

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХЕРСОНСЬКИЙ ДЕРЖАВНИЙ УНІВЕРСИТЕТ
Факультет комп'ютерних наук, фізики та математики
Кафедра комп'ютерних наук та програмної інженерії

Порівняльне оцінювання методів машинного навчання
для прогнозування цін на нерухомості

Кваліфікаційна робота (проект)
на здобуття ступеня вищої освіти «магістр»

Виконав: студент 2 курсу 261М групи
Спеціальності
126 «Інформаційні системи
та технології»
(шифр, назва)
Освітньо-професійної програми:
«Інформаційні системи та технології»
(назва)
Федченко Костянтин Олегович
Керівник: доктор економічних наук,
професор Кобець В.М.
Рецензент: Іванов О.Ю.
Senior Developer
ІТ компанія TurnKey Labs

ЗМІСТ

ВСТУП	3
РОЗДІЛ 1	
МОДЕЛЬ ВЕБ-СЕРВІСУ ПРОГНОЗУВАННЯ ЦІН, ЇЇ ХАРАКТЕРИСТИКИ ТА РОЛЬ	5
1.1. Основні положення	5
1.2. Огляд існуючих веб-сервіс з нерухомості	5
1.3 Функціональні вимоги до сервісу оренди житла	8
РОЗДІЛ 2	
ЦИФРОВА ТРАНСФОРМАЦІЯ ПРОГНОЗУВАННЯ ЦІН НА НЕРУХОМІСТЬ З ВИКОРИСТАННЯМ МАШИННОГО НАВЧАННЯ	
	11
2.1. Традиційні методи оцінки нерухомості	11
2.2. Просторова економетрика	13
2.3. Фактори впливу на ціни нерухомості	15
2.4. Машинне навчання у прогнозуванні цін на нерухомість	18
2.5. Змішана оцінка та індекси цін на житло	21
2.6. Моделі оцінки для окремих ринків	24
2.7. Актуальні тенденції в аналізі цін на нерухомість	27
РОЗДІЛ 3	
РОЗРОБКА ВЕБ-СЕРВІСУ ПРОГНОЗУВАННЯ ЦІН НА НЕРУХОМІСТЬ	31
3.1. Загальна функціональність веб-сервісу	31
3.2 Функціональні модулі системи веб-сервісу для прогнозування цін на нерухомість	32
3.3 Експериментальна частина	37
ВИСНОВОК	42
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	43

ВСТУП

Актуальність дослідження. На сьогодні ринок оренди нерухомості стрімко розвивається, і дедалі більше користувачів звертаються до онлайн-сервісів для пошуку та оцінки квартир. Проте більшість існуючих платформ зосереджуються лише на пропозиціях та не мають інструментів для прогнозування цін. Така ситуація ускладнює орендодавцям і орендарям прийняття зважених рішень, адже вони не мають можливості передбачити, як зміниться вартість оренди з часом, або на яку ціну очікувати при певних їх очікуваннях яка може задовольнити опираючись на їх можливості. В даний час оголошення про квартиру включають інформацію про квартиру, поверх, кількість кімнат, наявність тварин й тощо. Але немає прогнозування ціни на основі критеріїв пошуку, що може призвести до необґрунтованих пошуків, й можуть до того, що буде отримуватися суттєво відмінні результати, на відміну від пошукових.

Точне прогнозування цін на нерухомість з урахуванням характеристик квартир, таких як місцезнаходження, площа, стан будівлі та інші параметри, допоможе користувачам приймати обґрунтовані рішення і мінімізувати фінансові ризики.

Таким чином, на сьогодні актуальним є створення веб-сервісу, який забезпечить користувачам можливість отримувати точні прогнози цін на оренду нерухомості, з урахуванням специфічних характеристик кожного об'єкта.

Об'єкт дослідження — методи машинного навчання для прогнозування цін на нерухомість.

Предмет дослідження — прогнозування цін на нерухомість з використанням моделей машинного навчання.

Мета дослідження — розробка архітектури веб-сервісу для прогнозування цін на нерухомість, що враховує різні характеристики квартир та надає точні прогнози цін.

Завдання:

1. Провести огляд літератури щодо прогнозування цін на нерухомість за допомогою машинного навчання.
2. Дослідити та описати методи машинного навчання для прогнозування цін на нерухомість
3. Дослідити та описати основні показники й системи прогнозування.
4. Побудувати веб-сервіс для прогнозування цін на нерухомість.
5. Провести експериментальну оцінку моделі прогнозування та проаналізувати її точність.

Методи дослідження. Алгоритми та методи машинного навчання для прогнозування цін на нерухомість.

Апробація.

Апробацію роботи було проведено у 2024 році, за результатами якої підготовлено статтю на Міжнародну конференцію "Information Technology and Implementation" (IT&I-2024) Київського національного університет імені Тараса Шевченка і до альманаху «Магістерські студії» ХДУ.

Структура роботи. Дослідження складається зі вступу, трьох розділів, висновків та списку використаних джерел. У першому розділі аналізуються моделі веб-сервісів та основні характеристики. У другому розглядаються методи машинного навчання в прогнозуванні цін на нерухомість. Третій розділ присвячений розробці та тестуванню веб-сервісу для прогнозування цін на нерухомість.

РОЗДІЛ 1

МОДЕЛЬ ВЕБ-СЕРВІСУ ПРОГНОЗУВАННЯ ЦІН, ЇЇ ХАРАКТЕРИСТИКИ ТА РОЛЬ

1.1. Основні положення

З розвитком технологій, прискоренням міграційних процесів і зростанням міст виникає потреба у точному прогнозуванні цін на нерухомість. Багато сервісів для оренди нерухомості не пропонують можливості прогнозування цін, що є важливим для орендодавців та орендарів, які прагнуть мінімізувати ризики та оптимізувати свої обмежені ресурси. Тому наш проєкт спрямований на створення веб-сервісу, який прогнозує ціни на нерухомість, базуючись на аналізі різноманітних характеристик квартир.

Сьогодні існує багато веб-сервісів для оренди нерухомості, але жоден із них не пропонує прогнозування цін на нерухомість. Це є суттєвим недоліком, оскільки прогнозування цін на нерухомість дозволяє користувачам приймати більш обґрунтовані рішення. Наш веб-сервіс надає рішення цієї проблеми, пропонуючи точні прогнози цін на основі аналізу таких факторів, як кількість кімнат, поверх, загальна площа, дозвіл на домашніх тварин, район, матеріал стін і близькість до метро.

1.2. Огляд існуючих веб-сервісів з нерухомості

У процесі формування остаточних ідей та завдань для майбутнього веб-ресурсу було вирішено дослідити кілька ресурсів, які надають послуги оренди нерухомості. Перед вивченням цих ресурсів ми визначили 12 критеріїв, які характеризують якість і рівень послуг, пов'язаних з орендою нерухомості, на відповідних веб-сервісах в Україні (Таблиця 1.1).

Таблиця 1.1

Огляд існуючих веб-сервісів оренди нерухомості

Критерій	lun.ua	dom.ria.com	flatfy.ua	rieltor.ua	olx.ua	24realty.ua	100realty.ua
----------	------------------------------------	--	--	--	------------------------------------	--	--

1. Можливість прогнозування цін	ні	ні	ні	ні	ні	ні	ні
2. Можливість публікації оголошень	так	так	так	так	так	так	так
3. Фільтрація інформації за параметрами	так	так	так	так	так	так	так
4. Розділ для соціального житла	ні	ні	ні	ні	ні	ні	ні
5. Рейтинг орендодавців	ні	так	ні	ні	ні	ні	ні
6. Перевірка нерухомості	так	так	так	так	ні	ні	ні
7. Перевірка орендодавців	так	так	так	так	ні	ні	ні
8. Відгуки користувачів	так	так	так	так	так	ні	ні
9. Інтеграція з картами	так	так	так	так	так	так	ні
10. Підтримка мобільного додатку	так	так	так	так	так	ні	ні
11. Можливість вибору пріоритетних оголошень	так	так	ні	ні	так	ні	ні
12. Вартість послуг	ПП	ПП	безкоштовно	ПП	безкоштовно	ПП	ПП

Основними критеріями оцінки були можливість прогнозування цін та перевірка нерухомості. Дослідження показало, що жоден з ресурсів не має функції прогнозування цін орендованої нерухомості. Практично всі веб-сайти підтримують функцію публікації власних оголошень, а також фільтрацію інформації за необхідними параметрами. Крім того, компанії практично не мають можливості перевіряти існування орендарів, які зареєстровані на їхньому сайті.

Цей аналіз аналогів серед найбільш відомих веб-сайтів пошуку оренди нерухомості допоміг визначити функціональні можливості, сильні та слабкі сторони, а також перевірити новизну ідеї прогнозування цін. Ми

бачимо доцільність створення комплексної платформи, яка об'єднає ринкових агентів, таких як:

- Орендарі, зацікавлені в покращенні та прискоренні процесу пошуку та оренди житла.
- Власники та агенти з нерухомості, яким потрібно здати житло в оренду.

На нашу думку, тісна співпраця між зазначеними сторонами допоможе чітко визначити цільову аудиторію, виділити переваги системи веб-сервісу як для орендарів, так і для орендодавців (Таблиця 1.2), а також для нових користувачів і власників сервісу (Таблиця 1.3).

