

ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
КАФЕДРА ШТУЧНОГО ІНТЕЛЕКТУ

КВАЛІФІКАЦІЙНА РОБОТА

на тему: «Система прогнозування рівня попиту на житлову
нерухомість за допомогою методів машинного навчання»

на здобуття освітнього ступеня бакалавра
зі спеціальності 122 Комп'ютерні науки
(код, найменування спеціальності)
освітньо-професійної програми Штучний інтелект
(назва)

*Кваліфікаційна робота містить результати власних досліджень.
Використання ідей, результатів і текстів інших авторів мають посилання
на відповідне джерело*

(підпис)

Поліна КУРЧИНСЬКА
Ім'я, ПРІЗВИЩЕ здобувача

Виконала: здобувачка вищої освіти гр.ШІД-42
Поліна КУРЧИНСЬКА

Керівник: Ольга ЗІНЧЕНКО
Д.т.н, доц.

Рецензент:

Київ 2025

ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
Навчально-науковий інститут інформаційних технологій

Кафедра Штучного інтелекту

Ступінь вищої освіти Бакалавр

Спеціальність 122 Комп'ютерні науки

Освітньо-професійна програма Штучний інтелект

ЗАТВЕРДЖУЮ

Завідувач кафедри Штучного інтелекту

_____ Ольга ЗІНЧЕНКО

«_____» _____ 2025 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

Курчинська Поліна Валеріївна
(прізвище, ім'я, по батькові здобувача)

1. Тема кваліфікаційної роботи: Система прогнозування рівня попиту на житлову нерухомість за допомогою методів машинного навчання

керівник кваліфікаційної роботи Ольга ЗІНЧЕНКО, д.т.н, доц.
(Ім'я, ПРІЗВИЩЕ науковий ступінь, вчене звання)

затверджені наказом Державного університету інформаційно-комунікаційних технологій від «25» 02.2025р. № 56

2. Строк подання кваліфікаційної роботи «23» травня 2025р.

3. Вихідні дані до кваліфікаційної роботи: Вихідними даними для кваліфікаційної роботи є реальні ринкові дані з онлайн-платформи OLX, що включають інформацію про площу, кількість кімнат, розташування та ціну житла. Також використано програмні засоби Python, Flask, Scikit-learn та бібліотеки для вебпарсингу й візуалізації даних.

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

1. Аналіз сучасного стану ринку житлової нерухомості та факторів, що впливають на попит.
2. Огляд методів прогнозування, включаючи статистичні моделі та підходи машинного навчання.
3. Розробка архітектури інтелектуальної системи прогнозування попиту.
4. Побудова та навчання моделі машинного навчання на основі зібраного датасету.
5. Реалізація вебінтерфейсу, модуля візуалізації результатів та генерації PDF-звітів.

5. Перелік графічного матеріалу: *презентація*

6. Дата видачі завдання «25» лютого 2025 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів роботи	Примітка
1	Аналіз наявної науково-технічної літератури	25.02-05.03.25	Виконано
2	Вивчення моделей машинного навчання, що застосовуються для прогнозування	05.03-12.03.25	Виконано
3	Збір ринкових даних (OLX), попередня обробка та формування датасету	13.03-19.03.25	Виконано
4	Побудова та навчання моделі прогнозування (Linear Regression)	20.03-15.04.25	Виконано
5	Реалізація веб інтерфейсу системи (Flask, HTML/CSS, введення параметрів)	17.04-19.04.25	Виконано
6	Розробка модулів генерації PDF-звітів, збереження історії прогнозів	20.04-22.04.25	Виконано
7	Оформлення роботи: вступ, висновки, реферат	23.04-02.05.25	Виконано
8	Розробка демонстраційних матеріалів	03.05-23.05.25	Виконано

Здобувач вищої освіти

(підпис)

Поліна КУРЧИНСЬКА

(Ім'я, ПРІЗВИЩЕ)

Керівник

кваліфікаційної роботи

(підпис)

Ольга ЗІНЧЕНКО

(Ім'я, ПРІЗВИЩЕ)

РЕФЕРАТ

Текстова частина кваліфікаційної роботи на здобуття освітнього ступеня бакалавра : 64 стор., 20 рис., 1 таблиця , 20 джерел.

Мета роботи: покращення точності та ефективності процесу прогнозування попиту на житлову нерухомість шляхом розробки інтелектуальної системи, що використовує методи машинного навчання для обробки історичних даних.

Об'єкт дослідження: процес прогнозування попиту на житлову нерухомість на основі історичних ринкових даних.

Предмет дослідження: Програмна система, що реалізує прогнозування попиту на основі моделі машинного навчання, аналізує вхідні дані (ціна, площа, кількість кімнат, розташування) і формує результат у вигляді прогнозу з візуалізацією та можливістю експорту.

