюється ринкова рівновага, де єдиним засобом за допомогою якого можна забезпечити умови економічної рівноваги є механізм ціноутворення. Основні автори: А. Маршал, Дж. Мілль, Дж. Робінсон, Є. Чемберлен, А. Пігу, А. Чеканський, В. Репнева, І. Ліпсіц, Т. Миколаєва, І. Самуельсон, Ф. Котлер та інші.

Серед сучасних українських науковців внесок у розвиток теоретичних та практичних досліджень питань ціноутворення зробили В. Базилевич, Н. Меджибовська, М. Чайковська, Г. Чорноус та інші [8-11].

Згідно зі звітом Shopify «Future of Commerce Report 2022», ціна є ключовим фактором, який впливає на 74% споживачів [12].

Дослідження Namogo [13] підтверджують важливість ціни при прийнятті рішення про покупки в інтернет-магазинах (рис. 1.3). 50% опитуваних вважає найважливішим фактором ціну.

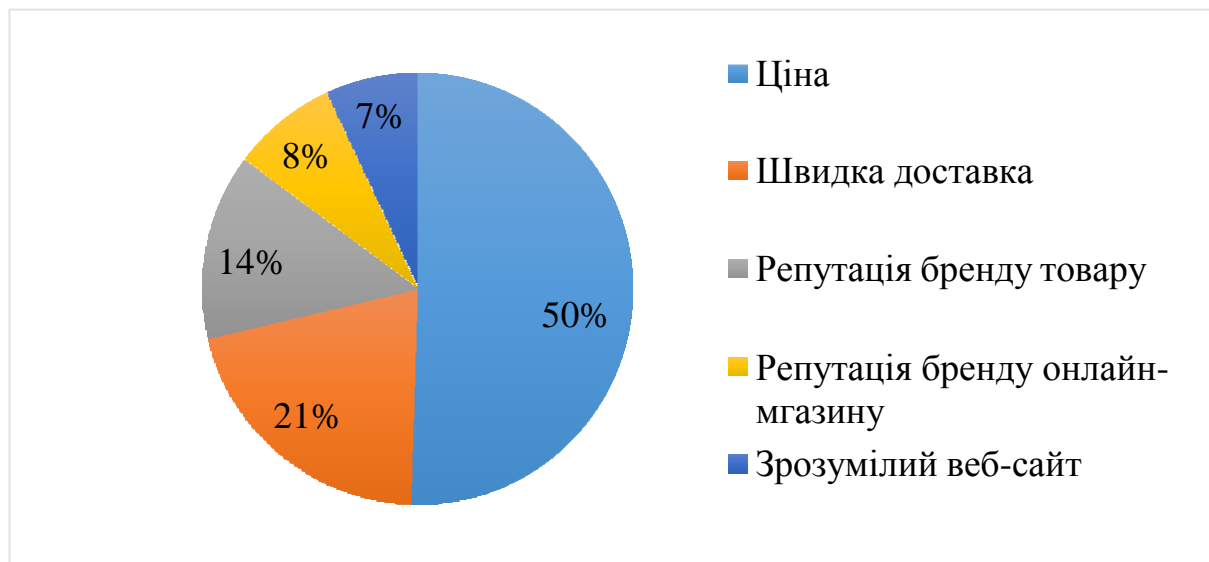


Рис. 1.3. Фактори впливу на онлайн покупку

Джерело: створено автором на основі [4,13]

Стратегія ціноутворення - це модель або метод, що використовується для встановлення найкращої ціни за товар чи послугу [14]. Стратегії ціноутворення допомагають вибрати ціни, які максимізують прибуток, урахувавши споживчий та ринковий попит.

Розглянемо основні стратегії для ціноутворення, які застосовуються сучасними бізнесами як в традиційній, так і в інформаційній економіці [15-18]:

1. Конкурентне ціноутворення – це стратегія, яка враховує поведінку споживача, устанавлюючи ціну на основі ціни за подібні товари чи послуги в конкурентів. Згідно з дослідженням, проведеним у 2016 році Forrester [19] 74% покупців використовують пошукові системи, щоб досліджувати товари та порівнювати їх від різних продавців і компаній.

2. Freemium - це стратегія, за якої компанії безкоштовно пропонують базову версію свого продукту, сподіваючись, що користувачі заплатять за оновлення або доступ до додаткових функцій.

3. Стратегія високої-низької ціни – стратегія, коли компанія спочатку продає продукт за високою ціною, а потім, коли продукт падає в новизні чи релевантності, знижує цю ціну. Знижки, розділи розпродажу та розпродажі наприкінці року — це приклади високої та низької ціни в дії.

4. «Зняття вершків» - стратегія, коли компанії встановлюють найвищу можливу ціну за новий продукт, а потім з часом знижують ціну, оскільки продукт стає все менш популярним.

5. Стратегія «проникнення» полягає в установленні низьких цін з метою заохочення покупців, які чутливі до цін. Однак ціна проникнення не є стійкою в довгостроковій перспективі і зазвичай застосовується на короткий час.

6. Престижна ціна - установлення високої ціни на товари високої якості, унікальні товари.

7. Стратегія ціноутворення на основі цінності — це коли компанії оцінюють свої продукти або послуги на основі того, що клієнт готовий платити.

8. Пакетна цінова стратегія — це коли пропонується або об'єднується два або більше додаткові продукти або послуги разом і продаються за єдиною ціною.

9. Психологічна стратегія ціноутворення - це стратегія, яка використовує ціноутворення, щоб впливати на витрати або покупки клієнта, щоб збільшити продажі.

10. Ціна проходження за лідером – установлення ціни на основі тих, які пропонує ведуча фірма в галузі.

11. Нейтральна стратегія – установлення ціни, за основі фактичних витрат виробництва, враховуючи середню норму прибутковості в даній галузі.

12. Динамічне ціноутворення – це гнучка цінова стратегія, коли ціни коливаються залежно від ринку та попиту споживачів.

13. Персоналізоване ціноутворення – стратегія встановлення цін і рівнів знижок на основі відмінностей у споживчих атрибутах, уподобаннях, факторах цінності, чутливості до ціни та купівельній поведінці.

Важливо відрізнити персоналізоване ціноутворення від динамічного ціноутворення, оскільки ці два терміни часто використовуються як взаємозамінні, незважаючи на те, що вони мають різні значення. Персоналізоване ціноутворення передбачає стягнення зі споживачів різних цін у залежності від їх особистих характеристик, динамічне ж ціноутворення передбачає коригування цін з урахуванням змін попиту та пропозиції, часто в режимі реального часу, не маючи на увазі будь-якої дискримінації між споживачами. Різницю між «динамічним ціноутворенням» та «персоналізованим ціноутворенням» описано в роботах Г. Чорноус та Ю. Ярмоленко [20]. Автори зазначають, що «динамічне ціноутворення – це стратегічний процес, при якому ціни синхронно змінюють для всіх клієнтів залежно від попиту за даних умов ринку з метою досягнення цілей підприємства» [20, с.126]. Наголос на синхронній зміні робиться для чіткого розмежування понять «персоналізоване ціноутворення» та «динамічне ціноутворення». Персоналізоване або кастомізоване ціноутворення – «це процес встановлення індивідуальних цін на основі поточних об'єктивних умов ринку та суб'єктивних характеристик споживача з метою досягнення цілей підприємства» [20, с.126].

Учені виділяють персоналізоване ціноутворення як альтернативу цінової дискримінації, інші використовують саме для 1-типу цінової дискримінації. Пігу у своїй роботі «The economics of welfare» окреслив три типи дискримінації цін [21], описані нижче.

Цінова дискримінація першого ступеня належить до ситуації, коли з кожного споживача стягується індивідуальна ціна, рівна його максимальній готовності платити. Для цього потрібна точна інформація про готовність покупця платити. Цінова дискримінація першого ступеня дозволяє продавцям отримати весь споживчий надлишок. Насправді така крайня форма цінової дискримінації в практиці майже не зустрічається, оскільки продавці неспроможні дізнатися точну ціну резервування покупців. Цінова дискримінація першого ступеня є своєрідним еталоном для оцінки інших схем ціноутворення.

Цінова дискримінація другого ступеня належить до схем ціноутворення, за яких ціна товару чи послуги залежить від купленої кількості. Такі схеми також називаються «нелінійним ціноутворенням» і можуть включати знижку за кількість або подвійний тариф із фіксованою та змінною платою. Для цінової дискримінації другого ступеня продавцю не потрібна інформація про покупця, оскільки покупці обирають самі: вони обирають іншу ціну, обираючи іншу кількість.

При ціновій дискримінації третього ступеня ціни різняться між групами чи типами покупців. Це, наприклад, добре відомі знижки для студентів, дітей або людей похилого віку. Компанія також може стягувати з людей різних географічних регіонів різні ціни. Для цінової дискримінації третього ступеня немає необхідності впізнавати окремих покупців: продавцям потрібно знати лише характеристики покупця, які використовуються для дискримінації цін.

Цінова дискримінація в Інтернеті зазвичай працює аналогічно: інтернет-магазин ідентифікує покупця на основі, наприклад, файлу cookie, IP-адреси або інформації для входу в систему.

Отже, ціноутворення з розвитком технологій набуває нових форм у вигляді стратегій. Кожна компанія в залежності від сфери своєї діяльності та мети може обрати зручний для себе тип ціноутворення. Більш детально розглянуто такі популярні типи ціноутворення як динамічне та персоналізоване, так як в літературі часто плутаються ці поняття і взагалі мало хто виділяє персоналізоване ціноутворення в окремий тип.

## 1.2. Персоналізоване ціноутворення та практика його застосування

Поява аналітики великої кількості даних сприяла появі форми цінової дискримінації на основі профілів споживачів та їх поведінки в Інтернеті – персоналізованого ціноутворення. Персоналізоване ціноутворення набуває популярності, особливо в електронній комерції, адже від поведінки людей в Інтернеті можна зібрати дані, за допомогою яких онлайн-магазини можуть

запропонувати найкращу ціну для покупця. Але проблемою є те, що не завжди можна помітити даний тип ціноутворення.

Проаналізуємо поняття «персоналізоване ціноутворення» у наукових працях. Управління справедливої торгівлі у своєму звіті 2013 року зазначає, що персоналізоване ціноутворення – це практика, коли підприємство використовує інформацію, яку вони спостерігали, припускали та збирали про поведінку або характеристики людей з метою встановлювати різні ціни для різних споживачів (на індивідуальній або груповій основі) на основі того, що компанія думає, що люди готові платити [22]. У звіті OECD від 2018 року визначають персоналізоване ціноутворення як будь-яку практику цінової дискримінації кінцевих споживачів на основі їх особистих характеристик і поведінки, внаслідок чого ціни встановлюються у зростаючій залежності від бажання споживачів платити [23].

Персоналізоване ціноутворення - це процес «встановлення індивідуальних цін на основі поточних об'єктивних умов ринку та суб'єктивних характеристик споживача з метою досягнення цілей підприємства» [20, с. 126].

Ринковими умовами для персоналізованого ціноутворення є [23-24]

1. виявлення споживчої оцінки (можливість завдяки інформації про клієнтів оцінити готовність споживачів платити);
2. відсутність арбітражу (неможливість перепродати товар або послугу дорожче);
3. наявність елемента ринкової влади, щоб мати можливість встановлювати ціни вищі за граничні витрати.

Є як мінімум три важливі кроки, які має виконати будь-яка фірма, щоб запровадити персоналізоване ціноутворення:

1. Фірма повинна збирати дані про особисті характеристики та поведінку споживачів.
2. Фірма повинна використовувати зібрані дані для оцінки готовності споживачів платити.

3. Фірма, враховуючи передбачувану готовність платити, повинна вибрати оптимальну ціну для кожного споживача та вирішити, як реалізувати персоналізоване ціноутворення.

У даний час переважна більшість підприємств мають веб-сайт або домашню сторінку, і близько половини з них використовують соціальні мережі, що дозволяє збирати інформацію про IP-адреси, відвідування веб-сторінок, історичні покупки, кількість «вподобайок» за допомогою файлів cookie та інших інструментів автоматичного збирання даних. Невелика кількість підприємств проводить аналіз великих даних або використовує програмне забезпечення для управління взаємовідносинами з клієнтами, що може додатково допомогти компаніям персоналізувати послуги та диференціювати ціни для споживачів на індивідуальній основі.

Існування персоналізованого ціноутворення в чистому вигляді поки немає, але з розвитком технологій, збором даних про клієнтів компанії до цього наближаються. Зараз через збільшення інформації про споживачів та їх поведінку компанії можуть ділити клієнтів за схожістю та пропонувати їм в залежності від груп різний порядок товарів по релевантності до їх становища або пропонувати ціни відмінні від іншої групи. Та, звичайно, базуючись на поведінці клієнтів, аналітика великих даних сприяє появі різних форм цінової дискримінації на цифрових ринках, як поєднання традиційної дискримінації, описаної Пігу, з різними формами персоналізації [25-26]:

1. Управління ціною (steering pricing): у залежності від запитів споживачів у пошуковій системі встановлювати бал для конкретного споживача і в залежності від його пошукової історії на один той самий запит пропонувати більш дорогі або дешеві продукти.

2. Відволікаючі ціни (decoy pricing): метод ціноутворення, призначений для «примусу» покупця до вибору. Коли клієнти купують, їм часто доводиться вибирати між продуктами з різними цінами та характеристиками. І коли компанія вирішує максимізувати продаж одного конкретного продукту, вона

часто вибирає так звану структуру ціноутворення-приманки, щоб вплинути на споживача в його рішенні про покупку.

3. Крапельне ціноутворення (drip pricing): платформа може ввести споживачів в оману, показуючи низьку "стартову ціну" товару, до якої платформа автоматично додає додаткові збори до завершення покупки. Класичним прикладом крапельного ціноутворення замовлення авіаквитків, первісна ціна яких зазвичай низька, щоб привернути увагу «економних споживачів», але пізніше у процесі покупки додаються додаткові збори (наприклад, збори аеропортів, плата за паливо, реєстрація багажу).

4. Повторні пропозиції (Re-offers): споживач шукає певний товар, який хоче придбати, але не купує його; і згодом компанія зв'язується з потенційним клієнтом з пропозицією придбати товар зі знижкою. Наприклад, після пошуку певного типу продукту на Amazon або eBay без завершення покупки платформа може зв'язатися з потенційним покупцем, запропонувавши цей продукт за зниженою ціною. Отже, «терплячіші» споживачі зазвичай отримують вигідніші пропозиції при здійсненні покупок в Інтернеті в порівнянні з «менш терплячими» споживачами, які купують продукт, як тільки знаходять те, що шукають.

5. Підроблені спеціальні пропозиції (Fake special offers): платформа може створювати спеціальні підроблені пропозиції для певних категорій споживачів. Наприклад, Amazon або eBay можуть вказати, що спеціальна пропозиція є індивідуальною і доступною лише протягом обмеженого періоду часу, тоді як насправді «спеціальна» пропозиція відповідає єдиній ціні, що стягується платформою за продукт. У рамках цього сценарію менш досвідчені споживачі з більшою ймовірністю потраплять у «пастку» і приймуть спеціальну пропозицію платформи.

Успіх таких компаній, як Google і Ryanair, частково пояснюється їх здатністю пропонувати різні ціни своїм клієнтам, завдяки цьому вони можуть відійти від галузевих стандартів і збільшити свої доходи.