Таблиця 1.2

Переваги системи веб-сервісу для оренди нерухомості для орендарів та орендодавців

Для потенційних орендарів	Для орендодавців
Можливість отримувати точні прогнози цін на орендовані об'єкти.	Збільшення кількості заявок завдяки поінформованим орендарям.
Доступ до перевіреної інформації про орендодавців та об'єкти.	Зміцнення довіри до їхнього профілю через систему верифікації та рейтинги.
Зручне використання фільтрів для пошуку об'єктів за критеріями.	Можливість аналізувати ринок і коригувати ставки оренди на основі прогнозів.
Доступ до інтерактивних карт для легкого вибору об'єктів за локацією.	Підвищення конкурентоспроможності через покращення якості послуг та прозорість.
Можливість переглядати відгуки та рейтинги інших орендарів.	

Таблиця 1.3.

Переваги системи веб-сервісу для оренди нерухомості для новачків і власників сервісу

Для новачків (орендарів)	Для власників сервісу
Зростання довіри до сервісу завдяки прозорій системі верифікації та рейтингів.	Залучення нових користувачів завдяки підвищенню якості та функціональності послуг.
Легкість пошуку об'єктів, що відповідають їхнім потребам.	Можливість розширення сервісу за рахунок нових функцій, таких як прогнозування цін та інтерактивні карти.
Підвищення зручності користування сервісом через інтеграцію з іншими корисними інструментами (наприклад, карти, транспорт).	Комерціалізація і монетизація сервісу шляхом впровадження платних функцій та підписок.

1.4 Функціональні вимоги до сервісу оренди житла

Були виділені три основні групи користувачів. Перша група складається з потенційних орендарів, які можуть створювати свої профілі, додавати та верифікувати інформацію про себе, відстежувати доступні пропозиції оренди на ринку житла та знаходити ті, що найбільше відповідають їхнім потребам. Друга група включає орендодавців, які хочуть взаємодіяти з потенційними орендарями, моніторити умови ринку оренди в конкретному сегменті та мати можливість просувати свою нерухомість і програми оренди. Третя група складається з адміністраторів, відповідальних за верифікацію інформації, підтвердження даних і маючих певні важелі для вирішення неетичних дій з боку профілів орендарів і орендодавців. Функціональні вимоги для першої та другої груп визначені в Таблиці 1.4.

Таблиця 1.4

Функціональні вимоги до сервісу оренди житла

Для орендаря	Для орендодавця
Авторизація та реєстрація в системі як користувача.	Авторизація та реєстрація в системі як користувача.

Створення/Редагування/Видалення власного профілю (номер телефону, email, фото, опис).	Доступ до контактів орендаря за взаємною зацікавленістю.
Додавання та верифікація вподобань і потреб (тип нерухомості, локація, кількість кімнат тощо).	Створення/Редагування/Приховування/Закриття профілів об'єктів (фото, описи, параметри).
Пошук та відстеження пропозицій оренди, що відповідають заданим критеріям.	Верифікація даних про нерухомість для підвищення довіри з боку орендарів.
Перегляд детальних описів квартир за їх характеристиками.	Прогнозування ціни квартири на основі її характеристик.
Доступ до контактів орендодавця за взаємною зацікавленістю.	Відстеження заявок від потенційних орендарів.
Перегляд профілів орендодавців.	Перегляд профілів орендарів та їхніх рейтингів.
Надсилання запитів на перегляд/оренду квартир.	Залишення відгуків про орендарів та перегляд їхніх відгуків про орендовану нерухомість.
Прогнозування ціни квартири на основі її характеристик.	Прогнозування ціни квартири на основі її характеристик.
Залишення відгуків та оцінок орендованої нерухомості та орендодавців.	

Функціональність для адміністраторів описана нижче:

Функціональні вимоги до адміністраторів сервісу оренди житла:

- Авторизація та реєстрація користувачів у системі.
- Перевірка та верифікація інформації від орендарів та орендодавців (наприклад, підтвердження сертифікатів, документів з публічно доступних реєстрів).
- Блокування та видалення профілів орендодавців і орендарів у разі неетичних дій або порушень правил сервісу.
- Моніторинг відповідності оголошень і профілів правилам сервісу.
- Управління та модерація контенту (оголошень, відгуків, рейтингів) для забезпечення якості та надійності сервісу.

Таким чином, система буде забезпечувати прозорість та надійність взаємодії між орендарями, орендодавцями та адміністраторами, сприяючи підвищенню якості послуг та задоволенню користувачів.

РОЗДІЛ 2

ЦИФРОВА ТРАНСФОРМАЦІЯ ПРОГНОЗУВАННЯ ЦІН НА НЕРУХОМІСТЬ З ВИКОРИСТАННЯМ МАШИННОГО НАВЧАННЯ

Прогнозування цін на житло є важливою частиною досліджень у галузі нерухомості, оскільки це дозволяє оцінювати ринкову вартість активів і прогнозувати майбутні зміни на ринку. Існує багато методів для оцінки вартості нерухомості, включаючи класичні регресійні моделі та новітні підходи, такі як машинне навчання та нейронні мережі.

Регресійні моделі, що базуються на статистичних методах, вже давно застосовуються для оцінки нерухомості, як це описано в роботах [2] і [3]. Однак ці методи мають обмеження, особливо коли йдеться про складні ринки з великою кількістю факторів, що впливають на ціни.

Сучасні тенденції включають використання алгоритмів машинного навчання, які дозволяють працювати з великими обсягами даних та враховувати складні зв'язки між змінними, як це зазначено у [1] і [22]. Ці нові технології відкривають широкі можливості для підвищення точності прогнозів цін на нерухомість і допомагають вирішувати проблеми, з якими стикаються традиційні методи.

2.1. Традиційні методи оцінки нерухомості

Регресійний аналіз є одним з найбільш поширених методів оцінки нерухомості, що використовується протягом десятиліть. Він дозволяє визначити взаємозв'язок між різними факторами, які впливають на ціну житла, такими як розмір, місцезнаходження, стан об'єкта, наявність інфраструктури тощо. Цей метод широко описаний у наукових працях і має свої переваги та недоліки.

Основи регресійного аналізу в оцінці нерухомості. Як зазначено у

роботах Mark і Goldberg (1988) [2], регресійний аналіз використовує математичну модель для прогнозування значення залежної змінної (ціни) на основі незалежних змінних (фактори, що впливають на ціну). Наприклад, площа житлового об'єкта або його віддаленість від центру міста можуть виступати незалежними змінними.

Isakson (1998) [4] зазначає, що метод регресійного аналізу є ефективним для масової оцінки нерухомості, оскільки дозволяє автоматизувати процес і обробляти велику кількість даних. Ця методика активно використовується у масовій оцінці нерухомості для податкових цілей та управління державними базами даних.

Використання регресії у реальних прикладах показується у роботі Wolverton (2009) [3] регресійний аналіз описується як один із найнадійніших методів для оцінки житлової нерухомості. Він особливо корисний у ситуаціях, коли необхідно враховувати численні фактори, які можуть впливати на ринкову вартість об'єкта. Автор наголошує, що точність регресійної моделі залежить від якості даних та правильного вибору незалежних змінних.

Shiller (1991) [5] досліджував використання регресійних моделей для аналізу повторних продажів житлової нерухомості. Його дослідження показало, що метод повторних продажів може дати більш точну оцінку вартості, оскільки він враховує зміни на ринку між продажами одного й того ж об'єкта.

Недоліки регресійного аналізу:

- Незважаючи на популярність регресійного аналізу, існує кілька проблем, пов'язаних із його використанням. Mark і Goldberg (1988) [2] вказують, що регресійні моделі можуть бути ненадійними, якщо дані, на яких вони базуються, містять помилки або якщо важливі змінні були

виключені з аналізу. Крім того, Isakson (1998) [4] підкреслює, що цей метод може бути обмежений у випадках, коли на ринку нерухомості відбуваються різкі зміни, які важко врахувати через лінійну природу регресії.

- Інша проблема полягає у припущенні, що взаємозв'язок між змінними є лінійним. На практиці ці зв'язки можуть бути складнішими, що знижує точність прогнозів. У випадку таких складних ринкових умов сучасні методи, як-от машинне навчання, можуть виявитися ефективнішими.

Регресійний аналіз залишається важливим інструментом в оцінці нерухомості, особливо для масових оцінок і побудови простих моделей. Однак його обмеження щодо врахування нелінійних зв'язків і зміни ринкових умов підкреслюють необхідність використання більш гнучких і складних підходів, таких як машинне навчання.

2.2. Просторова економетрика

Просторова економетрика — це набір методів, які враховують просторову залежність у даних, що особливо важливо в дослідженнях нерухомості. Ці методи дозволяють моделювати вплив просторових чинників, таких як близькість до центрів економічної активності або вплив сусідніх об'єктів на ринкову вартість житла. Просторова економетрика є важливою для точного прогнозування цін, оскільки нерухомість має чітко виражену географічну прив'язаність.

* Значення просторової залежності

Ключовою ідеєю просторової економетрики є те, що ціни на житло в одному місці залежать не тільки від характеристик самого об'єкта, але й від вартості сусідніх об'єктів. Anselin (1988) [10], який став піонером у розвитку

цієї галузі, описав базові просторові моделі та методи, такі як просторова авторегресійна модель (SAR) та модель просторових помилок (SEM). Ці моделі дозволяють врахувати взаємозв'язки між об'єктами нерухомості на певній території.

Wilhelmsson (2022) [9] досліджував застосування просторових моделей у сфері нерухомості і підкреслив важливість урахування географічного положення та просторової кореляції при аналізі ринку житла. Просторові моделі показують, що сусідні об'єкти нерухомості можуть мати значний вплив на ціну одного об'єкта, що пояснюється "ефектом сусідства".