Короткий зміст роботи: У рамках дипломної роботи розроблено та реалізовано інтелектуальну вебсистему для прогнозування рівня попиту на житлову нерухомість за допомогою методів машинного навчання. Система побудована з використанням фреймворку Flask і мови програмування Python. Основу аналітичного ядра становить модель лінійної регресії, яка прогнозує попит та орієнтовну вартість житла на основі вхідних параметрів: площа, кількість кімнат, район. Система включає модуль автоматичного збору ринкових даних через парсер оголошень з сайту OLX, інтерфейс для введення параметрів користувачем, візуалізацію результатів і можливість генерації звітів у форматі PDF. Також реалізовано модуль перенавчання моделі для адаптації до нових даних. У результаті тестування підтверджено практичну цінність і ефективність розробленої системи для аналітики ринку нерухомості.

КЛЮЧОВІ СЛОВА: ПРОГНОЗУВАННЯ ПОПИТУ, ЖИТЛОВА НЕРУХОМІСТЬ, МАШИННЕ НАВЧАННЯ, PYTHON, FLASK, LINEAR REGRESSION, АНАЛІТИКА, ВІЗУАЛІЗАЦІЯ ДАНИХ, PDF-ЗВІТИ, OLX-ПАРСИНГ.

ЗМІСТ

ВСТУП	7
1 АНАЛІЗ ІСНУЮЧИХ РІШЕНЬ У СФЕРІ ПРОГНОЗУВАННЯ НЕРУХОМОСТІ	
1.1 Огляд ринку житлової нерухомості та факторів, що впливають на попит	9
1.2 Сучасні підходи до прогнозування попиту	13
1.3 Порівняння моделей машинного навчання для прогнозу	25
Висновки до розділу 1	30
2 ПРОЄКТУВАННЯ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ	31
2.1 Архітектура системи прогнозування	31
2.2 Побудова моделі машинного навчання (Linear Regression)	34
2.3 Реалізація збору даних через парсер OLX	38
Висновок до розділу 2	43
3 РЕАЛІЗАЦІЯ ТА ОЦІНКА ЕФЕКТИВНОСТІ СИСТЕМИ	45
3.1 Опис технологічного стеку та інтерфейсу користувача	45
3.2 Тестування системи та приклади прогнозів	51
3.3 Автоматична генерація PDF-звітів і практична цінність	55
Висновок до розділу 3	59
ВИСНОВКИ	61
ПЕРЕЛІК ПОСИЛАНЬ	63
ДЕМОНСТРАЦІЙНІ МАТЕРІАЛИ (ПРЕЗЕНТАЦІЯ)	65

ВСТУП

У сучасних умовах нестабільної економічної ситуації, урбанізації та активного розвитку будівельної галузі особливого значення набувають інтелектуальні системи, здатні автоматизувати аналіз ринку нерухомості та сприяти прийняттю обґрунтованих рішень. Однією з ключових задач у цій сфері є точне прогнозування рівня попиту на житлову нерухомість, що дозволяє забудовникам, інвесторам та агентствам нерухомості ефективно планувати діяльність, оптимізувати ціни та зменшувати ризики. Традиційні підходи до аналізу ринку нерухомості часто ґрунтуються на експертних оцінках або статистичних звітах, що не завжди дозволяє оперативно реагувати на зміни в попиті. Використання методів машинного навчання відкриває нові можливості для точного та гнучкого прогнозування на основі реальних ринкових даних — таких як площа, кількість кімнат, розташування, середні ціни, що постійно оновлюються з відкритих джерел (наприклад, OLX).

Актуальність теми зумовлена потребою в автоматизованих рішеннях для аналітики ринку житлової нерухомості, які можуть ефективно працювати з великим обсягом даних, адаптуватися до змін попиту та забезпечувати користувача зрозумілими прогнозами в реальному часі.

Метою даної кваліфікаційної роботи є підвищення точності та ефективності процесу прогнозування попиту на житлову нерухомість шляхом розробки інтелектуальної вебсистеми, що використовує алгоритми машинного навчання для аналізу історичних ринкових даних.

Об'єктом дослідження є процес прогнозування попиту на житлову нерухомість на основі факторів, що впливають на його динаміку..

Предметом дослідження виступає програмна система, яка реалізує функцію прогнозування попиту на житло із застосуванням моделі машинного навчання та інтерфейсу для користувача.

Для досягнення поставленої мети були визначені такі завдання:

- проаналізувати сучасні рішення у сфері прогнозування попиту на нерухомість;
- дослідити методи машинного навчання, що застосовуються для побудови моделей прогнозування;
- реалізувати збір і обробку ринкових даних (OLX) з формуванням датасету;
- спроектувати архітектуру системи та розробити модель прогнозування (Linear Regression);
- створити вебінтерфейс з можливістю введення параметрів і відображення результатів;
- реалізувати генерацію PDF-звітів на основі результатів прогнозу;
- провести тестування системи та оцінити її точність і зручність використання.