Є багато досліджень, які є доказом існування персоналізованого ціноутворення. Приклади наведені нижче. Складністю дослідження даного питання є те, що компанії не вказують про застосування персоналізованого ціноутворенням, і найчастіше його можна виявити лише при наявності кейсів споживачів або при експериментальному дослідженні.

У своїй роботі Езрахі та Стукс припустили, що Coupons.com використовує власні дані про покупців з метою таргетування цифрових купонів [26].

Згідно з даними Організації економічного співробітництва та розвитку AirAsia використовує машинне навчання, щоб зрозуміти, скільки пасажери готові платити та запропонувати персоналізовані тарифи для багажу [27-28].

За даними журналу «Wall Street Journal», американський сайт для бронювання житла Orbitz.com показував більш дорогі пропозиції для споживачів, які користувалися Mac, ніж для відвідувачів з Windows. Клієнти Mac платили на 30% більше, ніж користувачі Windows за нічне проживання. [29].

Розслідування, проведене «Wall Street Journal» у 2012 році, показало, що роздрібні торговці Staples та Home Depot, освітня технологічна компанія Rosetta Stone та фінансова компанія Discover Financial Services персоналізували ціни на основі різних споживчих характеристик: геолокації покупців, рівня доходу, історії переглядів та близькості до магазинів конкурентів [30].

Тестування персоналізованого ціноутворення відбувалося й на такому відомому маркетплейсі, як Amazon. Користувачі чатів на DVD.talk.com дізналися, що споживачі Amazon за однакові dvd-диски платять більше, ніж інші покупці. Один зі споживачів заявив, що він замовив музичний диск за 24,49 доларів, а наступного тижня побачив, що ціна виросла на 26,24 долара. А коли він стер cookies-файли з комп'ютера, то ціна знизилася до 22,74 долара. Отже, ціни для постійних клієнтів вищі, адже Amazon вже завоював постійних покупців і тому може встановлювати більш високі ціни [31].

Books.com використовує дизайн свого сайту для визначення споживачів, які найбільш чутливі до зміни ціни, що в свою чергу дозволяє встановлювати різні ціни для різних споживачів на один і той же товар. Після того, як споживач

знайшов необхідну книгу із зазначенням ціни, на веб-сторінці з'являється кнопка порівняти ціни, яка визначає ціни на точно таку ж книгу на сайтах інших компаній таких, як amazon.com. Якщо раптом виявиться, що ціна на book.com вища, ніж на сайтах інших компаній, то система автоматично знизить ціну для покупця [32].

Виявили наявність персоналізованого ціноутворення в мережах супермаркетів Safeway та Kroger, коли ціни встановлювалися вищі для лояльних споживачів та нижчі для тих споживачів, які чергують конкуруючі бренди [33].

Застосування персоналізованого ціноутворення було виявлено в американській компанії з пошуку та спільного використання поїздок Uber. У 2018 році вона впроваджувала нову систему ціноутворення, яка, схоже, відповідає категоріям персоналізованого ціноутворення. Компанія застосовує методи машинного навчання, щоб оцінити, скільки груп клієнтів готові витратити на поїздку [34]. Для цього компанія поділяє споживачів залежно від конкретного маршруту та часу доби, у який вони подорожують. Як показало дослідження, компанія встановлює вищі ціни для людей, що подорожують між багатими районами. Було виявлено різниці в ціні, наприклад, під час оплати поїздки між особистою кредитною карткою та корпоративною кредитною карткою. Також було висловлено припущення, що ціни можуть ґрунтуватися на параметрах, що спостерігаються: історії поїздок, марці телефону або навіть рівню заряду батареї телефону, що може суттєво вплинути на готовність споживача платити [35].

Прикладом використання персоналізованого ціноутворення є додаток для знайомств Tinder. Додаток має безкоштовний функціонал та преміальний Tinder Plus або Tinder Gold, які пропонують спеціальні функції, такі як необмежену кількість «гортань», «супер вподобань» і «підвищення». Під час реєстрації в Tinder споживачі зобов'язані ділитися особистою інформацією. І у тарифі Tinder Plus пропонується індивідуальна ціна в залежності від персональної інформації про користувача. Існують переконливі докази того, що вік є важливим фактором, що впливає на індивідуальне ціноутворення, оскільки для користувачів у віці 18-

29 років у всіх країнах, окрім Бразилії, стягується плата значно менша, ніж у літніх користувачів, – у середньому в шести країнах з 30-49-річних людей стягується на 65,3 % більше, ніж 18-29-річних, у той час як різниця між 30-49-річними та людьми, старшими 50 років була мінімальною. Однак результати також чітко демонструють, що вік не є єдиним релевантним фактором, оскільки спостерігалися численні випадки, коли старші користувачі платили менше, ніж їхні молодші співвітчизники [36].

Дослідники [37] розробили сотні фальшивих облікових записів, щоб перевірити, чи вплинули історії перегляду та покупок, а також кліки на сайтах на ціни, які бачать користувачі. Вони виявили, що шість із шістнадцяти популярних веб-сайтів електронної комерції стягують із споживачів різні ціни за подібний продукт, і жоден із сайтів не попереджав своїх клієнтів про цей факт. Наприклад, туристичні онлайн-агентства Cheaptickets і Orbitz віддавали перевагу учасникам цих сайтів і стягували з користувачів, які шукають готелі, у середньому на 12 доларів більше за ніч. Тим часом Travelocity стягував з користувачів мобільної операційної системи iOS Apple Inc. на 15 доларів менше за готелі, ніж інші користувачі. Priceline персоналізував порядок результатів пошуку на основі історії кліків і покупок користувача. Згідно з тим же дослідженням туристичні онлайн-агентства Expedia і Hotels.com, схоже, розбивають своїх користувачів на групи і спрямовують користувачів з однієї групи випадковим чином до більш дорогих продуктів. Expedia і Hotels.com є підрозділами Expedia Inc., і компанія підтвердила, що вона постійно вдосконалює свої стратегії ціноутворення за допомогою A/B-тестування. Клієнти випадковим чином поміщаються в групу, яка виділяє дешевше або дорожче [38]

Американський магазин канцелярських товарів Staples адаптував ціни до області, де проживали покупці, на основі IP-адрес покупців, не пов'язаних із вартістю доставки [39]. Через цю схему ціноутворення магазин стягував нижчі ціни з людей із районів із високим доходом.

Хоча персоналізоване ціноутворення економічно вигідне, багато компаній не хочуть на нього переходити [22]:

1. Побоюються втратити свою репутацію чи викликати негативну реакцію споживачів [40].
2. Персоналізоване ціноутворення є непрозорим.
3. Опитування Deloitte, у якому взяли участь понад 500 компаній [41], показало, що серед усіх роздрібних продавців, які впровадили штучний інтелект (ШІ) для персоналізації споживчого досвіду, 40% з них використовували ШІ спеціально для налаштування цін та рекламних акцій у режимі реального часу. Крім того, опитування споживачів, проведене Європейською комісією, показало, що від 12% до 20% споживачів мали негативний досвід, пов'язаний із персоналізованим ціноутворенням. В іншому дослідженні використовувалися облікові записи та файли cookie більше 300 реальних користувачів, щоб перевірити наявність цінової дискримінації на 16 популярних веб-сайтах електронної комерції, з яких 9 мали деякі елементи персоналізації [37].

Через неоднозначне ставлення до персоналізованого ціноутворення дослідники [42] вирішили здійснити опитування вибірки населення старшого 18 років в Нідерландах щодо їх ставлення до персоналізованого ціноутворення. Результати опитування показали, що більше половини населення стверджують, що ніколи не стикалися з ціновою дискримінацією в Інтернеті, і лише один із двадцяти стверджує, що стикався з нею часто чи дуже часто. Переважна більшість вважає персоналізоване ціноутворення неприйнятним та несправедливим. Більшість турбується про те, щоб платити за продукт не більше, ніж інші. Згідно із статистикою опитуваних прийняття персоналізованого ціноутворення збільшується з рівнем доходу та освіти та знижується з віком. Можна припустити, що зв'язок з освітою обумовлений кращим розумінням ділових інтересів практики ціноутворення або навіть позитивного впливу таких стратегій на розподіл та добробут. Чоловіки в середньому більш терпимі до цінової дискримінації та динамічного ціноутворення, ніж жінки. Люди набагато охочіше погоджуються на цінову дискримінацію та динамічне ціноутворення, якщо воно оформлене як знижка.

Переважає більшість населення висловилося за заборону цінової дискримінації в Інтернеті та вважає, що інтернет-магазини, які практикують цінову дискримінацію в Інтернеті, зобов'язані інформувати про це покупців. Однак заборона на цінову дискримінацію в Інтернеті може бути нерозумною з погляду економічного добробуту.

Можна виділити основні переваги та недоліки персоналізованого ціноутворення [43]:

Переваги персоналізованого ціноутворення:

1. Покупці з низькими доходами можуть отримати товари за кращі ціни.
2. Вища виручка компанії за рахунок гнучкого ціноутворення, але не завжди, іноді люди відмовляються від покупок через несправедливість.
3. Цінова дискримінація посилює цінову конкуренцію, що приносить користь споживачу, так як в середньому вони мають нижчу ціну.
4. Цінова дискримінація, яка набуває форми кодів знижок/купонів, може стимулювати вищий попит та більш високу споживчу корисність.

Недоліки персоналізованого ціноутворення:

1. Цінова дискримінація може підірвати довіру споживачів до ринків.
2. Цінова дискримінація може викликати занепокоєння у споживачів щодо конфіденційності через надмірний збір даних. Як наслідок, споживачі можуть «приховувати» свої дані, що може негативно вплинути на ефективність персоналізації. Основні потенційні переваги та ризики персоналізованого ціноутворення для споживачів наведені у табл.1.1 [44]:

Таблиця 1.1

Переваги та ризики персоналізованого ціноутворення

	Переваги	Ризики
Більша різниця в цінах	Нижчі ціни вигідні для споживачів, які мають нижчу платіжну спроможність.	Вищі ціни для деяких клієнтів, особливо якщо вони перебувають у вразливих умовах.

	Переваги	Ризики
Розширення ринку довіра до ринку	Більше клієнтів обслуговується, оскільки ціни знижуються для тих, чия здатність/бажання платити в іншому випадку залишила б їх необслуговуваними (або недостатньо обслугованими).	Більші розбіжності в цінах вважаються несправедливими та підривають довіру та задоволеність.
Вартість пошуку угоди	Постачальники знижують відтік, завчасно пропонуючи кращі пропозиції клієнтам, які мають ймовірність змінити/переукласти переговори. Також створення пропозицій, які краще відповідають індивідуальним потребам.	Вищі витрати на пошук, якщо постачальники приховують інформацію про більш конкурентоспроможні ціни від власних існуючих клієнтів. Конкурентам або новим учасникам важче завоювати частку ринку, оскільки відтік знижується за рахунок персоналізованого ціноутворення, що потенційно послаблює конкуренцію.
Вплив на середню ціну	Нижча середня ціна, якщо персоналізація ціни призводить до посилення конкуренції.	Вища середня ціна, якщо основним ефектом є краща ідентифікація клієнтів, які навряд чи перейдуть і готові платити більше. Так само вища середня ціна, якщо персоналізація призводить до слабкішої конкуренції через зниження прозорості

Отже, компанії не хочуть афішувати застосування персоналізованого типу ціноутворення через те, що часто це призводило до обурення клієнтів, які негативно ставляться до цінової дискримінації. Тому й важко з'ясувати наявність даного типу ціноутворення. Проаналізовано практичне застосування компаніями персоналізованого ціноутворення. Персоналізоване ціноутворення має як свої переваги, так і недоліки для споживачів: вони окреслені в роботі.

## РОЗДІЛ 2. МЕТОДИ МОДЕЛЮВАННЯ ПЕРСОНАЛІЗОВАНОГО ЦІНОУТВОРЕННЯ

### 2.1 Підходи, моделі та інструменти Data Science для підтримки персоналізованого ціноутворення

Можливість компаній збирати дані про клієнтів, накопичувати їх сприяло розвитку методів для моделювання процесу ціноутворення. Інтерес до моделей зріс і з технологічним розвитком, адже більш точно можна спрогнозувати ціну, що призведе до більших доходів компанії.

Алгоритми персоналізованого ціноутворення забезпечують гнучкість, оскільки рітейлери можуть встановлювати ціни, орієнтуючись на різні групи покупців, створюючи оптимальну пропозицію за цінами, ґрунтуючись на тенденції ринку, коливаннях попиту, поведінці покупців, купівельної спроможності і безлічі інших чинників.

На практиці ціноутворення набагато складніше, ніж здається. Адже, крім собівартості товару чи послуги, включаються й інші фактори, які впливають на ціну.

У літературі не існує чіткої класифікації моделей персоналізованого ціноутворення й чітких визначень щодо економічної категорії «персоналізоване ціноутворення», і кожен із науковців вкладає різне в дане поняття. У загальному задача моделювання спрямована на розв'язання оптимізаційних завдань – найчастіше, максимізація прибутку.

Моделі персоналізованого ціноутворення можна класифікувати за ступенем етичності [45]:

1. Етичні - моделі, які використовують лише агреговану інформацію про клієнтів з їх дозволу, що не дозволяє використання різних стратегій до різних груп клієнтів.

$$\left\{ \begin{array}{l} \sum_{t=1}^T \sum_{i \in I(t)} \pi(P_t(i), t) \rightarrow \max \\ P_t(i) = f(h: h \in \mathbb{F} \cap h \notin \mathbb{C}, t) \end{array} \right.$$

де  $T$  - досліджуваний часовий горизонт,

$i \in I(t)$  – товари, продані у період  $t$ ,

$P_t$  – функція для встановлення ціни товару у період  $t$ ,

$f$  – деяка функція залежності ціни від факторів  $h$ ,

$\mathbb{F}$  – множина усіх факторів системи,

$\mathbb{C}$  – множина характеристик споживачів,

$\pi$  та  $r$  – функції комерційного зиску та соціальних втрат від продажу товару  $i$  за ціною  $P_t(i)$  у період  $t$ .

2. Псевдоетичні – моделі, у яких однією із оптимізаційних завдань є «максимізація суспільного блага», проте цінова стратегія дозволяє встановлення різних цін для покупців.

$$\begin{cases} \sum_{t=1}^T \sum_{i \in I(t)} r(P_t(i), t) \rightarrow \min \\ P_t(i) = f(h; h \in \mathbb{F}, t) \end{cases}$$

3. Неетичні – моделі, які використовують персональну інформацію про клієнтів без належного інформування про використання даних або про застосування різних цін для різних клієнтів.

$$\begin{cases} \sum_{t=1}^T \sum_{i \in I(t)} \pi(P_t(i), t) \rightarrow \max \\ P_t(i) = f(h; h \in \mathbb{F}, t) \end{cases}$$

Класифікація методів моделювання в залежності від математичного інструменту[46]:

1. Моделі на основі запасів: це моделі, які в основному базуються на рівнях запасів і рівнях обслуговування клієнтів.

2. Моделі, керовані даними: моделі, які використовують статистичні методи для використання наявних даних про переваги клієнтів і моделі купівлі для обчислення оптимальних цін.

3. Моделі теорії ігор: у сценарії з кількома продавцями продавці можуть конкурувати за один і той же пул клієнтів, і це викликає цінову гру між продавцями.

4. Моделі машинного навчання: ринок електронного бізнесу надає багатий майданчик для онлайн-навчання покупців і продавців. Продавці потенційно

можуть дізнатися про переваги покупців і моделі купівлі, а також використовувати алгоритми для динамічної оцінки своїх пропозицій, щоб максимізувати дохід або прибуток.