* Приклади використання просторової економетрики

У роботі Paez і Scott (2004) [11] були представлені різні статистичні методи для аналізу міських ринків, у тому числі з використанням просторових змінних. Автори підкреслюють, що традиційні регресійні моделі не можуть повністю врахувати просторову залежність, що може призвести до викривлення результатів. Використання просторових методів допомагає зменшити ці проблеми та підвищити точність прогнозів.

Bourassa, Cantoni і Hoesli (2010) [12] порівняли різні методи прогнозування цін на житло з урахуванням просторової залежності. Вони виявили, що просторова авторегресія і методи просторової кореляції надають більш точні прогнози у порівнянні з традиційними регресійними моделями, оскільки вони враховують вплив географічних сусідів.

* Переваги та виклики просторових моделей

Однією з основних переваг просторових моделей є те, що вони дозволяють врахувати локалізацію нерухомості і вплив зовнішніх факторів, які можуть значно відрізнятися на різних ринках. Це є критичним для ринків нерухомості, де ціни залежать від близькості до інфраструктурних

об'єктів, таких як школи, магазини, транспортні вузли тощо.

Однак просторові моделі також мають свої обмеження. Paetz і Scott (2004) [11] відзначають складність побудови моделей через велику кількість змінних і потребу в високоякісних просторових даних. Не всі ринки мають достатньо детальну інформацію для точного моделювання просторових взаємозв'язків.

Вплив транспортних інновацій на ринок житла. Один із прикладів впливу просторових факторів — це дослідження Gibbons і Machin (2005) [8], де вони проаналізували, як зміни в транспортній інфраструктурі впливають на ціни на житло. Автори показали, що нові транспортні розв'язки значно підвищують вартість житла, розташованого поблизу нових ліній метро або швидкісних доріг. Це підтверджує важливість врахування не тільки характеристик самого житла, але й його просторового положення відносно інфраструктурних об'єктів.

Просторова економетрика відіграє важливу роль у сучасних дослідженнях ринку нерухомості. Моделі, які враховують просторові залежності, дозволяють більш точно прогнозувати ціни на житло, зокрема в ситуаціях, коли важливими є зовнішні фактори, такі як близькість до інфраструктури або вплив сусідніх об'єктів. Однак для точного застосування цих моделей потрібні високоякісні просторові дані, що може бути проблемою на деяких ринках.

2.3. Фактори впливу на ціни нерухомості

Ціни на нерухомість залежать від багатьох факторів, включаючи як внутрішні характеристики самого об'єкта, так і зовнішні впливи, які пов'язані з економічними, соціальними, інфраструктурними та просторовими умовами. У цьому розділі розглянемо основні фактори, що визначають ціни на житло, а також приклади їх впливу з літератури.

* Внутрішні характеристики нерухомості

Основні характеристики нерухомості, такі як площа, кількість кімнат, тип будівлі, рік побудови та стан житла, є ключовими факторами, які визначають ринкову вартість об'єкта. Ці фактори зазвичай враховуються у більшості моделей оцінки нерухомості, включаючи традиційні регресійні моделі та моделі машинного навчання.

Mark і Goldberg (1988) [2] зазначають, що площа житлового приміщення та наявність додаткових зручностей (наприклад, гаража або басейну) суттєво впливають на кінцеву оцінку. Аналіз цих факторів є основою багатьох регресійних моделей, де ці характеристики виступають незалежними змінними.

* Локація та просторові фактори

Локація нерухомості є одним з найважливіших факторів, що впливають на її вартість. Як уже згадувалося в попередньому розділі про просторову економетрику, ціна на житло значною мірою залежить від географічного розташування об'єкта. Чим ближче житло до центрів економічної активності, транспортних вузлів, шкіл та інших важливих об'єктів інфраструктури, тим вища його вартість.

Gibbons і Machin (2005) [8] досліджували вплив транспортної доступності на ціни на житло, показуючи, що нові транспортні розв'язки значно підвищують вартість житлової нерухомості поблизу. Це особливо важливо для міст з високим рівнем урбанізації, де транспортна інфраструктура визначає зручність життя.

* Економічні умови

Економічна ситуація в регіоні або країні також значно впливає на ціни на житло. Інфляція, рівень зайнятості, ставки за іпотечними кредитами та

загальний економічний розвиток мають безпосередній вплив на попит і пропозицію на ринку нерухомості.

Wu та ін. (2018) [16] відзначають, що макроекономічні показники, такі як рівень інфляції та процентні ставки, мають критичний вплив на ринок житла. Коли економіка розвивається, підвищується попит на житло, що сприяє росту цін. Навпаки, під час економічної кризи попит на нерухомість падає, що призводить до зниження цін.

* Соціальні та демографічні фактори

Соціальні та демографічні фактори також грають важливу роль у визначенні вартості нерухомості. Зміни в демографії населення, такі як збільшення чисельності молодих сімей або старіючого населення, можуть суттєво вплинути на попит на житло в певних районах.

Дослідження Wolverton (2009) [3] показують, що соціальні тенденції, зокрема рівень народжуваності або збільшення кількості мігрантів, можуть створювати підвищений попит на нерухомість у певних регіонах. Це, в свою чергу, призводить до зростання цін на житло.

* Політичні та регуляторні фактори

Політика уряду та регулювання ринку нерухомості мають значний вплив на формування цін. Наприклад, зміни в податковому законодавстві або введення нових будівельних норм можуть стимулювати або обмежувати розвиток ринку. Програми підтримки покупців житла, такі як субсидії на іпотечні кредити, також можуть вплинути на ринок.

Штучні обмеження на будівництво в певних регіонах або введення зонування можуть суттєво вплинути на ціни через зміну пропозиції на ринку. Fleming і Nellis (1994) [7] підкреслюють, що регулювання забудови і доступ до земельних ресурсів часто впливають на ціноутворення в окремих

регіонах.

Як висновок ціни на житло залежать від багатьох факторів, які можна умовно поділити на внутрішні (характеристики об'єкта) і зовнішні (локація, економічні, соціальні та політичні фактори). Всі ці фактори мають значний вплив на ринок нерухомості і повинні враховуватися при моделюванні та прогнозуванні цін на житло. Сучасні підходи, такі як машинне навчання, допомагають інтегрувати більшу кількість факторів для точнішого прогнозування.

2.4. Машинне навчання у прогнозуванні цін на нерухомість

Останніми роками машинне навчання стало одним із найперспективніших підходів для аналізу і прогнозування цін на нерухомість. Завдяки здатності обробляти великі обсяги даних і виявляти складні закономірності між різними факторами, алгоритми машинного навчання пропонують точніші прогнози, ніж традиційні методи, такі як регресія. У цьому розділі розглянемо, як машинне навчання використовується для оцінки вартості житла, і які підходи при цьому застосовуються.

*** Порівняння машинного навчання з регресійними моделями**

Foryś (2022) [1] у своїй роботі порівняла використання регресійних моделей і методів машинного навчання для прогнозування цін на житло. Результати дослідження показали, що моделі машинного навчання забезпечують кращу точність у порівнянні з традиційними підходами, оскільки вони можуть обробляти більшу кількість змінних і виявляти складні нелінійні взаємозв'язки між факторами, які впливають на ціни.

Truong et al. (2020) [18] також зазначають, що регресійні моделі обмежені лінійними припущеннями, тоді як методи машинного навчання,

зокрема нейронні мережі та деревовидні моделі рішень, здатні обробляти більш складні дані і точніше прогнозувати зміни на ринку.

* Використання нейронних мереж

Одним з найпопулярніших підходів у сфері машинного навчання для прогнозування цін на житло є нейронні мережі. Ці моделі імітують роботу людського мозку і можуть навчатися на великій кількості даних, знаходячи приховані патерни і закономірності.

Peterson і Flanagan (2009) [13] дослідили застосування нейронних мереж для гедонічного ціноутворення, де ціна об'єкта визначається набором характеристик. Вони виявили, що нейронні мережі перевершують традиційні регресійні моделі в точності прогнозування завдяки здатності враховувати складні взаємозв'язки між змінними.

Wu та ін. (2018) [16] також використовували нейронні мережі для аналізу цін на нерухомість і підтвердили, що цей метод значно покращує результати у порівнянні з класичними моделями. Нейронні мережі ефективно обробляють великі обсяги даних і добре підходять для аналізу нелінійних змін на ринку.

* Інші методи машинного навчання

Окрім нейронних мереж, широко використовуються й інші методи машинного навчання, такі як випадкові ліси, градієнтний бустинг і методи підтримки векторів. Кожен із цих підходів має свої переваги залежно від типу даних і завдання.

Phan (2018) [21] застосував метод підтримки векторів для прогнозування цін на житло в Мельбурні, Австралія, і показав, що цей метод є ефективним при роботі з великими обсягами даних і може конкурувати з нейронними мережами за точністю.

Curry, Morgan і Silver (2022) [14] також порівняли нейронні мережі з нелінійними статистичними методами для моделювання цін на нерухомість і виявили, що машинне навчання дозволяє гнучкіше адаптуватися до змін ринку, забезпечуючи більш точні прогнози у складних умовах.

* Гібридні моделі машинного навчання

Окрім використання окремих методів машинного навчання, дослідники також працюють над створенням гібридних моделей, які поєднують кілька алгоритмів для підвищення точності. Zhan et al. (2023) [24] представили гібридну модель, яка поєднує методи глибокого навчання та традиційні статистичні підходи. Ця модель дозволяє краще прогнозувати ціни на житло завдяки інтеграції різних підходів і зменшенню похибок.

Такий підхід показує, що комбінування переваг різних моделей машинного навчання може значно покращити прогнозування цін на житло і допомагає обійти обмеження окремих алгоритмів.

* Переваги і виклики використання машинного навчання

Основними перевагами використання машинного навчання для прогнозування цін на нерухомість є висока точність і можливість враховувати складні взаємозв'язки між факторами. Алгоритми машинного навчання можуть обробляти великі набори даних, які містять багато змінних, і виявляти взаємозв'язки, які важко виявити за допомогою традиційних методів.

Однак є також виклики, пов'язані з використанням цих технологій. Fan та ін. (2018) [20] зазначають, що машинне навчання вимагає значних обчислювальних ресурсів і високоякісних даних. Крім того, побудова моделей може бути складною через необхідність коригування параметрів і уникнення перенавчання, коли модель надто добре підлаштовується під

дані, але втрачає здатність робити точні прогнози на нових даних.

Висновком буде що машинне навчання є потужним інструментом для прогнозування цін на нерухомість, дозволяючи враховувати складні нелінійні взаємозв'язки між факторами і працювати з великими даними. Методи, такі як нейронні мережі, випадкові ліси і гібридні моделі, показують високу ефективність у порівнянні з традиційними регресійними моделями. Однак для досягнення максимальної точності необхідні якісні дані і ретельна побудова моделей.

2.5. Змішана оцінка та індекси цін на житло

Змішана оцінка та індекси цін на житло є важливими інструментами для аналізу змін на ринку нерухомості. Вони дозволяють відслідковувати динаміку цін у часі, а також враховувати особливості різних сегментів ринку. Індекси цін використовуються для вимірювання загальної вартості нерухомості та можуть застосовуватися як для державного регулювання, так і для досліджень у сфері економіки житла.

*** Методи змішаної оцінки**

Метод змішаної оцінки полягає в тому, щоб поєднувати різні фактори, що впливають на ціни, та їхні ваги для створення індексу, який відображає середню вартість нерухомості з урахуванням різних змінних. Forú (2012) [6] представила метод змішаної корекції, який дозволяє коригувати індекси цін на житло, враховуючи типові характеристики нерухомості, такі як площа, кількість кімнат, стан тощо. Це дозволяє уникнути викривлень, викликаних змінами у структурі об'єктів, які входять до вибірки.

Методи змішаної оцінки також використовуються для побудови індексів цін на повторні продажі, як це описано в роботах Shiller (1991) [5].

Він показав, що змішана оцінка може покращити точність індексів, особливо для аналізу динаміки цін на ринках, де часто змінюється структура продажів.

* Індекси цін на житло

Індекси цін на житло є ключовими показниками для відстеження динаміки ринку нерухомості. Ці індекси враховують середні ціни на житло на основі даних про продажі нерухомості, що дозволяє спостерігати за загальними тенденціями цін.

Fleming і Nellis (1994) [7] досліджували різні підходи до вимірювання цін на житло у Великій Британії, підкреслюючи важливість побудови надійних індексів для відображення реальних змін на ринку. Вони розглянули основні джерела даних для вимірювання цін на житло та описали методи, які можуть покращити точність індексів, зокрема, корекцію даних щодо типу і стану нерухомості.

Shiller (1991) [5] запропонував методологію побудови індексів на основі повторних продажів, яка широко використовується для аналізу ринку житлової нерухомості. Цей підхід дозволяє порівнювати ціни тих самих об'єктів, які були продані кілька разів, що зменшує вплив зміни характеристик нерухомості.

* Вимірювання змін на ринку нерухомості

Індекси цін на житло дозволяють не лише аналізувати загальні тенденції на ринку, але й оцінювати вплив економічних факторів на ціни. Наприклад, Federal Housing Finance Agency (FHFA) [19] публікує регулярні індекси цін на житло в США, які використовуються для відстеження економічних змін і впливу державної політики на ринок нерухомості.

Індекси також дозволяють враховувати зміни в окремих сегментах

ринку. Foruś (2012) [6] зазначає, що корекція за типом об'єктів (наприклад, квартири або будинки) дозволяє будувати більш точні індекси, які відображають реальні зміни на ринку. Це є особливо важливим на ринках з великою різноманітністю нерухомості.

* Сучасні методи побудови індексів

Сучасні підходи до побудови індексів цін на житло використовують як традиційні методи статистичного аналізу, так і новітні технології, такі як машинне навчання. Роботи Zhan et al. (2023) [24] демонструють, що методи глибокого навчання можуть бути використані для побудови точних індексів, які враховують велику кількість факторів, включаючи просторові та часові залежності.

Крім того, використання великих даних (Big Data) дозволяє побудувати індекси, які більш точно відображають короткострокові та довгострокові зміни на ринку нерухомості. Mu, Wu і Zhang (2014) [22] показали, що використання відкритих даних з інтернету може значно покращити моделі прогнозування цін і побудову індексів.

Змішана оцінка та індекси цін на житло є важливими інструментами для аналізу ринку нерухомості, що дозволяють враховувати динаміку змін цін і коригувати дані відповідно до змін у структурі ринку. Сучасні методи побудови індексів, такі як машинне навчання та використання великих даних, дозволяють створювати більш точні й надійні інструменти для прогнозування цін.

2.6. Моделі оцінки для окремих ринків

Оцінка вартості житлової нерухомості може сильно варіюватися залежно від регіональних особливостей ринку. Кожен ринок нерухомості

має унікальні характеристики, які необхідно враховувати при розробці моделей прогнозування цін. Ці фактори можуть включати економічні умови, рівень урбанізації, демографічні зміни, місцеві політичні ініціативи та інфраструктурні проекти. У цьому розділі розглянемо дослідження оцінки цін для окремих ринків, включаючи приклади з Австралії, Китаю та інших країн.

* Оцінка цін на ринку Мельбурна, Австралія

Phan (2018) [21] досліджував ринок житлової нерухомості Мельбурна, застосовуючи методи машинного навчання для прогнозування цін на житло. Він відзначив, що ринок Мельбурна відрізняється високою конкуренцією та складною динамікою, тому традиційні методи оцінки не завжди дають точні результати.

У своїй роботі Phan застосував кілька алгоритмів машинного навчання, таких як метод підтримки векторів (SVM) і нейронні мережі, щоб проаналізувати величезні обсяги даних про продажі житлової нерухомості. Висновки дослідження показали, що використання машинного навчання надає точніші прогнози, оскільки ці методи краще справляються з великою кількістю змінних і можуть враховувати складні взаємозв'язки між факторами, що впливають на ціни.

* Ринки нерухомості в Китаї

Lulin і Li (2021) [17] досліджували вторинний ринок житлової нерухомості в перших містах Китаю, таких як Пекін, Шанхай та Шеньчжень. Їх дослідження фокусується на розробці нової моделі оцінки цін, заснованої на алгоритмах машинного навчання. Враховуючи специфіку швидкого розвитку китайських мегаполісів, їх модель враховувала фактори, пов'язані з урбанізацією, доходами населення, інфраструктурними змінами та міграційними потоками.

Автори виявили, що застосування гібридної моделі, яка поєднує методи регресії та машинного навчання, дозволяє значно підвищити точність прогнозування цін на житло. Це важливо на динамічних ринках, де ціни можуть змінюватися через економічні та політичні фактори, такі як державна політика щодо обмеження спекуляцій на ринку нерухомості.

* Інші регіональні ринки

Truong та ін. (2020) [18] провели дослідження на ринках житлової нерухомості в Азії, застосовуючи удосконалені методи машинного навчання. Їхнє дослідження показало, що ринки Азії, зокрема В'єтнам і Таїланд, мають унікальні особливості, пов'язані з швидкою урбанізацією та нерівномірним економічним зростанням.

Дослідники виявили, що традиційні методи оцінки вартості житла не дають точних результатів на таких ринках через високу нестабільність і значні соціально-економічні відмінності між регіонами. Застосування методів глибокого навчання та аналізу великих даних дозволило краще моделювати ціни на нерухомість в умовах високої мінливості ринку.

* Вплив локальних особливостей на прогнозування цін

Однією з основних проблем при оцінці вартості житлової нерухомості в окремих регіонах є необхідність врахування локальних особливостей, які можуть суттєво впливати на попит і пропозицію. Наприклад, інфраструктурні проекти, зміни в транспортній доступності, місцеві податкові ініціативи або екологічні фактори можуть мати значний вплив на ціни на житло в певному регіоні.

Gibbons і Machin (2005) [8] досліджували вплив транспортних інновацій на ціни на житло у Великій Британії. Вони показали, що нові транспортні розв'язки можуть суттєво підвищувати вартість житлової

нерухомості у районах, де покращилася транспортна інфраструктура. Це свідчить про те, що важливо враховувати регіональні особливості, коли розробляються моделі оцінки вартості нерухомості для окремих ринків.

* Застосування великих даних для аналізу регіональних ринків

Використання великих даних стає важливим інструментом для аналізу ринків нерухомості у різних регіонах. Mu, Wu і Zhang (2014) [22] показали, що великі дані з відкритих джерел, такі як онлайн-платформи продажу нерухомості, можуть значно покращити точність оцінки цін на житло. Це особливо актуально для ринків, де традиційні дані можуть бути обмежені або недоступні.

Застосування великих даних дозволяє враховувати більшу кількість змінних, включаючи регіональні та просторові фактори, що робить моделі оцінки більш гнучкими та точними. Це підхід, який може бути особливо корисним для аналізу ринків у країнах з різко мінливою економічною ситуацією.