5. Імітаційні моделі: імітаційна модель для персоналізованого ціноутворення може використовувати будь-яку з чотирьох моделей, зазначених вище, або використовувати прототип системи або будь-який інший спосіб імітації динаміки системи.

Розглянемо моделі, які застосували науковці в своїх працях для персоналізованого ціноутворення та дослідимо їх особливості.

А. Аккісті та Х. Р. Варіан [47] розглядають дві групи клієнтів, для кожної групи пропонується різна ціна.

#### **Припущення моделі:**

1. Нульова ставка дисконту.
2. Відсутній перепродаж.
3. Якщо споживач байдужий, він діятиме так, як вважає за краще продавець.
4. У продавця є механізми збору інформації про історію покупок клієнтів за допомогою програм лояльності, номерів кредитних карток, файлів «cookie».
5. Нехай,  $v_H$  – готовність платити за товар за вищою ціною (1 група клієнтів) та  $v_L$  – готовність платити за товар за нижчою ціною (2 група клієнтів).
6.  $\pi$  - частка населення, яка готова платити за вищою ціною.
7.  $p_H$  – теперішня ціна, яка стягується з покупців з 1 групи,  $p_L$  – теперішня ціна, яка стягується з покупців 2 групи.
8.  $x_H$  — загальна кількість, споживана 1 групою клієнтів, а  $x_L$  — загальна кількість, споживана 2 групою клієнтів.

**Метод дослідження:** задача нелінійного програмування - максимізація прибутку.

**Формалізація:**

$$\begin{aligned} \max_{x_H, x_L, p_H, p_L} \quad & \pi p_H + (1 - \pi) p_L, \\ & v_H x_H - p_H \geq v_H x_L - p_L, \\ & v_H x_H - p_H \geq 0, \\ & v_L x_L - p_L \geq v_L x_H - p_H, \\ & v_L x_L - p_L \geq 0 \end{aligned}$$

**Перевага** – приносить більший дохід, ніж при застосуванні фіксованої ціни для усіх груп.

**Недолік** – не гнучкість моделі – наявність усього двох груп споживачів (з високою та низькою очікуваною ціною за товар).

Р. Філіпс [48] у своїй роботі запропонував підхід до моделювання персоналізованого ціноутворення, яке передбачає сегментацію клієнтів та для кожного сегменту встановлювати індивідуальну ціну з метою максимізації прибутку.

**Припущення моделі:**

1. Кількість запитів клієнтів в сегменті не залежить від ціни.
2. Нехай  $N$  — кількість цінових сегментів.
3. Для кожного сегмента  $i$  полягає задача у визначенні ціни  $p_i$ .
4.  $D_i$  - кількість запитів клієнтів, які будуть отримані в сегменті  $i$  протягом деякого майбутнього періоду.
5.  $p_i(p_i)$  – функція запиту та відповіді, яка показує ймовірність того, що клієнти у сегменті  $i$  куплять продукт, якщо продавець поставить ціну  $p_i$ .
6.  $f_i(p_i)$  – функція приросту прибутку для сегмента  $i$  як функція ціни.

**Метод дослідження:** математичний - максимізація загального очікуваного прибутку для усіх сегментів.

**Формалізація:**

$$\max_p \sum_{i=1}^N D_i p_i(p_i) f_i(p_i)$$

**Переваги:** розглядається  $N$  сегментів клієнтів, для кожного з них встановлюється персональна ціна.

**Недоліки:** немає чітких факторів для сегментування клієнтів; відсутність обмежень у функції оптимізації щодо цін.

Ч. Нью, Ч. Чжен, Ф. ВуЮ Ш. Тан та Г. Чан [49] провели дослідження з ціноутворення для продажу персональних даних клієнтів (рейтинги продуктів, використання електрики, дані соціальних мереж, записи про стан здоров'я, фізичну активність та траєкторії руху). Автори розглядають залежність між ціною за дані та рівнем їх конфіденційності. Якщо витік даних має конфіденційний характер, то й ринкова вартість за них вища.

#### Припущення моделі:

1. Ринкова вартість товару є детермінованою функцією.
2. Враховується невизначеність ринкової вартості кожного запиту.
3.  $R_t$  – штрафна функція в період  $t$ .
4.  $p_t$  – ціна продажу в період  $t$ .
5.  $q_t$  – витрати на компенсацію витоку персональних даних в період  $t$ .
6.  $v_t$  – ринкова вартість персональних даних в період  $t$ .
7.  $\delta_t$  – ринкова невизначеність в період  $t$ .

**Метод дослідження:** метод нелінійного програмування - мінімізація штрафів від продажу персональних даних клієнтів.

#### Формалізація:

$$\sum_t R_t \rightarrow \min$$

$$R_t = \begin{cases} 0, & q_t > v_t \\ \max_{p_t^*} p_t^* \Pr(p_t^* \leq v_t) - p_t 1\{p_t \leq v_t\}, & q_t \leq v_t \end{cases}$$

**Переваги:** можливість підтримувати як лінійні, так і нелінійні залежності ринкової вартості, при цьому допускаючи певну невизначеність.

**Недоліки:** не чітко окреслено питання оцінювання витрат за компенсацію витоку персональних даних.

В. Дж. Аллендер, Й. Ляуконите, Ш. Насер та Т. Дж. Річардс [50] у своїй роботі моделюють персоналізоване ціноутворення на основі готовності платити за товар кожного покупця.

#### Припущення моделі:

1. Є два покупці  $a$  та  $b$ , які мають унікальну готовність платити за товар (WTP).
2. WTP покупців є незалежними та рівномірно розподіленими на проміжку  $[\underline{w}; \bar{w}]$ .
3. Продавець виробляє товари при постійних граничних витратах  $c \in [0; \underline{w}]$ .
4. Кожен покупець знає тільки свою WTP.
5. Покупець  $i$  знає ціну, запропоновану продавцем покупця  $j$ .
6.  $p_i$  - ціна, яку продавець пропонує покупцю  $i$  з набором інформації  $\Omega_i$ , прийнявши пропозицію, покупець отримує корисність  $U_i$ .
7.  $P_j$  – випадкова подія для покупця  $i$ .

**Метод дослідження:** математичний – максимізація прибутку.

#### Формалізація:

$$U_i(w_i, p_i, p_j) = \alpha(w_i - p_i) - \delta(p_i - c) - \rho_d \max\{0, p_i - p_j\} \geq 0$$

для  $\forall i, j \in \{a, b\}, i \neq j$

$\alpha(w_i - p_i)$  - об'єктивний надлишок покупця, зважений за параметром  $\alpha > 0$ .

$\delta(p_i - c)$  - інтереси покупця про справедливість розподілу, тобто наскільки він дбає про надлишок, який продавець залишає собі, зважений на  $\delta > 0$ .

$\rho_d \max\{0, p_i - p_j\}$  – занепокоєння справедливістю («peer-induced fairness»), що інший покупець отримає нижчу ціну, виважену на  $\rho_d > 0,2$ .

$$\max_{p_a, p_b} \pi = p_a + p_b$$

**Переваги:** розглядається занепокоєння справедливістю клієнта; враховується готовність покупця платити.

**Недоліки:** не враховані проблеми справедливості споживача: альтруїзм, зрада; розглядається всього два типи покупців.

Н. Каллус та А. Чжоу у своїй роботі [51] пропонують персоналізацію цін на основі прогнозованого моделювання попиту з урахуванням коваріатів: встановлення індивідуальних відсоткових ставок, цільові знижки на споживчі товари та персоналізовані субсидії дефіцитних ресурсів із позитивними зовнішніми ефектами, такими як вакцини.

#### **Припущення моделі:**

1.  $X$  - коваріати (особливості клієнта): вік, географічні фактори ризику, довіра, уявлення та знання про хвороби, від яких вакцинуються тощо
2.  $A$  - захищений атрибут, розглядається порівняння між двома захищеними групами  $A=a$ ,  $A=b$ .
3. Персоналізована цінова політика — це функція, яка відображає коваріати та, можливо, захищену інформацію про атрибути на реальну ціну  $p(X, A) \in \mathbb{R}^+$ .
4. Кожен продаж пов'язаний з невідомою кривою попиту, що представляє попит для кожної можливої ціни  $D(p) \in \mathbb{R}^+$ .
5. Оцінка персоналізованого попиту за ціною, заданою коваріатами позначається  $D(p|x, a) = \mathbb{E}[D(p)|X = x]$ .

**Метод дослідження:** параметрична, напівпараметрична модель.

#### **Формалізація:**

$$p^*(x, a) \in \operatorname{argmax}_{p(x,a)} \mathbb{E}[D(p(X, A))p(X, A)]$$

**Переваги:** враховуються персональні дані клієнта, що говорить про більш точне ціноутворення для кожного споживача

**Недоліки:** треба ретельно підбирати коваріати, щоб не виникло дискримінації серед клієнтів, що призведе до репутаційних втрат компанії.

У джерелі 52 зазначається, що популярними інструменти Data Science у 2022 році є SAS, Apache Spark, Tableau, BigML, D3.js, Excel, Google BigQuery, Java, Python, MySQL, Matplotlib, Sickit-learn, TensorFlow, ggplot2.

Отже, розглянуто методи, які використовуються при моделювання персоналізованого ціноутворення. Методи для моделювання було класифіковано

за етичністю: етичні, псевдоетичні, неетичні; у залежності від застосованого математичного інструментарію: моделі на основі запасів, моделі керовані даними, моделі теорії ігор, моделі машинного навчання та імітаційні моделі. Вивчено підходи до моделювання персоналізованого ціноутворення, запропоновані науковцями, зазначено їх переваги та недоліки.

## 2.2 Використання методів машинного навчання до персоналізованого ціноутворення

Стратегії ціноутворення, засновані на даних, стають усе більш поширеними, коли клієнтам пропонується персоналізована ціна на основі характеристик, які передбачають їхню оцінку продукту. Однак спроби включити машинне навчання до структури ціноутворення часто призводять до складної політики ціноутворення, яку неможливо інтерпретувати, що призводить до повільного впровадження на практиці.

З появою великих наборів даних (Big Data) про індивідуальні характеристики споживачів, покупки та поведінку в Інтернеті зросла кількість компаній, які вирішили запровадити персоналізоване ціноутворення на основі підходів Data Science з використанням машинного навчання та штучного інтелекту.

Наука про дані охоплює підготовку даних для аналізу, їх обробку, виконання розширеного аналізу даних і представлення результатів.

Продавці роздрібною торгівлі, які впроваджують персоналізоване ціноутворення, використовують потенціал машинного навчання для оптимізації цін, щоб задовольнити потреби своїх клієнтів за допомогою встановлення цін на основі даних, залишатися конкурентоспроможними та підвищувати прибуток.

Завдяки штучному інтелекту роздрібні продавці можуть розділити клієнтів на сотні груп на основі їхньої готовності платити. Алгоритм використовує інформацію про покупців, що допомагає класифікувати клієнтів, приналежних до певної групи. Для аналізу беруться такі показники споживачів: вік, стать, місцезнаходження та поведінки в Інтернеті. При використанні технологій

машинного навчання враховують як цінові, так і нецінові фактори. Це дозволяє роздрібним торговцям встановлювати збалансовані ціни в режимі реального часу.

У літературі виділяють найчастіше 3 типи алгоритмів машинного навчання [53-54]:

1. Навчання з вчителем (Supervised learning) - алгоритм складається із залежної змінної, яку необхідно передбачити на основі заданого набору незалежних змінних. Використовуючи набір змінних, дослідник генерує функцію, яка співставляє вхідні дані та бажані вихідні дані. Процес навчання триває до тих пір, поки модель не досягне бажаного рівня точності на даних навчання. Поширені алгоритми: лінійна регресія, дерево рішень, випадковий ліс, KNN, SVM, наївний Байєс, логістична регресія, нейронні мережі тощо

2. Навчання без вчителя (Unsupervised Learning) - алгоритм не має цільової або кінцевої змінної для прогнозування або оцінки. Алгоритм намагається використовувати методи обробки вхідних даних з метою виявлення закономірностей, а також підсумовування та угруповання точок даних, які допомагають отримати інформацію та краще описують дані для користувачів. Поширені алгоритми: кластеризація k-середніх, правила асоціації.

3. Навчання з підкріпленням (Reinforcement Learning) – метод спрямований на використання спостережень, отриманих при взаємодії з навколишнім середовищем, для вжиття заходів, що максимізують винагороду або мінімізують ризик. Поширені алгоритми: Q-Learning, Temporal Difference (TD), Deep Adversarial Networks.

Дж. Браунлі у своїй книзі «Machine Learning Mastery» [55] наводить ширшу класифікацію алгоритмів машинного навчання: глибоке навчання, ансамблі, нейронні мережі, регресії, дерево рішень, кластеризація, моделі регуляризації, системи правил, Баєсові алгоритми, алгоритми зменшення розмірності, навчання на прикладах.

Розглянемо алгоритми, які застосовувалися при моделюванні персоналізованого ціноутворення.

Залежно від математичної моделі підприємства можуть створювати численні алгоритми, які відповідають їхній стратегії динамічного та персоналізованого ціноутворення. Кілька підходів, зібраних із дослідницьких статей [56]:

### 1. Байєсова модель

У байєсівській моделі користувач вибирає апріорне значення, що вказує вихідну думку про можливу ціну або значення, скільки споживач бажає заплатити за товар чи послугу (WTP). Потім, щоразу, як у алгоритм вводиться нова точка даних, вихідне переконання зміщується або вище, або нижче. Більшість байєсівських моделей використовують історичні дані про ціну як найважливішу функцію визначення остаточної ціни.

### 2. Модель навчання з підкріпленням

Модель ціноутворення навчання з підкріпленням досліджує дані про попит клієнтів, беручи до уваги сезонність, ціни конкурентів та невизначеність ринку, при цьому кінцевою метою є високий дохід.

### 3. Модель дерева рішень

Дерева рішень – це моделі машинного навчання класифікації, які виводять деревоподібну модель рішень та їх можливих наслідків, включаючи можливість певного результату, витрати ресурсів та корисність. Алгоритми дерев рішень для ціноутворення допомагають підприємствам зрозуміти, які параметри найбільше впливають на ціни, і яка ціна передбачає найвищий дохід. Використовуючи цю інформацію, алгоритм прогнозує найкращий діапазон цін для кожного товару.

У табл. 2.1 наведено алгоритми Data Science, які використовувалися в наукових працях для моделювання персоналізованого ціноутворення.

Таблиця 2.1

Алгоритми машинного навчання, які застосовувалися до персоналізованого ціноутворення

Дослідники, рік	Алгоритм
Й. Бауер та інші, 2017 [57]	Kernel Regression
Н. Каллус, 2017 [58]	CART, OLS, logistic regression, KNN, PT, PF, OPT

Дослідники, рік	Алгоритм
Ч. Чжоу та інші, 2018 [59]	Cross-fitted Augmented Inverse Propensity Weighted Learning (CAIPWL), Inverse Propensity Weighted Learning (IPWL)
Р. Маестре та інші, 2018 [60]	Q-Learning, Reinforcement Learning
Д. Берцимас та інші, 2019 [61]	Optimal Prescription Trees (OPT), Random Forest, CART, Boosting
Н. Шукла, 2019 [62]	Gaussian Naive Bayes with clustered features (GNBC), Deep Neural Network, Random Forest
В. Місик, 2020 [63]	Random Forest, Hierarchical Bayesian
Я. Сатсангі та інші, 2020 [64]	Linear Regression, XGBoost, Random Forest, SVR, MLP
М. Дж. Аван та інші, 2020 [65]	Random Forest, Linear Regression
А. Елмахтуб, 2020 [66]	Smart Predict-then-Optimize Tree (SPOT), CART
М. Біггс та інші, 2020 [67]	LightGBM, Personalization Tree, Causal Tree, Student Prescriptive Tree
М. Д. Іслам, 2021 [68]	Linear Regression, Random Forest, XGBoost
А. Ахуджа, 2021 [69]	Boosting, Ridge, Deep Neural Network, XGBoost, LightGBM

Отже, як показують дані в таблиці, найпопулярнішими є алгоритми дерев рішень.