Ринки нерухомості в різних регіонах мають свої унікальні особливості, які необхідно враховувати при прогнозуванні цін. Використання машинного навчання, великих даних і гібридних моделей дозволяє точніше оцінювати вартість житла на специфічних ринках, враховуючи такі фактори, як урбанізація, економічне зростання та інфраструктурні зміни. Однак важливо пам'ятати, що моделі оцінки повинні адаптуватися до локальних умов, щоб забезпечити максимальну точність.

2.7. Актуальні тенденції в аналізі цін на нерухомість

Сучасні тенденції в аналізі цін на нерухомість зосереджені на застосуванні новітніх технологій, таких як машинне навчання, використання великих даних, а також вдосконалення інструментів

прогнозування з урахуванням соціально-економічних і екологічних факторів. У цьому розділі розглянемо основні сучасні підходи, які визначають розвиток оцінки вартості житлової нерухомості.

* Використання великих даних (Big Data)

Однією з найбільших змін на ринку нерухомості стало широке застосування великих даних для аналізу та прогнозування цін на житло. Великі дані дозволяють обробляти величезні обсяги інформації з різних джерел, таких як онлайн-платформи продажу житла, соціальні мережі, державні бази даних та супутникові знімки. Використання таких даних значно підвищує точність прогнозів, оскільки враховує ширший спектр змінних, які можуть впливати на ціни.

Wu та ін. (2018) [16] описують використання відкритих даних з інтернету для побудови моделей прогнозування цін на житло в міських умовах. Вони виявили, що великі дані допомагають виявляти тренди, які не можуть бути визначені через обмеження традиційних наборів даних. Це дозволяє швидше реагувати на зміни на ринку нерухомості та робити точніші прогнози.

Інший приклад — Mu, Wu і Zhang (2014) [22], які показали, що використання даних з онлайн-платформ може значно покращити оцінку цін на житло. Вони наголошують на тому, що такі дані дають більше деталей про транзакції та дозволяють побачити в реальному часі, як змінюються ціни в різних регіонах.

* Використання машинного навчання та штучного інтелекту

Машинне навчання та штучний інтелект (AI) продовжують відігравати ключову роль в удосконаленні моделей оцінки вартості житла. Здатність цих технологій обробляти великі обсяги даних, виявляти складні

закономірності та забезпечувати високу точність прогнозування робить їх незамінними в сучасних умовах.

Fan, Cui і Zhong (2018) [20] використовували методи машинного навчання для прогнозування цін на житло і показали, що ці методи значно перевершують традиційні моделі, особливо в умовах високої динаміки ринку. Їх дослідження показує, що алгоритми машинного навчання, такі як нейронні мережі, забезпечують гнучкість у роботі з різними типами даних і дозволяють адаптувати моделі до різних ринкових умов.

Zhan та ін. (2023) [24] також представили гібридну модель машинного навчання, яка поєднує кілька алгоритмів для підвищення точності прогнозування. Такий підхід дозволяє зменшити похибки і покращує здатність моделі передбачати короткострокові й довгострокові зміни на ринку.

* Врахування екологічних і соціальних факторів

Все більше дослідників приділяють увагу соціальним та екологічним факторам при оцінці вартості житлової нерухомості. Це включає такі аспекти, як енергоефективність будинків, близькість до зелених зон, доступ до громадського транспорту, а також екологічні ризики (наприклад, близькість до районів, схильних до затоплень або забруднення).

Врахування екологічних факторів стає дедалі актуальнішим у міру зростання попиту на "зелені" будинки та більш екологічно стійкі райони. Gibbons і Machin (2005) [8] підкреслюють, що наявність зелених зон та доступність екологічної інфраструктури значно впливають на ціни на житло, особливо у великих містах.

* Зростаюча роль автоматизації в оцінці нерухомості

Автоматизовані системи оцінки нерухомості (AVM — Automated

Valuation Models) набирають популярності завдяки своїй здатності швидко обробляти дані і генерувати точні оцінки вартості житлової нерухомості. Такі системи, засновані на алгоритмах машинного навчання, можуть автоматично аналізувати велику кількість даних і видавати результати, які часто перевищують точність традиційних підходів.

Worzala та ін. (2021) [15] досліджують технологічний вплив на оцінку нерухомості, наголошуючи, що автоматизовані моделі оцінки стають важливим інструментом для інвесторів і агентств з нерухомості. Вони дозволяють швидко оцінювати великі портфелі нерухомості, використовуючи мінімум людських ресурсів, і водночас забезпечують високу точність прогнозів.

* Використання відкритих даних для створення індексів цін

Іншою важливою тенденцією є створення індексів цін на основі відкритих даних. Це дозволяє аналізувати ринок у режимі реального часу і забезпечує кращу прозорість для всіх учасників ринку.

Federal Housing Finance Agency (2019) [19] використовує індекси цін на житло для моніторингу ринкових тенденцій і оцінки змін у вартості житла на національному та регіональному рівнях. Відкриті дані та індекси стають важливим інструментом для економістів і державних органів, дозволяючи відслідковувати зміни на ринку і приймати обґрунтовані рішення.

РОЗДІЛ 3

РОЗРОБКА ВЕБ-СЕРВІСУ ПРОГНОЗУВАННЯ ЦІН НА НЕРУХОМІСТЬ

3.1. Загальна функціональність веб-сервісу

Для детального опису загальної функціональної системи на концептуальному рівні ми використали діаграму варіантів використання (Рис. 3.1). Вона показує зовнішні сутності (акторів), які взаємодіють із системою та використовують її функціональність, представляючи всі групи користувачів сервісу. Кожен варіант використання на діаграмі представляє окремі функціональні блоки системи.

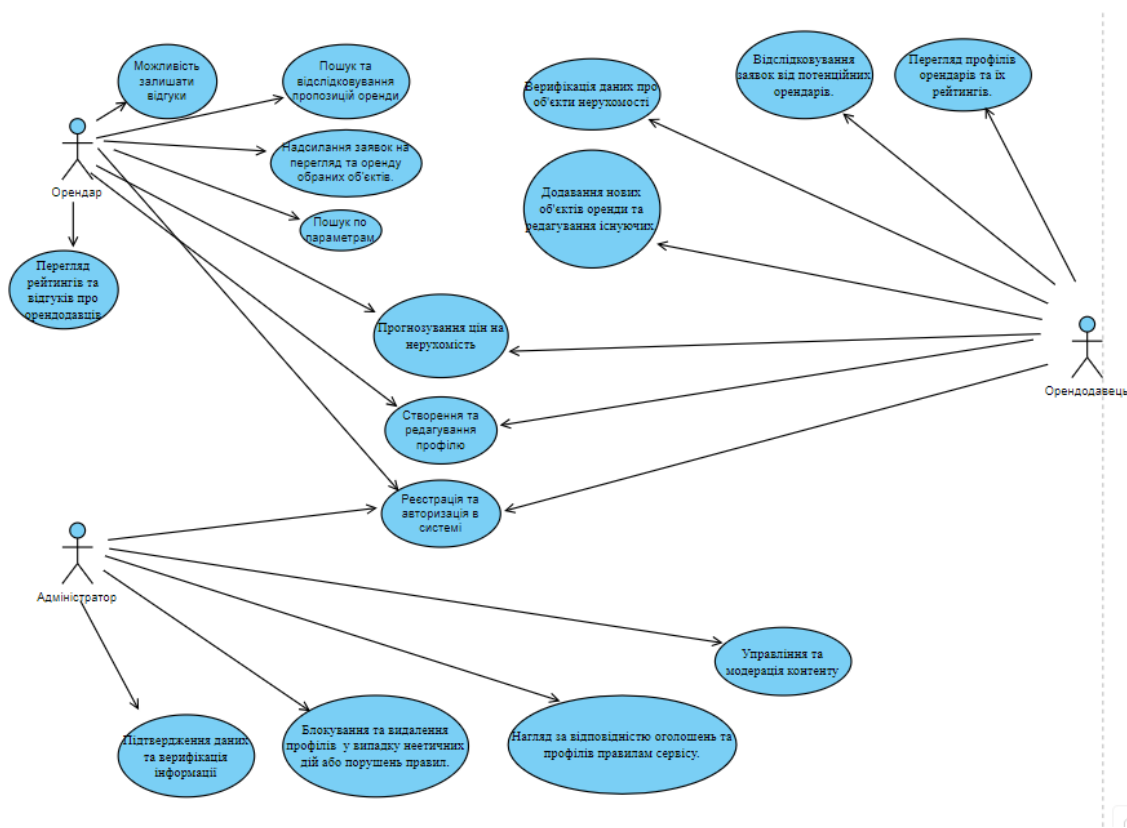


Рис. 3.1. Діаграма варіантів використання функцій веб-сервісу прогнозування цін на нерухомість

Діаграма варіантів використання надала чітке розуміння можливостей кожної групи агентів та дозволила визначити функціональні взаємодії між ними. Це дозволило згрупувати функціональні вимоги у

модулі на основі споріднених функцій агентів. Високорівнева архітектура майбутнього веб-сервісу ілюструє зв'язки між цими модулями. Декомпозиція системи допомогла визначити 11 функціональних модулів, деякі з яких є універсальними і обслуговують кілька груп агентів (Рис. 3.2).