К. Намара [70] виділяє найкращими алгоритми для прогнозу цін Xgboost, Catboost та Lightgbm.

Більш детально розглянемо алгоритми, які є популярними в наукових працях для моделювання персоналізованого ціноутворення та зацікавили автора роботи.

У роботі [71] підкреслюється, що використання Ridge regression є доцільним при наявності мультиколінеарності в даних, що є частою проблемою при персоналізованому ціноутворенні. При зборі інформації про клієнтів може бути схожа інформація, яка сильно корелює між собою. Взагалі Ridge regression була розроблена як можливе вирішення проблеми неточності оцінок методом

найменших квадратів, коли моделі лінійної регресії мають мультиколінеарні (сильно корельовані) незалежні змінні. За допомогою методу регуляризації L2 вирішується проблема мультиколінеарності множинної регресії [72].

**Переваги:** працює з мультиколінеарними даними.

**Недоліки:** в даних має бути лінійна залежність.

Частим є застосування Random Forest Regression для персоналізованого ціноутворення. У роботі [73] зазначається, що аргументами для використання даного алгоритму є те, що модель вимагає мінімальної підготовки даних та легко працює з категоріальними, числовими та логічними даними без нормалізації їх. Особливо це є вигідним, якщо компанія напямую хоче тягнути дані з сайту та будувати моделі ціноутворення. Також модель Random Forest може допомогти з вибором факторів, які впливають на ціну та автоматично змінювати їх в залежності від даних.

Random Forest Regression - це ансамбль дерев рішень. Це означає, що множина дерев, побудованих певним «випадковим чином», утворюють випадковий ліс. Кожне дерево створюється з різних вибірок рядків, і в кожному вузлі для поділу вибирається різна вибірка ознак. Кожне з дерев робить свій індивідуальний прогноз. Потім ці прогнози усереднюються для отримання єдиного результату [74].

**Переваги:** гнучкий алгоритм, не потребує нормалізації даних.

**Недоліки:** потребує високої обчислюваної потужності, необхідно багато часу на навчання, важкий в інтерпретації.

Моделі градієнтного бустінгу являються одними з найпопулярніших методів у сучасному машинному навчанні, в тому числі й при персоналізованому ціноутворенні. Взагалі виділяють три найпопулярніших алгоритму: XGBoost, LightGBM, CatBoost.

Gradient Boosting Regression - це метод машинного навчання, який подає модель прогнозування у вигляді ансамблю слабких моделей, похідний від дерева рішень. Модель дерев з градієнтним посиленням будується поетапно, як і в

інших методах підвищення, але вона узагальнює інші методи, дозволяючи оптимізувати довільну функцію, що диференціюється, втрат [75].

XGBoost Regression - це алгоритм градієнтного бустингу, який заснований на деревах рішень. При використанні підвищення градієнта для регресії слабкими учнями є дерева регресії, і кожне дерево регресії зіставляє точку вхідних даних з одним із його листя, що містить безперервну оцінку. XGBoost мінімізує регуляризовану (L1 та L2) цільову функцію, яка поєднує в собі опуклу функцію втрат (засновану на різниці між прогнозованими та цільовими вихідними даними) та штрафний член за складність моделі (іншими словами, функції дерева регресії). Навчання продовжується ітеративно, додаючи нові дерева, які передбачають залишки або помилки попередніх дерев, які потім поєднуються з попередніми деревами для отримання остаточного прогнозу. Це називається підвищенням градієнта, тому що воно використовує алгоритм градієнтного спуску для мінімізації втрат при додаванні нових моделей [76].

**Переваги:** ефективно працює з великим набором даних, нелінійними даними; не потребують нормалізації даних.

**Недоліки:** можуть перенавчатися, особливо якщо дерева глибокі; використовує лише числові дані.

LightGBM Regression - це алгоритм градієнтного бустингу, який заснований на деревах рішень. Light GBM розбиває дерево по листах, тоді як інші алгоритми підвищення розбивають дерево по глибині або за рівнями, а не по листах. Інакше кажучи, Light GBM вирощує дерева вертикально, тоді як інші алгоритми вирощують дерева горизонтально. [77-78]

**Переваги:** швидке навчання, менше використання пам'яті, висока точність, використовується при великому об'ємі даних.

**Недоліки:** можливе перенавчання, погано працює з малим об'ємом даних.

CatBoost регресія - це алгоритм градієнтного бустингу, який заснований на деревах рішень. CatBoost будує симетричні (збалансовані) дерева. На кожному кроці листя з попереднього дерева розщеплюються за однаковою умовою. Для всіх вузлів рівня вибирається та використовується пара функція-розділ, на яку

припадає найменша втрата. Ця збалансована архітектура дерева сприяє ефективній реалізації ЦП, зменшує час передбачення, робить швидкі застосування моделі та контролює переобладнання, оскільки структура служить регуляризацією [79-80].

**Переваги:** здатність інтегрувати різні типи даних, такі як зображення, аудіо або текстові функції в одну структуру; швидкість обрахунку.

**Недоліки:** необхідне більше сховище даних через наявність категоріальних даних.

Отже, підходи Data Science, в особливості алгоритми машинного навчання набирають популярності у сфері ціноутворення. Через збільшення об'єму інформації про клієнтів, яку можуть збирати компанії на власних сайтах, вони можуть пропонувати персоналізовані ціни для клієнтів, а для цього потрібні потужніші обчислювальні можливості, які можуть бути вирішені за допомогою застосування підходів машинного навчання. Також Data Science в сфері персоналізованого ціноутворення може запропонувати нові метод для виявлення важливих факторів, які впливають на ціну. Було розглянуто основні алгоритми Data Science, підходи по класифікації: навчання під наглядом, навчання без нагляду, підкріплене навчання. Проаналізовано алгоритми машинного навчання, які застосовувалися в роботах до персоналізованого ціноутворення, описано найпопулярніші з них та зазначено їх переваги та недоліки.

## РОЗДІЛ 3. МОДЕЛЮВАННЯ ЦІНОУТВОРЕННЯ НА ПРИКЛАДІ МАРКЕТПЛЕЙСУ «AIRBNB»

### 3.1 Постановка задачі й методологія

Персоналізоване ціноутворення у переважній більшості використовує персональну інформацію про клієнтів та традиційні фактори ціноутворення для досягнення конкретних цілей продажу, таких як максимізація прибутку, максимізація обсягу продажу та мінімізація часу продажу. Це призводить до встановлення різних цін на одні й ті ж товари різним клієнтам.

Для моделювання персоналізованого ціноутворення було обрано дані відомої платформи бронювання житла Airbnb [1].

Airbnb - це онлайн-платформа для перегляду та бронювання житла. Це маркетплейс, де взаємодіють господарі житла та люди, які хочуть забронювати житло. Airbnb з'єднує господарів і мандрівників, що полегшує процес оренди. Крім того, цей сервіс розвиває економіку спільного використання, дозволяючи власникам нерухомості здавати в оренду приватні квартири.

Якщо розглядати маркетплейс як об'єкт персоналізованого ціноутворення, то на формування ціни менеджментом платформи може впливати не тільки інформація про споживачів, як у звичайному інтернет-магазині, а й інформація про продавців. Тобто в маркетплейсі можна виділити два компоненти персоналізації: продавці та покупці. Якщо розглядати Airbnb, то продавцями є господарі, які здають в оренду своє житло, а покупцями – гості, які бажають забронювати житло. Визначення інструментів підтримки персоналізованого ціноутворення для маркетплейсів практично не розглядалося у науковій літературі, що підтверджує новизну досліджень, проведених в цій роботі, та їх результатів.

У цій роботі ми розглядаємо лише вплив інформації щодо господарів житла на ціну, але їх специфікою є те, що в деяких з досліджуваних даних опосередковано враховується й інформація про споживачів (оцінка споживачем конкретного господаря, його житла, тощо).

При розв'язанні задачі персоналізованого ціноутворення необхідно:

- з’ясувати, від яких факторів залежить ціна, як вона змінюється залежно від інформації про господаря житла (персональних даних), динамічних та традиційних факторів;
- розглянути можливість застосування персоналізованого ціноутворення для маркетплейсів;
- підтвердити чи спростувати гіпотезу щодо можливості використання персоналізованого ціноутворення із врахуванням інформації про господаря житла.

Тобто, базовим завданням є з’ясувати, чи впливає інформація про господаря житла на ціну. А також запропонувати методи аналізу та побудови моделей для персоналізованого ціноутворення.

Моделювання персоналізованого ціноутворення відбувалося у декілька етапів:

1. Завантаження та аналіз даних.
2. Обробка даних.
3. Побудова та відбір моделей.
4. Аналіз отриманих результатів.

Для побудови моделі були розглянуті дані про майбутнє бронювання житла в м. Берлін (17.12.2021-17.12.2022) з ціною за житло та інформацію, чи ще наявне житло (таблиця з даними «Calendar»), а також зібрані із сайту Airbnb за 17.12.2021-18.12.2021 дані про опис житла та господаря житла (тип житла, кімнати, відгуки, розміщення та характеристики). [таблиця з даними «Listing»]

Першим етапом обробки було з’єднання даних двох таблиць, як показано на рис. 3.1.



Рис. 3. 1. Схема зв’язків з даними

Джерело: розроблено автором

Об'єднавши дані, провели аналіз та обробку даних: проаналізовано наявність пропущених значень та видалено ті стовпці, що мають у собі більше 30% пустих значень; змінено пропущені значення на середнє значення. Потім проведено декодування даних у логічні змінні; перекодування категоріальних даних для подальшого застосування їх у моделях у логічні або числові значення. Переведено дати в числові значення, у значення місяця, дня тижня з метою аналізу динамічних факторів та видалені всі рядки, які не мають великий обсяг пропусків.

На рис. 3.2 проілюстровано кількість наявного житла на дату. Найбільше заброньовано житло на найближчий місяць, згодом зростає кількість вільних місць до березня тому, що в Берліні в цей період холодно й мало туристів його відвідує, а на весну-літо кількість зайнятих місць житла зростає, адже цей період вважаються найбільш популярним для відпочинку у туристів та завчасне бронювання житла дає змогу туристам зекономити.

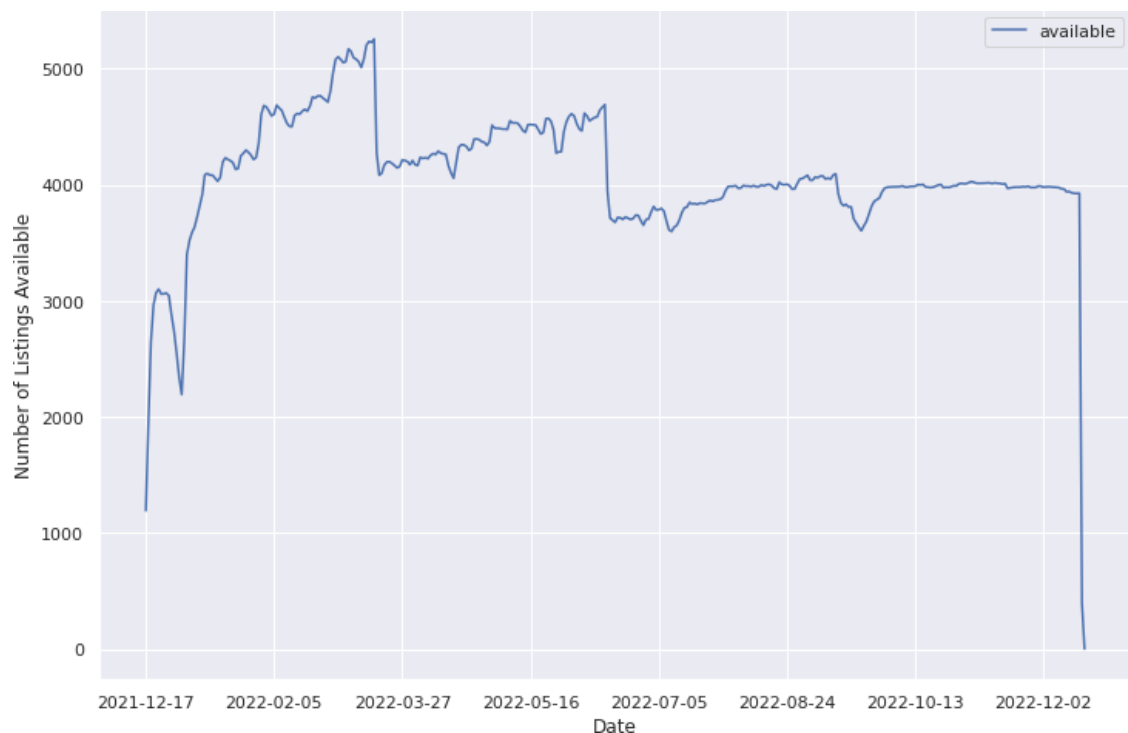


Рис. 3. 2. Кількість вільного житла на дату

Джерело: створена автором на основі даних [1] у програмному середовищі Python

Досліджуючи середньомісячну ціну та кількість вільного житла помісячно, ми помітили, що існує тенденція зміни ціни та кількості бронювання протягом року. Це показано на рис 3.3. найдешевші ціни в лютому, вони зростають протягом року, а щодо вільного житла, то зворотна тенденція. найдешевшими місяцями вважаються січень-квітень, а найдорожчими вересень та грудень.

Якщо досліджувати наявність вільних місць протягом року, то до липня поступово зменшується їх кількість, протягом липня-листопада тримається на одному рівні, а в грудні спостерігається значний стрибок униз. Різка зміна ціни та кількості вільних місць у грудні пояснюється тим, що туристи часто приїжджають у Берлін на Різдво, щоб відвідати різдвяні ярмарки.