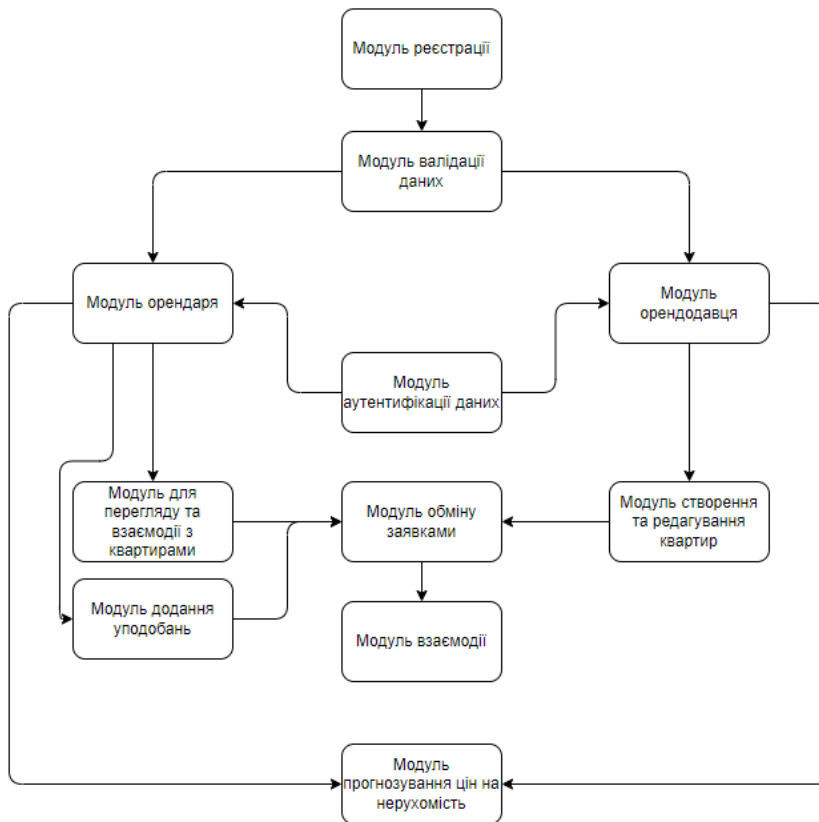


Рис. 3.2: Високорівнева архітектура веб-сервісу прогнозування цін на нерухомість

3.2 Функціональні модулі системи веб-сервісу для прогнозування цін на нерухомість

Для забезпечення повноцінного функціонування веб-сервісу для оренди нерухомості з функцією прогнозування цін було розроблено ряд функціональних модулів. Кожен з них виконує специфічні завдання, що полегшують процес взаємодії між орендарями та орендодавцями, підвищують точність прогнозів, забезпечують зручний інтерфейс для управління оголошеннями та здійснюють обмін заявками. Нижче

представлено детальний опис кожного модуля, їхні функціональні можливості та внесок у роботу системи

Модуль реєстрації - Цей модуль відповідає за створення нових облікових записів користувачів у системі. Він надає форму для введення даних, таких як ім'я користувача, пароль, тип користувача (орендар або орендодавець) та інші базові дані. Після підтвердження реєстрації дані зберігаються в базі даних, а користувач отримує доступ до системи.

Модуль валідації даних - Модуль валідації перевіряє правильність і повноту даних, які вводять користувачі. Для цього використовуються як автоматичні, так і ручні перевірки. Наприклад, модуль може перевірити коректність заповнення профілю орендодавця та наданої інформації про нерухомість, що підвищує рівень довіри до даних у системі.

Модуль аутентифікації - Цей модуль відповідає за підтвердження особи користувача, забезпечуючи безпеку системи. Він включає функціонал для логіну, перевірки правильності введених даних, а також системи відновлення паролю. Аутентифікація дає можливість лише авторизованим користувачам взаємодіяти з системою.

Модуль орендаря - Модуль орендаря надає функції, які доступні лише орендарям. Користувачі цього типу можуть переглядати пропозиції квартир, додавати свої уподобання, взаємодіяти з орендодавцями та надсилати заявки на перегляд або оренду квартир. Модуль забезпечує спеціальний інтерфейс для зручного пошуку житла відповідно до потреб орендаря.

Модуль орендодавця - Модуль орендодавця надає орендодавцям можливість керувати своїми оголошеннями. Власники нерухомості можуть створювати, редагувати та видаляти оголошення про оренду квартир, отримувати та обробляти заявки від орендарів і переглядати профілі потенційних орендарів. Цей модуль покращує процес управління

пропозиціями для орендодавців.

Модуль перегляду та взаємодії з квартирами - Дозволяє користувачам переглядати доступні оголошення про оренду квартир, з повною інформацією про кожну квартиру. Цей модуль забезпечує сортування та фільтрацію (Рис. 3.3) за різними критеріями (кількість кімнат, район, площа тощо), а також інтерактивність, дозволяючи орендарям відзначати цікаві пропозиції та надсилати заявки орендодавцю.

Available Apartments

Район: Мин. ціна: Макс. ціна: Кількість кімнат: Можна з тваринами:

[Add New Apartment](#)

- [Urvivska №1](#) - 2 кімнат, 3.0 кв.м., 20000.00 грн
- [Ahmatovoi 43 №1](#) - 2 кімнат, 50.0 кв.м., 24000.00 грн
- [Ahmatovoi №1](#) - 1 кімнат, 21.0 кв.м., 17000.00 грн
- [Knyazhyi zaton №99](#) - 4 кімнат, 103.0 кв.м., 70400.00 грн
- [Hnyuri №49](#) - 1 кімнат, 35.0 кв.м., 40030.00 грн
- [Докотинний 3-й №13](#) - 2 кімнат, 20.9 кв.м., 31790.00 грн
- [Деревообробний 5-й №31](#) - 1 кімнат, 128.64 кв.м., 93094.00 грн
- [Манежна №83](#) - 5 кімнат, 173.97 кв.м., 28893.00 грн
- [Солощоватий №83](#) - 1 кімнат, 32.08 кв.м., 12554.00 грн
- [19-й км Старокілівської дороги №10](#) - 4 кімнат, 117.45 кв.м., 94919.00 грн
- [Дазурний 2-й №31](#) - 3 кімнат, 189.48 кв.м., 11976.00 грн
- [Рівний №10](#) - 3 кімнат, 194.4 кв.м., 20041.00 грн
- [Головна №78](#) - 1 кімнат, 52.02 кв.м., 96897.00 грн
- [Солоний 5-й №2](#) - 3 кімнат, 60.33 кв.м., 16196.00 грн
- [Бороткий №14](#) - 5 кімнат, 157.37 кв.м., 65477.00 грн
- [Василия Кандіського 2-й №12](#) - 2 кімнат, 131.09 кв.м., 50757.00 грн
- [Транспортна №18](#) - 1 кімнат, 26.81 кв.м., 72465.00 грн
- [Грузинська №74](#) - 4 кімнат, 87.18 кв.м., 18432.00 грн
- [Прибережний №27](#) - 4 кімнат, 81.99 кв.м., 72121.00 грн
- [Дня 31-ша №46](#) - 3 кімнат, 150.84 кв.м., 75387.00 грн
- [Журавлина №77](#) - 4 кімнат, 51.62 кв.м., 88854.00 грн
- [Заводська 1-ша №77](#) - 4 кімнат, 72.69 кв.м., 43822.00 грн
- [Космонавтів №99](#) - 4 кімнат, 111.9 кв.м., 72049.00 грн
- [Конструкторський №3](#) - 2 кімнат, 164.02 кв.м., 26281.00 грн
- [Солоний 7-й №61](#) - 5 кімнат, 165.12 кв.м., 77917.00 грн
- [Виноградна №42](#) - 4 кімнат, 21.45 кв.м., 15626.00 грн

Рис. 3.3. Модуль перегляду й фільтрування нерухомості

Модуль додання уподобань - Орендарі можуть вказувати свої особисті вподобання та вимоги до житла (наприклад, район, наявність дозволу на проживання з тваринами, близькість до метро тощо). Цей модуль дозволяє фільтрувати оголошення відповідно до уподобань орендаря, що значно скорочує час на пошук.

Модуль створення та редагування квартир - Забезпечує орендодавцям можливість створювати та редагувати оголошення про квартири. Користувачі можуть додавати детальну інформацію про житло, включаючи адресу, площу, кількість кімнат, поверх, наявність дозволу на тварин та близькість до метро. Завдяки цьому модулю оголошення завжди актуальні

та відповідають очікуванням орендарів.

Модуль обміну заявками - Забезпечує систему запитів і відповідей між орендарями та орендодавцями. Наприклад, орендар може надіслати запит на перегляд квартири, а орендодавець має можливість схвалити або відхилити цей запит. Таким чином, цей модуль забезпечує прозору та зручну комунікацію між сторонами (Рис. 3.4).

Ваші запити на перегляд контактної інформації

- Квартира: [Кnyazhiy zaton](#)
 Статус: **Погоджено**
 Контактний телефон орендодавця: 103
 Контактний емейл: user3@test.com
- Квартира: [Кnyazhiy zaton](#)
 Статус: **Погоджено**
 Контактний телефон орендодавця: 103
 Контактний емейл: user3@test.com

Рис. 3.4. Модуль обміну заявками на спілкування

Модуль взаємодії - Модуль взаємодії дозволяє орендарям і орендодавцям обмінюватися контактною інформацією після схвалення заявки. Після підтвердження запиту орендодавець може отримати контактні дані орендаря, а орендар – доступ до контактів орендодавця. Це полегшує подальше спілкування між сторонами та допомагає швидко організувати перегляд житла (Рис. 3.5).