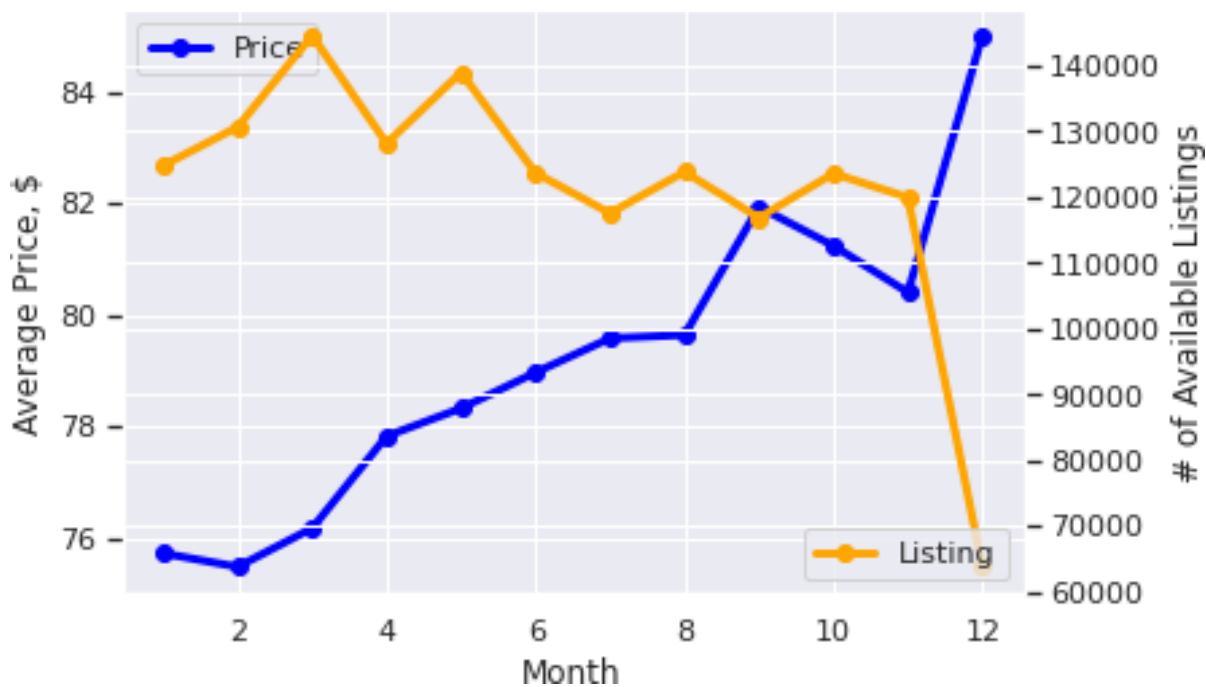


Рис. 3. 3. Сезонна зміна ціни та вільного житла

Джерело: створена автором на основі даних [1] в програмному середовищі Python

Проаналізуємо залежність ціни від дня тижня (рис. 3.4). Графік показує незначне коливання ціни та кількості заброньованих місць протягом тижня. Коливання протягом тижня спостерігається в розмірі 2 доларів США.

Найдешевшим житло є, якщо бронювати з понеділка-середи та неділі. Підвищення ціни спостерігається з четверга на п'ятницю та становить 2 долари

США. Найдорожчими є ціни з п'ятниці-суботи. Залежність по кількості вільних місць має зворотний зв'язок. Найбільша кількість вільних місць спостерігається у понеділок та вівторок, потім поступово падає до четверга. Найбільше заброньованого житла з суботи, бо це вихідний день і туристам зручно заселятися в цей день, а з неділі мала кількість споживачів бронює житло, тому кількість вільних місць висока.

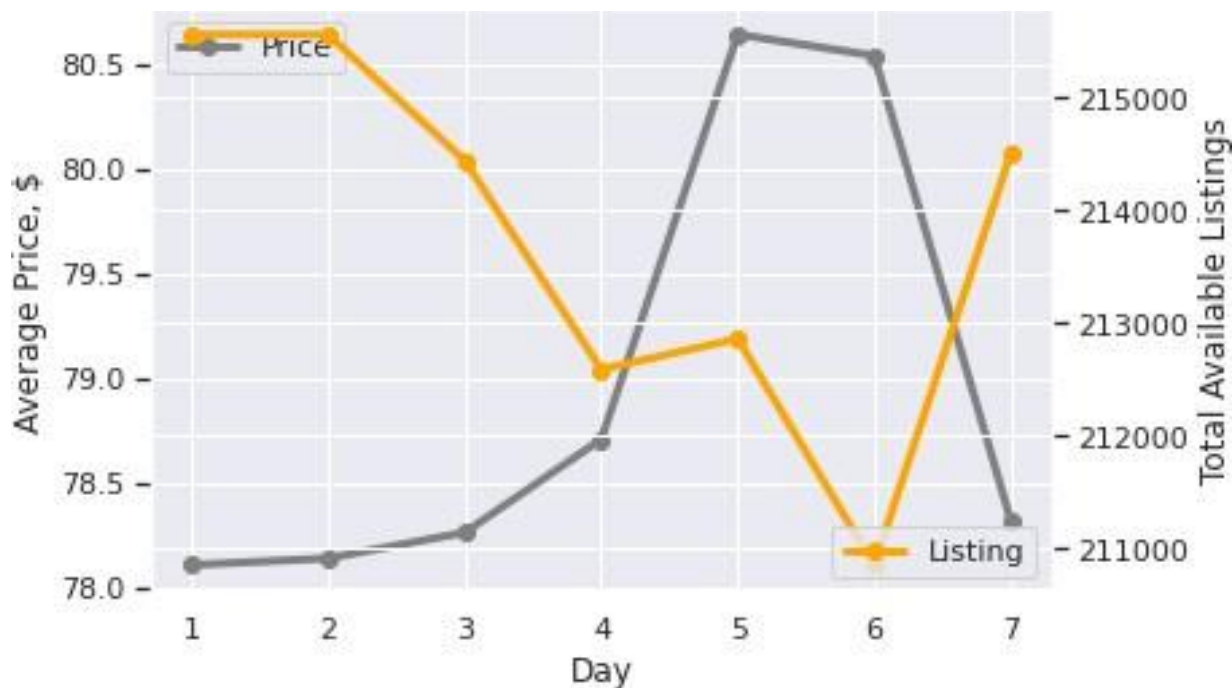


Рис. 3.4. Тижнева зміна ціни та кількість вільних місць

Джерело: створена автором на основі даних [1] в програмному середовищі Python

У Додатку А представлено середню ціну бронювання житла в Берліні на двох осіб за регіонами. Якщо розглянути ціну бронювання за місцезорозташуванням, то найдорожчим є житло в Mitte й коштуватиме 80 доларів США за ніч на двох осіб, що не дивно, адже це центр Берліна. найдешевшим є житло в Marzahn-Hellersdorf, адже воно знаходиться ближче до околиці міста.

Найдорожчими є ціни для готелів та власного будинку, а найдешевшими - для спільної кімнати (рис.3.5). Разом з тим найменший розкид у цінах спостерігається для власного будинку: ціна змінюється в межах 10 доларів США.

А найбільше коливання в цінах спостерігає у готелях; ціни змінюються в розмірі 40 доларів США.

Якщо переглянути дані за кількістю заброньованого житла, то найбільше на Airbnb розміщуються господарями власних будинків (64%), потім йдуть приватні кімнати (32%) та в незначній кількості готелі та спільні кімнати (4%).

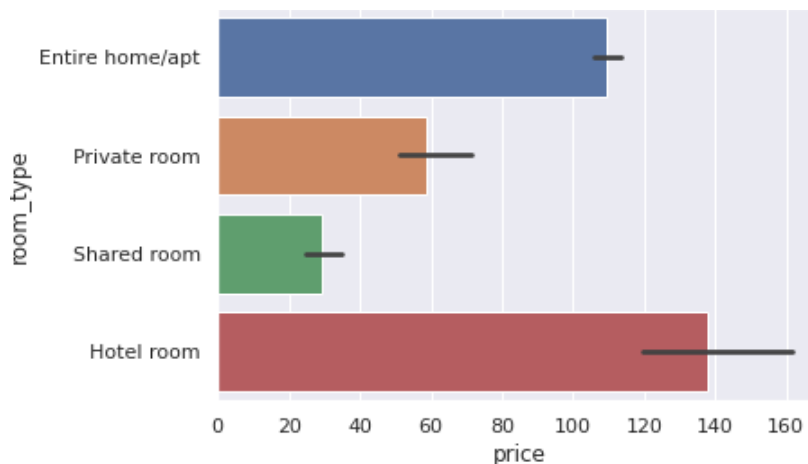


Рис. 3.5. Залежність ціни від типу житла

Джерело: створена автором на основі даних [1] у програмному середовищі Python

Дослідимо, чи є якась залежність між ціною та інформацією про господаря житла. Рис. 3.6 показує, чим швидша відповідь господаря, тим вища ціна за житло, однаковими є ціни при відповіді протягом кількох годин або протягом дня й найдешевшими, якщо господар відповідає лише через декілька днів.

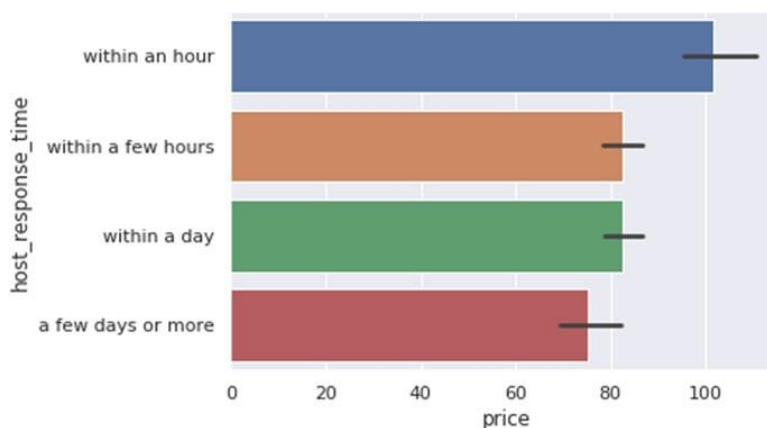


Рис. 3.6. Залежність ціни від часу відповіді

Джерело: створена автором на основі даних [1] у програмному середовищі Python

У господаря, у якого миттєво бронюється житло, ціни вищі, а також у власника житла, у якого є фото власного профайла на Airbnb (рис. 3.7).

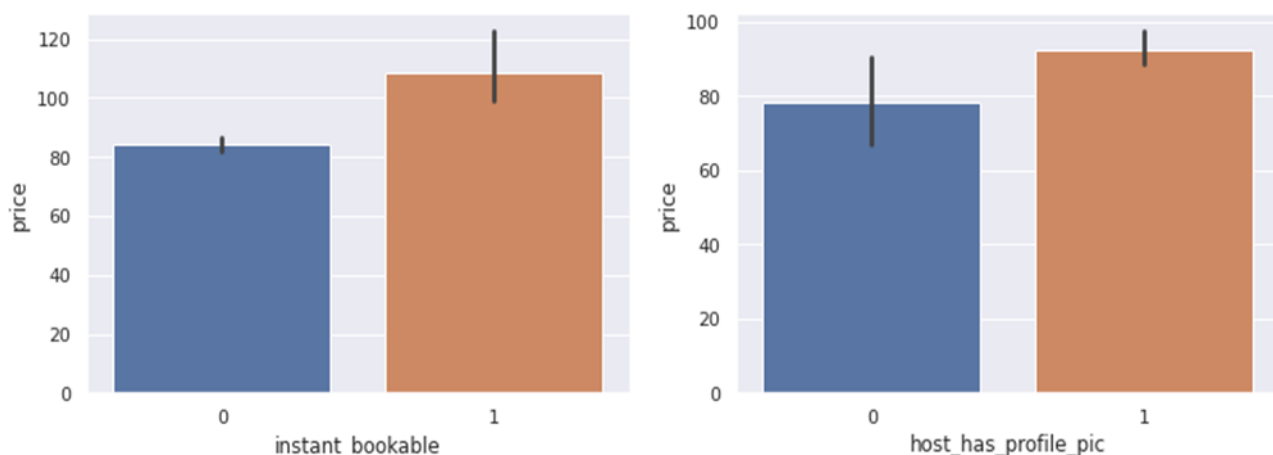


Рис. 3.7. Залежність ціни і миттєвого бронювання та наявності фото господаря. Джерело: створена автором на основі даних [1] у програмному середовищі Python

Аналіз показує, що є залежність між інформацією про господаря та ціною. і це логічно, адже якщо профіль господаря більш відкритий, то йому більше довіряють та швидше забронюють житло в нього. Господар може користуватися великим попитом на житло та встановлювати вищі.

Отже, для моделювання було обрано дані бронювання житла на Airbnb, проведено необхідну підготовку даних та проаналізовано залежності в них, пов'язані з ціною, наприклад, найдешевшими ціни в Берліні є в лютому та якщо бронювати на початку тижня (понеділок-середа). Найдорожчими є ціни в готелях, середніми у приватних кімнатах, та найдешевшими в спільних кімнатах. Було помічено залежність між інформацією про господаря та ціною: чим швидша відповідь господаря, тим ціна вища; наявність фото господаря також підвищує ціну на житло.

### 3.2 Моделювання та інтерпретування результатів

Перші етапи дослідження, а саме, завантаження даних, їх поєднання, проведення статистичного дослідження та обробка їх, були описані в пункті 3.1.

У цьому пункті розглянемо наступні 2 етапи:

3. Побудова та відбір моделей.
4. Аналіз отриманих результатів.

Після обробки даних залишилося 70 колонок з 60 049 рядками даних. Перелік використаних даних та їх опис можна переглянути в Додатку Б. У додатку наведена класифікація досліджуваних факторів у відповідності до трьох груп даних: традиційні – звичайні фактори, які впливають на ціну, такі як місткість житла, кількість кімнат; динамічні – фактори, які залежать від дня тижня, місяця. та персональні – інформація про господаря житла.

Аналіз підходів до моделювання персоналізованого ціноутворення показав, що найрелевантнішими моделями є Ridge, Random Forest, Gradient Boosting, XGBoost, LightGBM, CatBoost регресії.

На основі результатів реалізованих проаналізуємо, які важливі фактори впливу на ціну пропонує кожна з них.

Результати моделювання CatBoost регресії зазначені на рис. 3.8.



Рис. 3.8. Важливі фактори впливу на ціни для CatBoost регресії

Джерело: створена автором у програмному середовищі Python

Результати показують, що топ-5 найважливіших факторів на ціну є кількість спальних кімнат, наявність власної ванної кімнати, місткість житла, кількість заброньованого житла в господаря, розташування. Також на ціну впливають динамічні фактори (кількість вільного житла за 365, 30, 60, 90 днів) і персональні дані про господаря (скільки він уже користується Airbnb, його рейтинг та дата останнього відгуку).

Модель XGBoost регресія в топ-5 найвпливовіших факторів на ціну має лише традиційні дані. У топ-15 факторів усе ж наявні персональні дані господаря: кількість відвіданих гостей, чи належить господар до суперхосту та оцінка чистоти житла. Серед динамічних можна виділити кількість наявних місць за 60 та 365 днів. Результати показують, що вплив інформації про господаря значно менший у порівнянні з результатами в першій моделі.

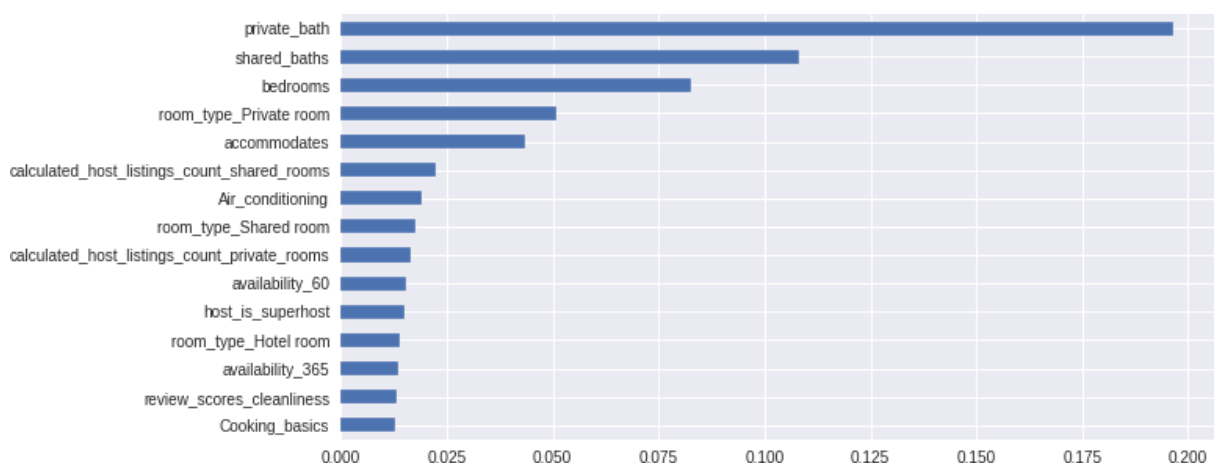


Рис. 3.9. Важливі фактори впливу на ціни для XGBoost регресії

Джерело: створена автором у програмному середовищі Python

У модель Random Forest переважають традиційні фактори, а також в топ-5 потрапляє динамічний фактор: кількість вільних місць при бронюванні за 365 днів, а в топ-15 входять такі персональні дані, як кількість днів, скільки господар зареєстрований на Airbnb; кількість бронювань житла; кількість днів від останнього та першого відгуку та кількість відгуків за тиждень.