Send Rental Request for Knyazhiy zaton

Message:

Рис. 3.5. Модуль взаємодії між орендарями та орендодавцями

Модуль прогнозування цін - Цей модуль використовує алгоритми машинного навчання для оцінки вартості оренди квартир на основі заданих параметрів, таких як кількість кімнат, площа, район, поверховість, близькість до метро тощо. Він дозволяє орендарям і орендодавцям отримати уявлення про ринкову ціну оренди для подібних об'єктів і приймати обґрунтовані рішення (Рис. 3.6).

Прогнозування ціни квартири

Кількість кімнат:

Поверх:

Загальна кількість поверхів:

Площа (м²):

Район:

Близькість до метро:

Можливість проживання з тваринами:

Прогнозована ціна: 51868.42 грн

Рис. 3.6. Модуль прогнозування цін

3.3 Експериментальна частина

Використовуючи модуль прогнозування цін на нерухомість, можна обрати модель машинного навчання, яка забезпечує найкращу якість прогнозування ціни оренди квартири за критерієм RMSE (для моделей прогнозування та класифікації). Для прогнозування ціни оренди квартири залежно від предикторів можна розглянути використання лінійної множинної регресії, випадкового лісу, методів дерева рішень. Для цього необхідно врахувати залежну змінну (ціна оренди на місяць) і 9 незалежних змінних (предикторів) на основі набору даних з 2040 зразків по Києву (Україна):

- **price** - ціна квартири в \$1000
- **totsp** - загальна площа квартири, кв.м.
- **livesp** - житлова площа квартири, кв.м.
- **kitsp** - площа кухні, кв.м.
- **dist** - відстань від квартири до центру міста в км.
- **metrdist** - відстань до станції метро в хвилинали
- **walk**: 1 - пішки від метро, 0 - транспортом
- **brick**: 1 - цегляна будівля, 0 - інша будівля
- **floor**: 1 - поверх, відмінний від першого та останнього, 0 - перший або останній поверх
- **code** - райони міста від 1 до 8

Кореляція між змінними показана на Рис. 3.7

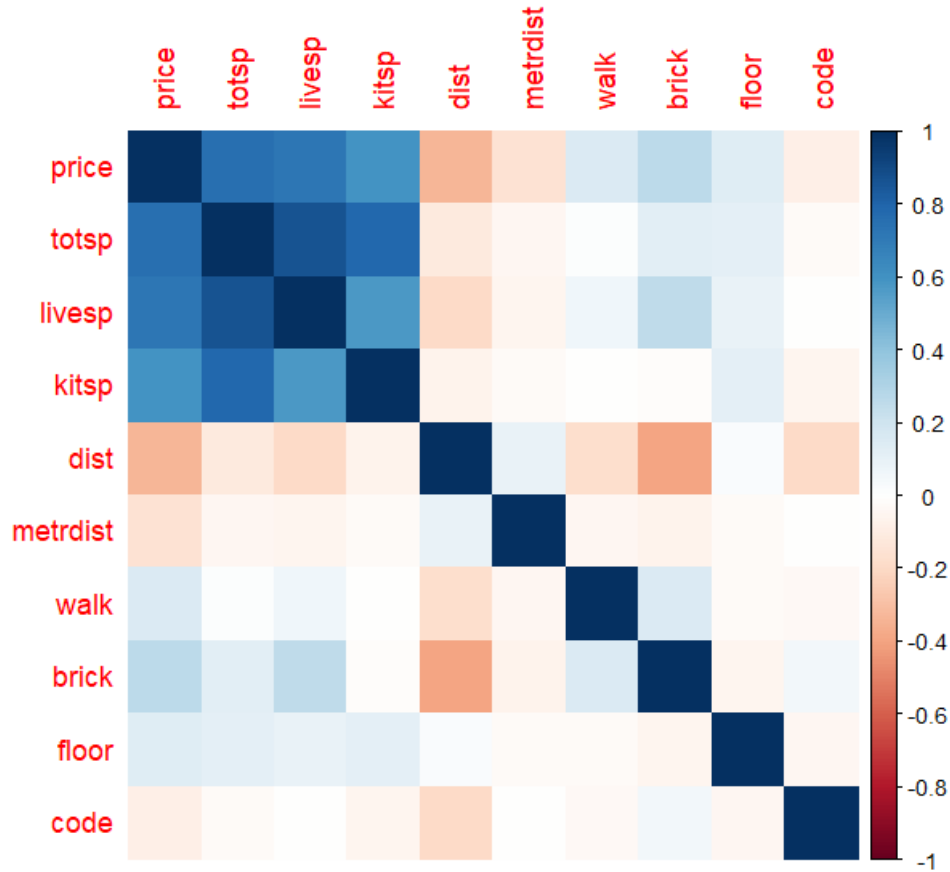


Рис. 3.7. Кореляція між змінними (синій колір відповідає позитивній кореляції, червоний - негативній кореляції)

Гістограма цін на оренду квартир у Києві показана на Рис. 3.8 (n=2040 пропозицій оренди).

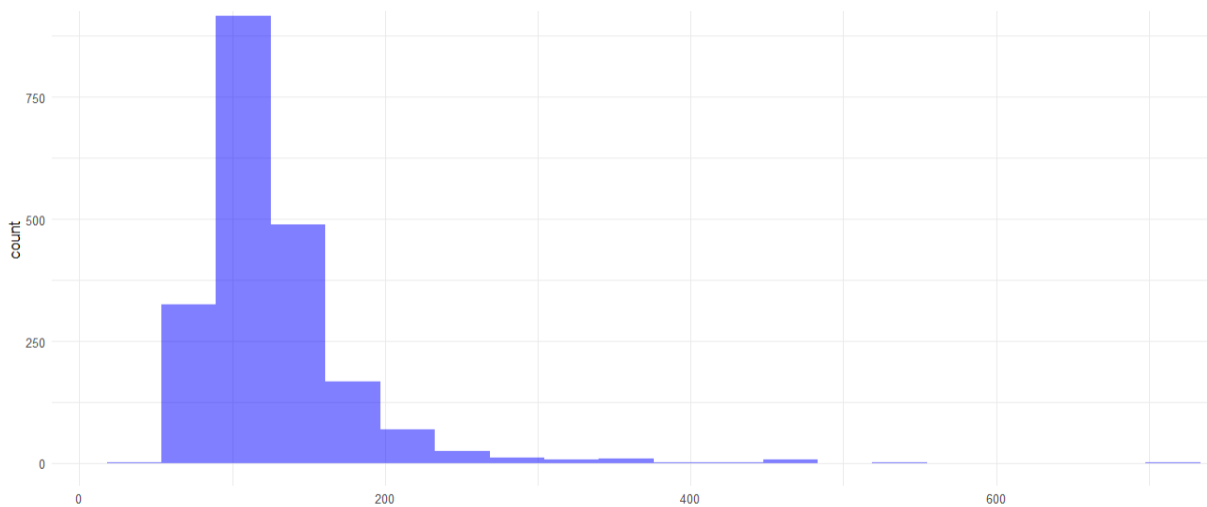


Рис. 3.8: Гістограма цін на оренду квартир у Києві

Розрахунок VIF для предикторів показує, що мультиколінеарності немає ($VIF < 10$), тобто всі предиктори можна використовувати для прогнозування ціни оренди квартири (Рис. 3.9):

```

totsp  livesp  kitsp  dist  metrdist  walk  brick  floor
7.738744 4.922180 2.851522 1.268720 1.016622 1.056505 1.297474 1.023024
code
1.041753

```

Рис. 3.9: Коефіцієнти VIF для предикторів цін на оренду квартир

Результати лінійної регресії показують, що всі предиктори мають статистично значущий вплив на ціну оренди квартири (Рис. 3.10):

```

lm(formula = price ~ ., data = train)

Residuals:
    Min       1Q   Median       3Q      Max
-110.10  -14.88   -0.70   10.66   410.51

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -26.9336     6.7600  -3.984 7.11e-05 ***
totsp         1.8086     0.1416  12.772 < 2e-16 ***
livesp        1.1661     0.2134   5.463 5.52e-08 ***
kitsp         1.2010     0.4775   2.515 0.012006 *
dist         -3.1897     0.2631 -12.125 < 2e-16 ***
metrdist     -1.3464     0.2081  -6.470 1.34e-10 ***
walk         10.0979     1.7394   5.805 7.90e-09 ***
brick         8.6527     1.9443   4.450 9.24e-06 ***
floor         6.8975     1.9744   3.493 0.000491 ***
code         -2.5899     0.3668  -7.061 2.58e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 30.03 on 1424 degrees of freedom
Multiple R-squared:  0.6976,    Adjusted R-squared:  0.6957
F-statistic: 365.1 on 9 and 1424 DF,  p-value: < 2.2e-16

```

Рис. 3.10: Модель лінійної регресії для прогнозування ціни оренди квартири

При цьому предиктори **totsp**, **livesp**, **kitsp**, **walk**, **brick**, **floor** мають прямий вплив на ціну оренди квартири, тоді як предиктори **dist**, **metrdist**, **code** мають негативний вплив на ціну оренди квартири.

При застосуванні методу випадкового лісу найважливішими факторами впливу є загальна та житлова площа орендованої квартири, площа кухні та відстань від квартири до центру міста (Рис. 3.11).

	%IncMSE
totsp	1855.58661
livesp	816.92381
kitsp	515.73011
dist	402.23938
metrdist	51.80257
walk	63.21499
brick	78.14269
floor	16.17641
code	147.56076

Рис. 3.11. Важливість предикторів у моделі випадкового лісу

Метод дерева рішень розподіляє квартири за найважливішими предикторами (загальна та житлова площа, відстань від квартири до центру міста) (Рис. 3.12).