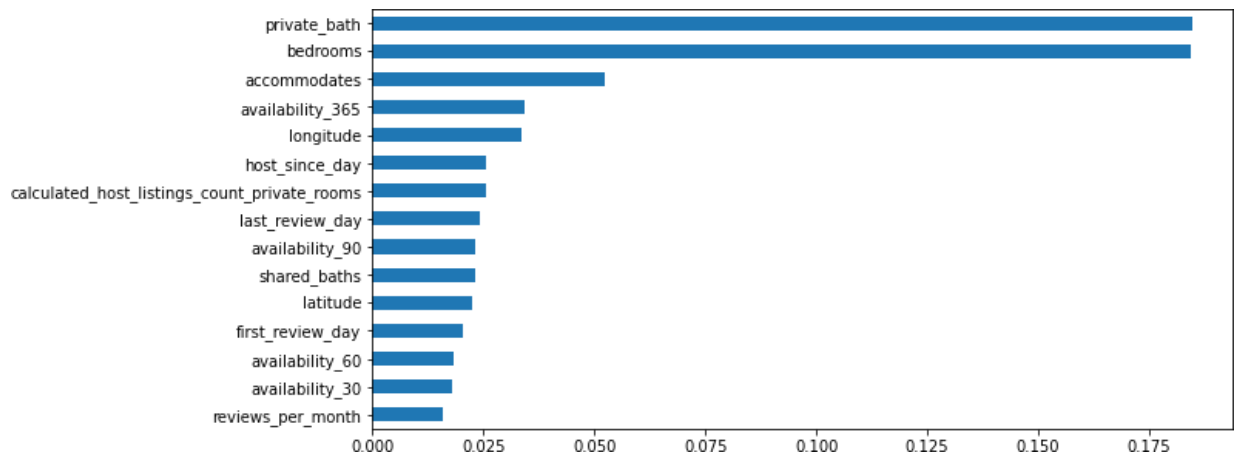


Рис. 3.10. Важливі фактори впливу на ціни для Random Forest

Джерело: створена автором у програмному середовищі Python

Регресія градієнтного бустінгу має схожу картину з попередніми моделями, у топ-5 входять лише традиційні фактори, нижче є як персональні (кількість днів з останнього відгуку, кількість разів заброньоване житло в господаря, рейтинг житла, кількість днів реєстрації господаря, оцінка чистоти житла), так і динамічні фактори.

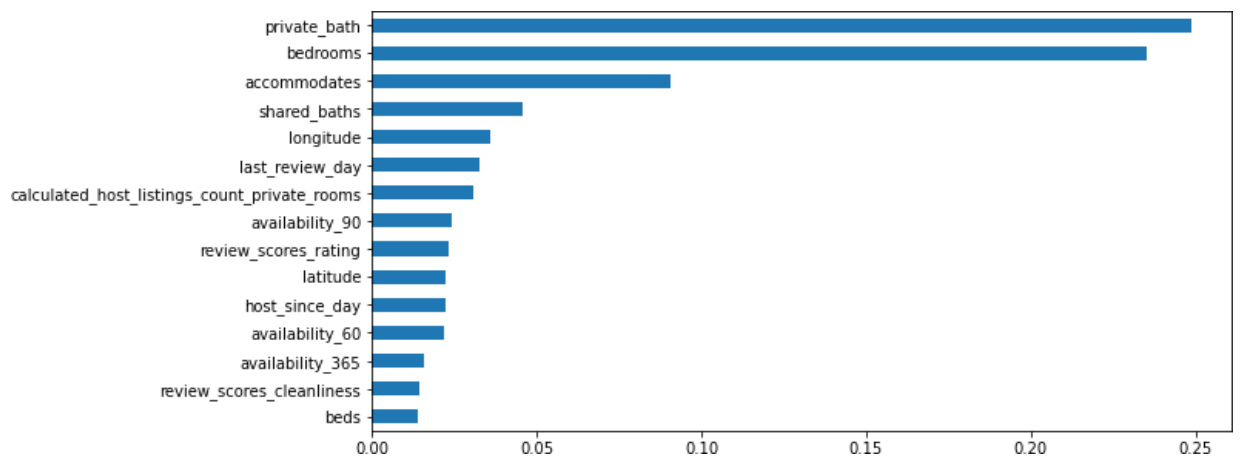


Рис. 3. 11 Важливі фактори впливу на ціни для Gradient Boosting регресії

Джерело: створена автором у програмному середовищі Python

Регресія LightGBM найвпливовішим вважає фактор кількості заброньованого житла в господаря, тобто персональні дані, а нижче лише традиційні фактори. Серед персональних також можна виділити кількість днів

від останнього відгуку, кількість розміщеного житла господарем, кількість відгуків за місяць, загальна оцінка господаря, оцінка чистоти житла та кількість днів від першого відгуку.

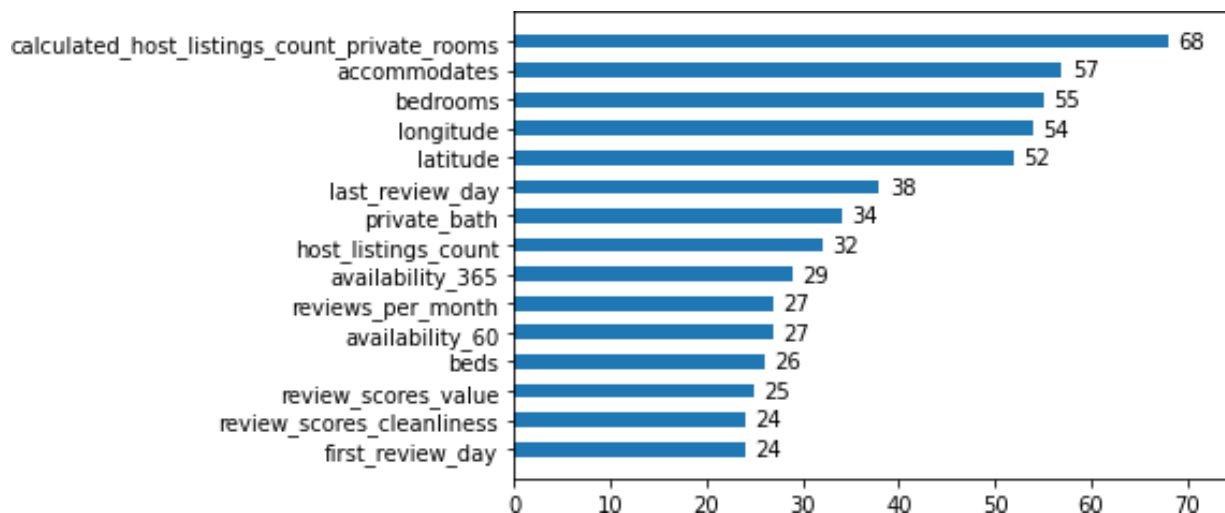


Рис. 3.12. Важливі фактори впливу на ціни для LightGBM регресії

Джерело: створена автором у програмному середовищі Python

Отже, персональні фактори господаря впливають на ціну, і найпопулярнішими є кількість бронювань у господаря, кількість днів від останнього відгуку, рейтинг чистоти житла, кількість відгуків за місяць, загальний рейтинг господаря.

Якщо проаналізувати кількість персональних факторів у перерахованих моделях, які входять в топ-15 найвпливовіших факторів, то отримаємо рейтинг моделей за кількістю персональних даних (Табл. 3.1). Найбільше вплив персональних даних підтверджують моделі CatBoost та LightGBM.

Таблиця 3.1

#### Рейтинг моделей за кількістю персональних даних

Модель	Кількість персональних факторів в:		
	Топ-5	Топ-10	Топ-15
LightGBM Regression	1	4	7
CatBoost Regression	1	1	6
Random Forest Regression	0	3	5
Gradient Boosting Regression	0	3	5
XGBoost Regression	0	2	4

При оцінці якості моделей використовувалося декілька показників: MAE (середня абсолютна похибка) – представляє середнє значення абсолютної різниці між фактичним і прогнозованим значеннями в наборі даних та вимірює середнє значення залишків у наборі даних, MSE (середня квадратична помилка) - середнє квадратичне значення різниці між фактичними і прогнозованим значеннями в наборі даних та вимірює дисперсію залишків, RMSE – це корінь від середньоквадратичних помилок та вимірює дисперсію залишків. та  $R^2$  - вимірює частку дисперсії, пояснену моделлю в загальній дисперсії цільової змінної. Фактично цей показник якості - це нормована середньоквадратична помилка.

Якщо аналізувати якість представлених моделей, то найкраще значення коефіцієнта детермінації в CatBoost Regression ( $R^2=0.624$ ) за середньоквадратичною помилкою - Gradient Boosting Regression (RMSE=42.531).

Результати якості моделей, можна переглянути у Табл. 3.2.

Таблиця 3.2

## Оцінка якості моделей

Модель	MAE	MSE	RMSE	$R^2$
CatBoost Regression	24.883	2064.379	45.435	0.624
XGBoost Regression	26.151	2051.225	45.290	0.616
Random Forest	28.129	2387.875	48.866	0.547
Ridge regression	30.642	2750.637	52.446	0.499
Gradient Boosting Regression	25.963	1808.908	42.531	0.576
LightGBM Regression	27.274	2182.998	46.722	0.584

Отже, було побудовано моделі та визначені найвпливовіші фактори на ціну. Моделі підтвердили гіпотезу про те, що на ціну впливає інформація про господаря (в топ-15 найвпливовіших факторів входять персональні дані господаря, а в моделях CatBoost та LightGBM в топ-5) та доречне дослідження персоналізованого ціноутворення для маркетплейсів загалом та Airbnb зокрема. Найкращими моделями з точки зору коефіцієнта детермінації є CatBoost та XGBoost регресії. Якщо за основу оцінки якості моделі брати RMSE, то це буде Gradient Boosting Regression.

Було показано доречність використання методів градієнтного бустінгу для аналізу впливу персональних факторів на ціну та побудові моделей прогнозу ціни. На основі отриманих результатів можна порадити менеджменту Airbnb покращити їх стратегію ціноутворення з використанням персональних даних господарів та споживачів послуг, за умови наявності згоди на їх використання та повідомлення про застосування стратегії персоналізованого ціноутворення. Також можна використати результати дослідження на Airbnb для інших маркетплейсів.

## ВИСНОВКИ

Результатом кваліфікаційної роботи магістра з економічної кібернетики є систематизація теоретично-методологічних напрацювань щодо ціноутворення в інформаційній економіці, визначення можливостей і перспектив персоналізованого ціноутворення. Запропоновано економіко-математичний інструментарій на основі Data Science для визначення факторів, що мають найбільше значення при формуванні ціни при персоналізованій стратегії. Проведене дослідження дозволяє сформулювати такі висновки та пропозиції:

1. Узагальнено теоретичні підходи щодо ціноутворення в інформаційній економіці та виділено популярні стратегії ціноутворення: «зняття вершків», «проникнення», Freemium, високої-низької ціни, динамічне ціноутворення, конкурентне ціноутворення, нейтральна, пакетна, персоналізоване ціноутворення престижна ціна, проходження за лідером, психологічна, ціноутворення на основі цінності.

2. Персоналізоване ціноутворення - це процес встановлення індивідуальних цін на основі поведінки або характеристики людей з метою досягнення цілей підприємства. Особливостями персоналізованого ціноутворення є вміння розрізняти покупців з метою оцінки їх готовності платити та пропонувати різні ціни індивідуально чи групі осіб; відсутність або не доцільність перепродажу товару або послуги; наявність елементів ринкової влади, щоб мати можливість встановлювати ціни вище за граничні витрати. Серед дискримінації цін з поєднанням персоналізації виділяють такі форми ціноутворення: управління ціною, відволікаючі ціни, крапельне ціноутворення, повторні пропозиції, підроблені спеціальні пропозиції.

3. Існує велика кількість переконливих доказів впровадження персоналізованого ціноутворення в компаніях. Дискримінація споживачів відбувається на основі операційної системи пристрою, з якого заходить клієнт, рівень заряду батареї мобільного пристрою, марці телефону, рівня доходу, історії переглядів, cookies-файлів, частота покупок у конкурентів, вік тощо.

4. Серед популярних економіко-математичних підходів, які застосовуються при персоналізованому ціноутворенні, визначено: моделі нелінійного програмування, математичні моделі, параметричні та напівпараметричні моделі та економетричні моделі. Високої популярності набули алгоритми машинного навчання, як одного з напрямків Data Science. А саме, алгоритми підвищеного градієнту (XGBoost, LightGBM), дерева рішень (Random forest, CART, Personalization Tree, Causal Tree, Student Prescriptive Tree), моделі Байєса, регресії (Linear regression, Ridge regression, Kernel Regression, Logistic regression), нейронні мережі (Deep-Neural Network), моделі кластеризації (SVR, KNN), алгоритми з підкріпленням (Q-Learning).

5. Проаналізовано переваги і недоліки алгоритмів машинного навчання, що можуть застосовуватися при персоналізованому ціноутворенні, а саме Ridge regression, яка створена для роботи з мультиколінеарними даними, що актуально при великому об'ємі схожих даних. Алгоритми, які добре себе показали в сучасних дослідженнях науковців: Random Forest Regression, Gradient Boosting Regression, XGBoost Regression, LightGBM Regression та CatBoost Regression, яка працює з категоріальними даними.

6. Запропоновано використовувати визначені інструменти підтримки персоналізованого ціноутворення на маркетплейсах як специфічних підприємствах електронної комерції. При цьому рекомендовано враховувати персональну інформацію про господарів житла, як продавців, з одного боку, та гостей, як споживачів послуг, з іншого. Такий концептуальний підхід відрізняє дане дослідження від багатьох інших, де розглядається лише використання персональних даних про споживачів при персоналізації цін. Підхід дозволяє сформувати ефективну ціну, що буде раціональною як з точки зору компанії-маркетплейсу, так і з точки зору покупців і продавців.

7. Визначено ключові фактори, що впливають на ціну, при бронюванні житла на платформі «Airbnb»: місткість житла, кількість спальних та ванних кімнат, розташування житла, оцінка господаря житла в рейтингу, наявність житла за 30, 60, 90 та 365 днів, термін, протягом якого господар користується

платформою, термін від першого та останнього відгуку, кількість заброньованого житла.

8. Проаналізовано вплив персональної інформації про господаря на ціноутворення при бронюванні житла на платформі «Airbnb» та визначено, що впливовими чинниками є кількість заброньованого житла у господаря, кількість зареєстрованих днів господаря на платформі, загальна оцінка господаря від користувачів, оцінка чистоти житла, чи належить господар до супергосподаря, кількість днів від першого та останнього відгуку на господаря, кількість відгуків за місяць. Також за допомогою методів машинного навчання показано значимість динамічних (кількість вільних місць на житло за 30, 60, 90 та 365 днів) характеристик при формуванні ціни.

9. Запропоновано та побудовано рейтинг для набору релевантних алгоритмів машинного навчання: LightGBM Regression, CatBoost Regression, Random Forest Regression, Gradient Boosting Regression, XGBoost Regression, що підтримують процеси формування раціональної персональної ціни. Найкраще вплив персональних факторів на ціноутворення продемонструвала модель LightGBM Regression. Найвищу якість моделі з описаних показали CatBoost Regression та Gradient Boosting Regression.

10. На основі аналізу результатів моделювання на даних маркетплейсу «Airbnb» представлено дослідження рекомендує агентам, які займаються питаннями персоналізованого ціноутворення, використовувати запропоновані вище методи для визначення впливовості персональних факторів (даних про продавця), що дозволить розрахувати раціональні персональні ціни на житло, сприятиме удосконаленню стратегії ціноутворення, а значить і покращенню результатів діяльності маркетплейсу. При цьому про використання персональних даних мають бути повідомлені продавці (господарі житла), а про використання стратегії персоналізованого ціноутворення – споживачі (гості). Подібною методикою можуть скористатися будь-які маркетплейси, що прагнуть удосконалити свою стратегію ціноутворення через персоналізацію цін.

11. Розвиток запропонованого підходу вбачається у етичному залученні до дослідження другого компонента персоналізації – даних про покупця.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Inside Airbnb. URL: <http://insideairbnb.com/get-the-data.html>
2. Плєскач В.Л., Затонацька Т.Г. Електронна комерція: підручник / За ред. Л.А.Пономаренко, О.І.Черняка, А.В.Шегда, Київ: Знання, 2007. 535 с.
3. Digital Economy Report 2019. United Nations Conference on Trade and Development – 2019. URL: [https://unctad.org/en/PublicationsLibrary/der2019\\_en.pdf](https://unctad.org/en/PublicationsLibrary/der2019_en.pdf)
4. Statista. URL: <https://www.statista.com/>
5. Trushkina, N. Development of the information economy under the conditions of global economic transformations: features, factors and prospects. *Virtual Economics*. 2019. 2(4). P. 7-25. doi:10.34021/ve.2019.02.04(1)
6. World Bank. URL: <https://www.worldbank.org>
7. Брижань І. А., Ватуля А. О. Сучасні підходи до трактування економічної категорії «ціна». *Економічний форум*. 2015. № 3. С. 6-15. URL: [http://nbuv.gov.ua/UJRN/ecfor\\_2015\\_3\\_3](http://nbuv.gov.ua/UJRN/ecfor_2015_3_3)
8. Bazylevych V., Ihnatiuk V. Metaorder limit prices in evaluating expected market impact and assessing execution service quality. *Investment Management and Financial Innovations*. 2019. 16(2). P. 355-369. doi: 10.21511/imfi.16(2).2019.30
9. Меджибовська Н. С. Електронні системи закупівель: модель поведінки підприємств з урахуванням нецінових показників. *Актуальні проблеми розвитку регіону*. 2017. № 13(1). С. 30-37. URL: [https://www.researchgate.net/publication/322110186\\_ELECTRONIC\\_PROCUREMENT\\_SYSTEMS\\_ENTERPRISES\\_BEHAVIOR\\_MODEL\\_WITH\\_Non-price\\_INDICATORS/fulltext/5a45a7a3a6fdcce1971a635d/ELECTRONIC-PROCUREMENT-SYSTEMS-ENTERPRISES-BEHAVIOR-MODEL-WITH-Non-price-INDICATORS.pdf](https://www.researchgate.net/publication/322110186_ELECTRONIC_PROCUREMENT_SYSTEMS_ENTERPRISES_BEHAVIOR_MODEL_WITH_Non-price_INDICATORS/fulltext/5a45a7a3a6fdcce1971a635d/ELECTRONIC-PROCUREMENT-SYSTEMS-ENTERPRISES-BEHAVIOR-MODEL-WITH-Non-price-INDICATORS.pdf)
10. Чайковська М. П., Медведь Т. С. Модель оптимального ціноутворення в режимі реального часу на основі методів динамічного програмування. *Економічний вісник НТУУ КПІ*. 2016. № 13. doi: 10.20535/2307-5651.13.2016

11. Черноус Г. О., Рибальченко С. А. Оптимізація ціноутворення на основі моделей інтелектуального аналізу даних. *Вісник Київського національного університету. Економіка*. 2015. № 172. С. 52-58. doi:10.17721/1728-2667.2015/172-7/7
12. The Future of Commerce Trend Report 2022. Shopify – 2021. [URL: https://www.shopify.com/research/future-of-commerce/future-of-ecommerce](https://www.shopify.com/research/future-of-commerce/future-of-ecommerce)
13. Namogoo. URL: <https://www.namogoo.com/>
14. Decker A. The Ultimate Guide to Pricing Strategies. URL: <https://blog.hubspot.com/sales/pricing-strategy>
15. Сорокин М. А., Вершинина Ю. С. Поведенческие аспекты ценообразования в электронной торговле. Сборник материалов XII Международной научно-практической конференции (26–27 октября 2017 г.). Томск, 2018. 258 с.
16. Deshpande S. S. Various Pricing Strategies: A Review. *Journal of Business and Management*. 2018. 20(2). P. 75-79. URL: <https://iosrjournals.org/iosr-jbm/papers/Vol20-issue2/Version-8/K2002087579.pdf>
17. Lockhart J. W. Ecommerce Pricing Strategies: How to Scale and Grow Without Losing Profit. *Shopify*. 2022. URL: <https://www.shopify.com/enterprise/ecommerce-pricing-strategy>
18. Wells T. Leadership Guide to Machine Learning Pricing Algorithms. *Taylorwells*. 2020. URL: <https://taylorwells.com.au/machine-learning-pricing-algorithms/>
19. The role of search and social in the customer life cycle. Forrester – 2016. URL: <https://www.catalystdigital.com/wp-content/uploads/WhySearchPlusSocialEqualsSuccess-Catalyst.pdf>
20. Ярмоленко Ю. А., Черноус Г. О. Співвідношення між різновидами ціноутворення в інформаційній економіці В кн.: Инструментальные средства моделирования систем в информационной экономике / Под ред. В.С. Пономаренко, Т. С. Клебановой, Харьков, ВШЭМ – ХНЭУ им. С. Кузнеця, 2019. С. 120-135., 452 с.

21. Pigou, A.C. The economics of welfare. *The Economic Journal*. 1921. 31(122). P. 209-214. doi: 10.2307/2222816
22. Personalised Pricing - Increasing Transparency to Improve Trust. Office of Fair Trading – 2013. URL: [https://webarchive.nationalarchives.gov.uk/ukgwa/20140402165101/http://oft.gov.uk/shared\\_of/market-work/personalised-pricing/oft1489.pdf](https://webarchive.nationalarchives.gov.uk/ukgwa/20140402165101/http://oft.gov.uk/shared_of/market-work/personalised-pricing/oft1489.pdf)
23. Personalised Pricing in the Digital Era. Organisation for Economic Co-operation and Development – 2018. URL: [https://one.oecd.org/document/DAF/COMP\(2018\)13/en/pdf](https://one.oecd.org/document/DAF/COMP(2018)13/en/pdf)
24. Poort J., Zuiderveen Borgesius F. J. Does everyone have a price? Understanding people's attitude towards online and offline price discrimination. *Internet Policy Review*. 2019. 8(1). P. 1-20. doi: 10.14763/2019.1.1383
25. Filistrucchi, L., Geradin, D., Van Damme, E. Identifying two-sided markets. *SSRN Electronic Journal*. 2013. 36. P. 33–59. doi: 10.2139/ssrn.2008661
26. Ezrachi, A., Stucke M. The Rise of Behavioural Discrimination. *European Competition Law Review*. 2016. 37/2. P. 485-492. URL: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2830206](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2830206)
27. Freed J. AirAsia Testing Personalised Baggage Pricing, Eyes More Add-On Revenues. *Reuters*. 2017. URL: <https://www.reuters.com/article/airasia-strategy/airasia-testing-personalised-baggage-pricingeyes-more-add-on-revenues-idUSL3N1NM2DS>
28. OECD. Personalised Pricing in the Digital Era. Background note by the secretariat. 2018. DAF/COMP(2018)13'. URL: [https://one.oecd.org/document/DAF/COMP\(2018\)13/en/pdf](https://one.oecd.org/document/DAF/COMP(2018)13/en/pdf)
29. Mattioli D. On Orbitz, Mac Users Steered to Pricier Hotels. *Wall Street Journal*. 2012. URL: <https://www.wsj.com/articles/SB10001424052702304458604577488822667325882>
30. Valentino-DeVries, J., Singer-Vine J., Soltani A. Websites Vary Prices, Deals Based on Users' Information. *Wall Street Journal*. 2012. URL:

<https://www.wsj.com/articles/SB1000142412788732377720457818939181388153>

[4](#)

31. Zuiderveen Borgesius F., Poort J. Online Price Discrimination and EU Data Privacy Law. *Consum Policy*. 2017. 40. P. 347–366. doi: 10.1007/s10603-017-9354-z
32. Joseph P. Bailey. *Intermediation and Electronic Markets: Aggregation and Pricing at the Internet*. 1998. 180 p.
33. Clifford, S. Shopper Alert: Price May Drop for You Alone. *The New York Times*. 2012. URL: <https://www.nytimes.com/2012/08/10/business/supermarkets-try-customizing-prices-forshoppers.html>.
34. Newcomer, E. Uber Starts Charging What It Thinks You're Willing to Pay. *Bloomberg*. 2017. URL: <https://www.bloomberg.com/news/articles/2017-05-19/uber-s-future-may-rely-on-predictinghow-much-you-re-willing-to-pay>.
35. Mahdawi, A. Is Your Friend Getting a Cheaper Uber Fare Than You Are? *The Guardian*. 2018. URL: <https://www.theguardian.com/commentisfree/2018/apr/13/uber-lyft-prices-personalized-data>
36. Consumers International, the Mozilla Foundation. A consumer investigation into personalized pricing. 2022. URL: [https://www.consumersinternational.org/media/369078/personalised\\_pricing\\_15\\_02\\_2022.pdf](https://www.consumersinternational.org/media/369078/personalised_pricing_15_02_2022.pdf)
37. Hannak, A. Measuring Price Discrimination and Steering on E-commerce Web Sites. IMC '14: Proceedings of the 2014 Conference on Internet Measurement Conference (05 November 2014). New York. P. 305–318. doi:10.1145/2663716.2663744
38. Chulkov, D., Nizovtsev D. Exploring Price Discrimination in an e-Commerce Environment. *Journal of the International Academy for Case Studies*. 2016. 22(3). P. 157-164. URL: [https://www.researchgate.net/publication/320212710\\_Exploring\\_Price\\_Discrimination\\_in\\_an\\_e-Commerce\\_Environment](https://www.researchgate.net/publication/320212710_Exploring_Price_Discrimination_in_an_e-Commerce_Environment)

39. Valentino-Devries, J., Singer-Vine, J., Soltani, A. (2012, December 23). Websites vary prices, deals based on users' information. *Wall Street Journal*. 2012. URL:  
<http://online.wsj.com/article/SB10001424127887323777204578189391813881534.html>
40. Council of Economic Advisers. Big Data and Differential Pricing, *The White House*. 2015. URL:  
[https://obamawhitehouse.archives.gov/sites/default/files/whitehouse\\_files/docs/Big\\_Data\\_Report\\_Nonembargo\\_v2.pdf](https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/docs/Big_Data_Report_Nonembargo_v2.pdf)
41. Hogan, K. Consumer Experience in the Retail Renaissance: How Leading Brands Build a Bedrock with Data. *Deloitte Digital*. 2018. URL:  
<https://www.deloittedigital.com/us/en/blog-list/2018/consumerexperience-in-the-retail-renaissance--how-leading-brand.html>
42. Poort J., Frederik J. Zuiderveen Borgesius. Does everyone have a price? Understanding people's attitude towards online and offline price discrimination. *Internet Policy Rev*. 2019. doi:10.14763/2019.1.1383
43. Ennis S. F, Lam W. Personalised Pricing and Disclosure. *BEIS Research Paper Number*. 2021. 8. 28 pages. URL:  
<https://www.gov.uk/government/publications/personalised-pricing-and-disclosure>
44. Personalised pricing for communications: making data work for consumers. *Ofcom*. 2020. URL:  
[https://www.ofcom.org.uk/data/assets/pdf\\_file/0033/199248/personalised-pricing-discussion.pdf](https://www.ofcom.org.uk/data/assets/pdf_file/0033/199248/personalised-pricing-discussion.pdf)
45. Яромоленко Ю. А., Черноус Г. О. Моделивання транзакційного ціноутворення в інформаційній економіці. Київський національний університет ім. Тараса Шевченка. 2020. 217 с.
46. Narahari Y., Raju C.V.L., Ravikumar K. Dynamic pricing models for electronic business. *Sadhana*. 2005. 30. P. 231–256. doi:10.1007/BF02706246
47. Acquisti A., Varian H. R. Conditioning Prices on Purchase History. *Marketing Science*. 2005. 24(3). P. 367-381. doi:10.1287/mksc.1040.0103

48. Phillips R. Customized Pricing. Columbia University and Nomis Solutions. 2010. 34 pages.
49. Niu C., Zheng Z., Wu F., Tang S. Online Pricing with Reserve Price Constraint for Personal Data Markets. *IEEE Transactions on Knowledge and Data Engineering*. 2022. 34(4). P. 1928-1943. doi: 10.1109/TKDE.2020.3000262
50. Allender W. J., Liaukonyte J., Nasser S., Richards T. J. Price Fairness and Strategic Obfuscation. *Marketing Science*. 2021. 40(1). P. 122-146. doi: 10.1287/mksc.2020.1244
51. Kallus N., Zhou A. Fairness, Welfare, and Equity in Personalized Pricing. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency Association for Computing Machinery. P. 296–314. doi: 10.1145/3442188.3445895
52. 14 Most Used Data Science Tools for 2022 - Essential Data Science Ingredients. URL: <https://data-flair.training/blogs/data-science-tools/>
53. Fumo D. Types of Machine Learning Algorithms. *Towards Data Science*. URL: <https://towardsdatascience.com/types-of-machine-learning-algorithms-you-should-know-953a08248861>
54. Commonly used Machine Learning Algorithms. Analytics Vidhya. URL: <https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/>
55. Brownlee J. Machine Learning Mastery with Python: Understand Your Data, Create Accurate Models and Work Projects End-to-end. 2016. 170 pages.
56. Hajjar A. J. Dynamic Pricing Algorithms: Top 3 Models in 2022. AIMultiple. 2022. URL: <https://research.aimultiple.com/dynamic-pricing-algorithm/>
57. Bauer J., Jannach D. Optimal Pricing in E-Commerce Based on Sparse and Noisy Data. *Decision Support Systems*. 2017, doi:10.1016/j.dss.2017.12.00
58. Kallus, N. Recursive Partitioning for Personalization using Observational Data. *ICML*. 2017.
59. Zhou Zh. Athey, Wager, S. (2018). Offline Multi-Action Policy Learning: Generalization and Optimization. 2018. URL: <https://arxiv.org/pdf/1810.04778.pdf>