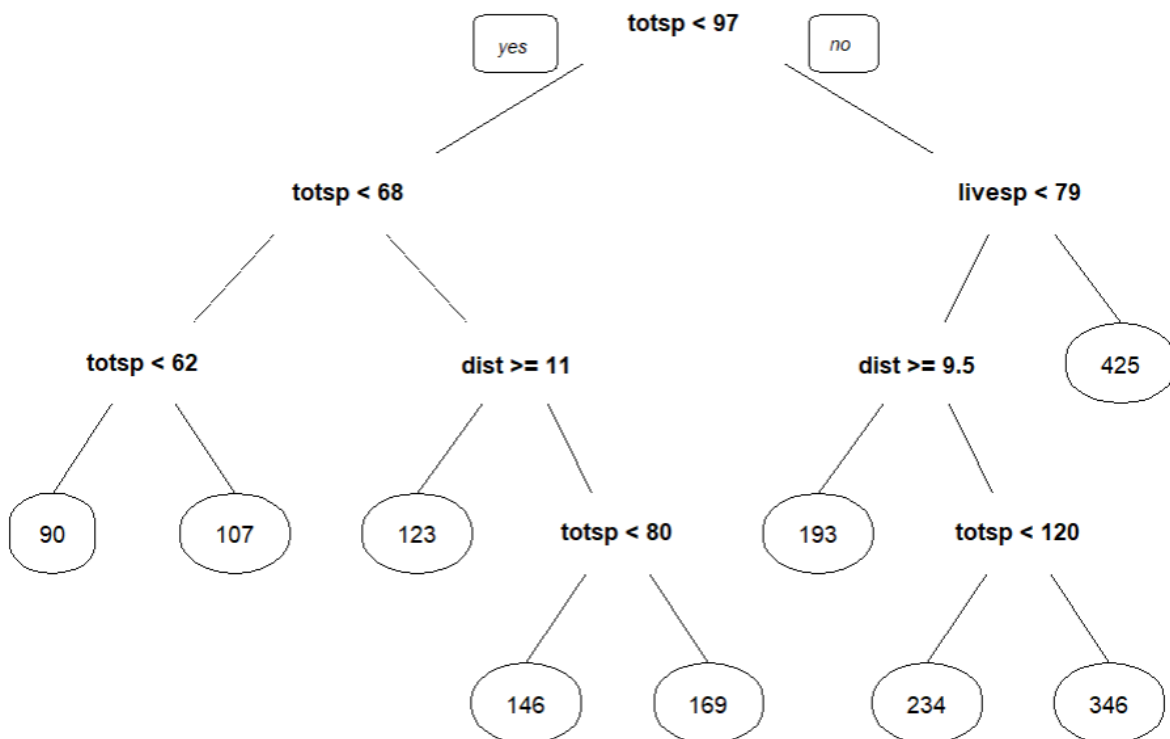


Рис. 3.12. Розподіл квартир за найважливішими предикторами моделі дерева рішень

Після порівняння якості прогнозу для ціни оренди квартири були отримані наступні результати (Таблиця 3.1):

Таблиця 3.1

RMSE методів машинного навчання

Методи	Лінійна регресія	Випадковий ліс	Дерева рі...
RMSE	27.68	23.56	73.82

На основі результатів застосування модуля прогнозування цін на нерухомість було встановлено, що найкращий прогноз для ціни оренди квартири в Києві за критерієм RMSE та визначеними предикторами забезпечує метод випадкового лісу.

ВИСНОВКИ

У результаті дослідження було розроблено веб-додаток, який надає можливість користувачам для пошуку житла поєднувати пошук квартир для оренди та прогнозування цін на орендну оплату. Були розроблені технічні специфікації для орендарів, орендодавців та адміністраторів, що окреслюють їхні ролі в даній системі.

Система включає функціональні модулі для аутентифікації, фільтрації інформації за критеріями, перегляду профілів орендарів і орендодавців, обробки відгуків, а також модуль прогнозування вартості орендної плати.

Крім того, було проведено аналіз прогнозування цін за допомогою програмного забезпечення R. Найкращий результат для прогнозування вартості оренди квартир у Києві показав метод випадкового лісу, але в той же час він більш складний для реалізації по ресурсам й часу, ніж метод лінійної регресії, тому для зручності інтерпретації результатів було вирішено інтегрувати метод лінійної регресії.

У майбутніх етапах роботи над веб-сервісом можливе вдосконалення модуля прогнозування шляхом додавання нових, більш значущих предикторів, які можуть суттєво впливати на точність прогнозу. Крім того, замість лінійної регресії доцільно розглянути ширший набір методів машинного навчання, серед яких можна буде обрати найбільш відповідний для точності прогнозу ціни оренди за нерухомість у системі. Такий підхід дозволить підвищити точність прогнозів, що, у свою чергу, сприятиме підвищенню якості та ефективності веб-сервісу для користувачів.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Forys, I. (2022). Machine learning in house price analysis: regression models versus neural networks. *Procedia Computer Science*, 207, 435–445.
<https://doi.org/10.1016/j.procs.2022.09.078/>
2. Mark, J., & Goldberg, M. (1988). Multiple regression analysis and mass assessment: a review of the issues. *The Appraisal Journal*, 56(1), 89–109.
3. Wolverton, L. M. (2009). *Introduction to Statistics for Appraisers*. Appraisal Institute, Chicago.
4. Isakson, R. H. (1998). The review of real estate appraisals using multiple regression analysis. *Journal of Real Estate Research*, 15(2), 177-190.
5. Shiller, J. R. (1991). Arithmetic Repeat Sales Price Estimators. *Journal of Housing Economics*, 1(1), 110–126.
6. Forys, I. (2012). Mix-adjustment method of determining residential real estate price indices on the example of cooperative premises. *Studies and Materials of the Scientific Society for Real Estate*, 20(1), 41–52.
7. Fleming, C. M., & Nellis, G. J. (1994). The Measurement of UK House Prices: a review and Aprisal of the Principlal Sources. *Journal of Housing Finance*, 24, 6-16.
8. Gibbons, S., & Machin, S. (2005). Valuing Rail Access Using Transport Innovations. *Journal of Urban Economics*, 57(1), 148–69.
9. Wilhelmsson, M. (2022). Spatial Model in Real Estate Economics. *Housing Theory and Society*, 9(2), 92–101.
10. Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers.
11. Páez, A., & Scott, M. D. (2004). Spatial statistics for urban analysis: A review of techniques with examples. *GeoJournal*, 61, 53-67.
12. Bourassa, S., Cantoni, E., & Hoesli, M. (2010). Predicting house prices with spatial dependence: a comparison of alternative methods. *Journal of Real Estate*

Research, 32(2), 139-160.

13. Peterson, S., & Flanagan, A. (2009). Neural network hedonic pricing models in mass real estate appraisal. *Journal of Real Estate Research*, 31(2), 147-164.

14. Curry, B., Morgan, P., & Silver, M. (2022). Neural networks and non-linear statistical methods: an application to the modelling of price–quality relationships. *Computers & Operations Research*, 29(8), 951-969.

15. Worzala, E., Souza, L. A., Koroleva, O., China, M., Becker, A., & Derrick, N. (2021). The technological impact on real estate investing: Robots vs humans: New applications for organizational and portfolio strategies. *Journal of Property Investment & Finance*, 39(2), 170-177.

16. Wu, H., Jiao, H., Yu, Y., Li, Z., Peng, Z., Liu, L., & Zeng, Z. (2018). Influence Factors and Regression Model of Urban Housing Prices Based on Internet Open Access Data. *Sustainability*, 10(5), 1676.

17. Lulin, X., & Li, Z. (2021). A new appraisal model of second-hand housing prices in China's first-tier cities based on machine learning algorithms. *Computational Economics*, 57(2), 617-637.

18. Truong, Q., Nguyen, M., Dang, H., & Mei, B. (2020). Housing Price Prediction via Improved Machine Learning Techniques. *Procedia Computer Science*, 174, 433–442. <https://doi.org/10.1016/j.procs.2020.06.111>

19. Federal Housing Finance Agency. House Price Index. Retrieved September 1, 2019, from <https://www.fhfa.gov/>

20. Fan, C., Cui, Z., & Zhong, X. (2018). House Prices Prediction with Machine Learning Algorithms. *Proceedings of the 2018 10th International Conference on Machine Learning and Computing - ICMLC 2018*.

<https://doi.org/10.1145/3195106.3195133>

21. Phan, T. D. (2018). Housing Price Prediction Using Machine Learning Algorithms: The Case of Melbourne City, Australia. *2018 International Conference on Machine Learning and Data Engineering (ICMLDE)*.

<https://doi.org/10.1109/icmlde.2018.00017>

22. Mu, J., Wu, F., & Zhang, A. (2014). Housing Value Forecasting Based on Machine Learning Methods. *Abstract and Applied Analysis*, 2014, 1–7. <https://doi.org/10.1155/2014/648047>
23. Lu, S., Li, Z., Qin, Z., Yang, X., & Goh, R. S. M. (2017). A hybrid regression technique for house prices prediction. 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). <https://doi.org/10.1109/ieem.2017.8289904>
24. Zhan, C., Liu, Y., Wu, Z., Zhao, M., & Chow, T. W. (2023). A hybrid machine learning framework for forecasting house price. *Expert Systems With Applications*, 233, 120981. <https://doi.org/10.1016/j.eswa.2023.120981>