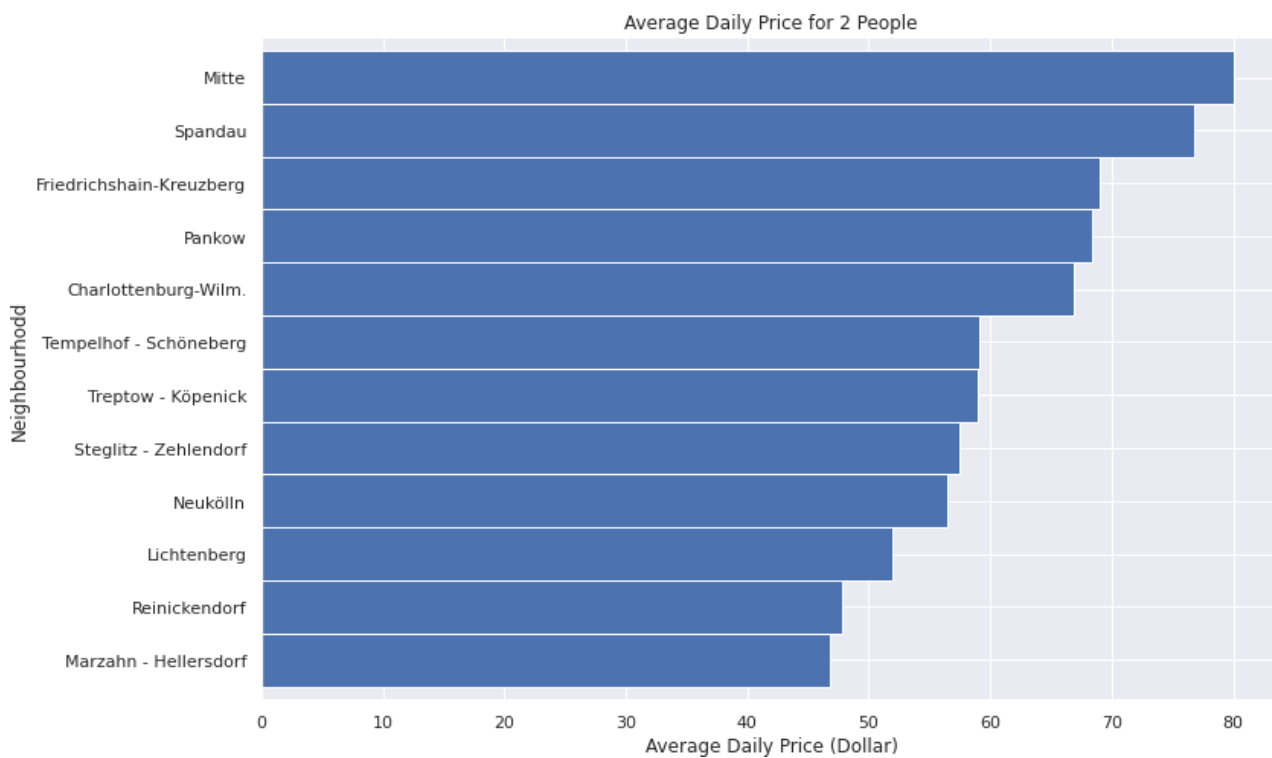
60. Maestre R., Duque J., Rubio A., Arevalo J. Reinforcement Learning for Fair Dynamic Pricing. 2018
61. Bertsimas D., Dunn J., Mundru N.(2019) Optimal Prescriptive Trees. *INFORMS Journal on Optimization*. 2019. 1(2). P. 164-183. doi: 10.1287/ijoo.2018.0005
62. Shukla N., Kolbeinsson A., Otwell K., Marla L., Yellepeddi K. 2019. Dynamic Pricing for Airline Ancillaries with Customer Context. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19). 2019. Association for Computing Machinery. New York. P. 2174–2182. doi: 10.1145/3292500.3330746
63. Mišić V. V. Optimization of Tree Ensembles. *Operations Research*. 2020. 68(5). P.1605-1624. doi: 10.1287/opre.2019.1928
64. Satsangi Y., Schwarzova L. Predicting Airbnb Prices with Neighborhood Characteristics: Machine Learning Approach. Tilburg University. 2020. 43 pages.
65. Awan M. J., Rahim M. S. M., Nobanee H., Yasin A., Khalaf. O. I. A Big Data Approach to Black Friday Sales. *Intelligent Automation and Soft Computing*. 2021. 27. P. 785–797. doi: 10.32604/iasc.2021.014216.
66. Elmachtoub A., Jason L., Ryan M. Decision Trees for Decision-Making under the Predict-then-Optimize Framework. 2020
67. Biggs M., Sun W., Ettl M. (2020). Model Distillation for Revenue Optimization: Interpretable Personalized Pricing. 2020. URL: [https://www.researchgate.net/publication/342733678\\_Model\\_Distillation\\_for\\_Revenue\\_Optimization\\_Interpretable\\_Personalized\\_Pricing](https://www.researchgate.net/publication/342733678_Model_Distillation_for_Revenue_Optimization_Interpretable_Personalized_Pricing)
68. Islam M. D., Li B. Saiful I. K., Ahasan R., Mia, M. R., Haque M. (2021). Airbnb rental price modeling based on Latent Dirichlet Allocation and MESF-XGBoost composite model. *Machine Learning with Applications*. 2021. 7. doi: 10.1016/j.mlwa.2021.100208.
69. Ahuja A. & Lahiri A., Das A. Predicting Airbnb Rental Prices Using Multiple Feature Modalities. 2021. URL: [https://www.researchgate.net/publication/357014036\\_Predicting\\_Airbnb\\_Rental\\_Prices\\_Using\\_Multiple\\_Feature\\_Modalities](https://www.researchgate.net/publication/357014036_Predicting_Airbnb_Rental_Prices_Using_Multiple_Feature_Modalities)

70. McNamara C. Xgboost vs Catboost vs Lightgbm: which is best for price prediction? 2018. URL: <https://blog.griddynamics.com/xgboost-vs-catboost-vs-lightgbm-which-is-best-for-price-prediction/>
71. Xin S., Khalid K. Modelling House Price Using Ridge Regression and Lasso Regression. *International Journal of Engineering & Technology*. 2018. 7. P. 498. doi: 10.14419/ijet.v7i4.30.22378.
72. Kargin K. Ridge Regression Fundamentals. URL: <https://keremkargin.medium.com/ridge-regression-fundamentals-and-modeling-in-python-bb56f4301f62>
73. Mwitid D. Random Forest Regression: When Does It Fail and Why? 2021. URL: <https://neptune.ai/blog/random-forest-regression-when-does-it-fail-and-why>
74. Yadaw S. Predicting Housing Prices Using Scikit-Learn's Random Forest Model. Towards Data Science. 2020. URL: <https://towardsdatascience.com/predicting-housing-prices-using-a-scikit-learns-random-forest-model-e736b59d56c5>
75. Piryonesi, S. M., El-Diraby T. E. Data Analytics in Asset Management: Cost-Effective Prediction of the Pavement Condition Index. *Journal of Infrastructure Systems*. 2020
76. Chen T., Guestrin C. XGBoost: A Scalable Tree Boosting System. 2016
77. LightGBM. URL: <https://lightgbm.readthedocs.io/en/latest/index.html>
78. LightGBM (Light Gradient Boosting Machine). URL: <https://www.geeksforgeeks.org/lightgbm-light-gradient-boosting-machine/>
79. Thiesen S. CatBoost regression. 2021. URL: <https://towardsdatascience.com/catboost-regression-in-6-minutes-3487f3e5b329>
80. John B. When to Choose CatBoost Over XGBoost or LightGBM. Neptune. 2022. URL: <https://neptune.ai/blog/when-to-choose-catboost-over-xgboost-or-lightgbm>

## ДОДАТКИ

## Додаток А

## Середня ціна бронювання житла по регіонам м. Берліна



## Типи даних

Дані	Класифікація	Тип даних	Опис
listing_id	Традиційні	Числовий	Код бронювання
host_since_day	Персональні	Числовий	Кількість днів з дати реєстрації господаря
host_response_rate	Персональні	Числовий	Швидкість відповіді господаря
host_acceptance_rate	Персональні	Числовий	Швидкість прийняття господарем запит на бронювання
host_is_superhost	Персональні	Логічний	Чи є господар суперхостом
host_listings_count	Персональні	Числовий	Кількість записів, які має хост (за розрахунками Airbnb)
host_has_profile_pic	Персональні	Логічний	Наявність фота господаря
host_identity_verified	Персональні	Логічний	Наявність підтвердження особистості господаря
Latitude	Традиційні	Числовий	Розміщення житла: широта
Longitude	Традиційні	Числовий	Розміщення житла: довжина
Accommodates	Традиційні	Числовий	Місткість житла
Bedrooms	Традиційні	Числовий	Кількість спальних кімнат
Beds	Традиційні	Числовий	Кількість ліжок
has_availability	Традиційні	Логічний	Вільне чи заброньоване житло на конкретну дату
availability_30	Динамічні	Числовий	На скільки днів з 30 днів заброньоване житло на конкретну дату
availability_60	Динамічні	Числовий	На скільки днів з 60 днів заброньоване житло на конкретну дату
availability_90	Динамічні	Числовий	На скільки днів з 90 днів заброньоване житло на конкретну дату
availability_365	Динамічні	Числовий	На скільки днів з 365 днів заброньоване житло на конкретну дату
number_of_reviews	Традиційні	Числовий	Кількість відгуків
first_review_day	Персональні	Числовий	Кількість днів з першого відгуку господареві
last_review_day	Персональні	Числовий	Кількість днів з останнього відгуку господареві
review_scores_rating	Персональні	Числовий	Загальна цінка мандрівника
review_scores_accuracy	Персональні	Числовий	Оцінка мандрівника відносно точності розташування житла
review_scores_cleanliness	Персональні	Числовий	Оцінка мандрівника відносно чистоти житла
review_scores_checkin	Персональні	Числовий	Оцінка мандрівника відносно заселення в житло
review_scores_communication	Персональні	Числовий	Оцінка мандрівника відносно комунікації з господарем
review_scores_location	Традиційні	Числовий	Оцінка мандрівника відносно розташування житла
instant_bookable	Персональні	Логічний	Можливість гостя автоматично забронювати оголошення, не вимагаючи від господаря прийняти його запит на бронювання.
calculated_host_listings_*	Персональні	Числовий	Кількість запитів, які має хост у поточному записі.
reviews_per_month	Персональні	Числовий	Кількість відгуків за місяць
shared_baths	Традиційні	Логічний	Кількість спільних ванних кімнат
private_bath	Традиційні	Логічний	Кількість окремих ванних кімнат
Email	Персональні	Логічний	Верифікація господаря через e-mail
Phone	Персональні	Логічний	Верифікація господаря через телефон
Facebook	Персональні	Логічний	Верифікація господаря через facebook
government_id	Персональні	Логічний	Верифікація господаря через паспорт

## Продовження до Додатку Б

Дані	Класифікація	Тип даних	Опис
Heating	Традиційні	Логічний	Наявність опалення
TV	Традиційні	Логічний	Наявність телевізора
AirCondition	Традиційні	Логічний	Наявність кондиціонера
Microwave	Традиційні	Логічний	Наявність мікрохвильовки
Cooking_basics	Традиційні	Логічний	Наявність базових речей для приготування їжі
Hair_dryer	Традиційні	Логічний	Наявність фена
Refrigerator	Традиційні	Логічний	Наявність холодильника
Iron	Традиційні	Логічний	Наявність праски
Wifi	Традиційні	Логічний	Наявність Wi-fi
Hot_water	Традиційні	Логічний	Наявність гарячої води
Kitchen	Традиційні	Логічний	Наявність кухні
Breakfast	Традиційні	Логічний	Включення сніданку
month	Динамічні	Логічний	Місяць бронювання житла
Day	Динамічні	Логічний	День тижня бронювання житла
day_num	Динамічні	Логічний	День місяці бронювання житла
neighbourhood_group*	Традиційні	Логічний	Район місцезнаходження житла
room_type_*	Традиційні	Логічний	Тип кімнати
host_response_time_*	Персональні	Логічний	Швидкість відповіді господаря

**Київський національний університет імені Тараса Шевченка**  
**Економічний факультет**  
**Кафедра економічної кібернетики**

**ЗАВДАННЯ**  
**на кваліфікаційну роботу магістра**  
студентки 2 курсу магістратури спеціальності 051 «Економіка»,  
ОНП «Економічна кібернетика»  
**Горбунової Ярослави Миколаївни**

1. Тема роботи: «Моделювання персоналізованого ціноутворення на основі підходів Data Science»
2. Термін завершення роботи: 9 травня 2022 року.
3. Об'єкт дослідження: персоналізоване ціноутворення в інформаційній економіці.
4. Предмет дослідження: економіко-математичні методи та моделі для підтримки персоналізованого ціноутворення.
5. Мета дослідження: систематизація теоретично-методологічні напрацювання щодо ціноутворення в інформаційній економіці, визначення можливостей й перспективи персоналізованого ціноутворення, обґрунтування економіко-математичного інструментарію на основі Data Science для визначення факторів, що мають найбільше значення при формуванні ціни при персоналізованому ціноутворенні.
6. Методи дослідження: методи наукової абстракції, порівнянь, індукції та дедукції, аналізу та синтезу, економіко-математичні методи та ІТ.
7. Завдання дослідження:
  - 7.1 узагальнити теоретичні підходи щодо ціноутворення;
  - 7.2 визначити сутність та особливості персоналізованого ціноутворення;
  - 7.3 дослідити практику впровадження персоналізованого ціноутворення;
  - 7.4 вивчити економіко-математичні підходи, моделі та інструменти Data Science, які застосовуються при персоналізованому ціноутворенні;
  - 7.5 проаналізувати вплив персональної інформації про господаря на ціноутворення при бронюванні житла на платформі «Airbnb»;
  - 7.6 показати значимість динамічних та персональних характеристик при формуванні ціни;
  - 7.7 запропонувати набір релевантних алгоритмів машинного навчання для дослідження факторів впливу на ціну, що підтримують формування раціональної персональної ціни.

Науковий керівник: д.е.н., професор Черноус Галина Олександрівна  
Студент: Горбунова Ярослава Миколаївна

Затверджено на засіданні кафедри економічної кібернетики  
Протокол № 3 від 12 жовтня 2021 року

## Календарний план виконання кваліфікаційної роботи магістра

№	Етапи бакалаврської роботи	Терміни виконання	Відмітка керівника про виконання
1	Вибір теми роботи магістра	01.10.2021 - 07.10.2021	
2	Розробка та затвердження завдання на роботу магістра	08.10.2021-12.10.2021	
3	Опрацювання зарубіжних і українських джерел з вибраної тематики	13.10.2021-15.12.2021	
4	Дослідження теоретично-методологічні напрацювання щодо ціноутворення в інформаційній економіці. Написання Розділу 1	16.12.2021-30.01.2022	
5	Опрацювання методів для моделювання персоналізованого ціноутворення Написання Розділу 2	31.01.2021-20.02.2021	
6	Формування бази даних для дослідження. Обробка та аналіз даних	21.02.2022-20.03.2022	
7	Побудова та відбір моделей для визначення найвпливовіших факторів для персоналізованого ціноутворення	20.03.2022-20.04.2022	
8	Опрацювання отриманих результатів. Написання Розділу 3	19.04.2022-08.05.2022	
9	Подання роботи до захисту	15.05.2022	

**Науковий керівник:** д. е. н., професор. Черноус Г. О.

**Студент:** Горбунова Я. М.