Методи дослідження: аналіз наукових публікацій і ринку, розробка клієнт-серверної архітектури, застосування бібліотек Python (Flask, scikit-learn, pandas, BeautifulSoup), парсинг вебданих, побудова та навчання моделей машинного навчання.

Наукова новизна полягає у поєднанні моделі прогнозування попиту з автоматичним збором даних з онлайн-ресурсів, реалізації вебінтерфейсу та механізму виведення результатів у вигляді PDF-звітів.

Практичне значення розробки полягає в можливості використання системи учасниками ринку нерухомості для об'єктивного та швидкого аналізу ринкової ситуації, що сприяє підвищенню ефективності прийняття рішень у будівельній та інвестиційній діяльності.

1 АНАЛІЗ ІСНУЮЧИХ РІШЕНЬ У СФЕРІ ПРОГНОЗУВАННЯ НЕРУХОМОСТІ

1.1 Огляд ринку житлової нерухомості та факторів, що впливають на попит

Ринок житлової нерухомості є складною багаторівневою системою, що поєднує економічні, соціальні, правові та технологічні аспекти функціонування суспільства. Його роль в економіці країни є стратегічною, оскільки він забезпечує не лише життєво необхідну потребу населення в житлі, але й активізує суміжні галузі, такі як будівництво, виробництво будматеріалів, архітектура, фінанси, страхування, юридичні послуги тощо. Функціонування цього ринку залежить від тісної взаємодії великого спектра змінних: макроекономічних показників (ВВП, інфляція, рівень безробіття), соціальних параметрів (рівень життя, міграція, сімейна структура), демографічної ситуації (кількість домогосподарств, темпи урбанізації, середній вік населення), політичної стабільності (стабільність інститутів, державна політика в галузі житлового будівництва), а також наявності фінансових механізмів підтримки населення (іпотека, державні програми доступного житла). Житлова нерухомість у країнах із розвиненою ринковою економікою відіграє подвійну роль. З одного боку — це базова потреба кожної людини у безпечному, комфортному житлі, що є ключовим компонентом гідного рівня життя. З іншого — це об'єкт довгострокових інвестицій, інструмент накопичення та збереження капіталу, який у багатьох випадках перевищує банківські депозити чи інші форми фінансових вкладень. Особливо в умовах високої інфляції, девальвації національної валюти або нестабільності фінансових ринків, інвестування в нерухомість вважається одним із найбільш надійних способів збереження вартості активів.

Житлова нерухомість є важливою складовою активів домогосподарств: на її частку припадає значний відсоток загального багатства населення. Від рівня забезпеченості населення житлом та умов проживання безпосередньо залежить соціальна стабільність, рівень задоволеності життям та ефективність міграційної

політики. У багатьох країнах саме сектор житлової нерухомості визначає архітектурне обличчя міст, формує локальні спільноти, а також значною мірою впливає на якість навколишнього середовища через стандарти енергоефективності, щільність забудови та інфраструктурні рішення. У цьому контексті прогнозування попиту на житло є не лише інструментом ринкової аналітики, а й основою для прийняття управлінських рішень забудовниками, органами державного управління, інвесторами та фінансовими установами. Висока точність такого прогнозу дозволяє забезпечити ефективний розподіл ресурсів, запобігти дефіциту або перенасиченню ринку, а також сприяє стабільному розвитку галузі.

В Україні ринок нерухомості демонструє циклічну змінність, що характеризується періодами активного зростання і спадів. До основних факторів, які впливають на попит на житлову нерухомість, належать:

Рівень доходів населення. Один із найважливіших факторів, що визначає попит на житло. У випадках, коли доходи населення зростають, підвищується купівельна спроможність, що, у свою чергу, розширює аудиторію потенційних покупців житла. З іншого боку, при зниженні реального доходу, навіть за умови стабільних цін на нерухомість, попит різко падає. Доходи також впливають на вибір типу житла (економ, комфорт чи бізнес-клас), а також на рішення про участь у кредитних програмах. Варто зазначити, що нерівномірність доходів у різних регіонах країни створює дисбаланс попиту між мегаполісами і провінціями.

Доступність іпотечного кредитування. Іпотека є головним фінансовим інструментом придбання житла для значної частини населення. Умови кредитування, зокрема розмір процентної ставки, початковий внесок, термін кредиту та валютний ризик, безпосередньо впливають на рішення громадян щодо купівлі нерухомості. Підвищення ставок або ускладнення умов доступу до кредитів (наприклад, збільшення вимог до позичальника) призводить до зниження активності на ринку. Навпаки, стимулюючі державні програми (наприклад, "Доступне житло" або програми молодіжного кредитування) можуть істотно підвищити попит, особливо серед соціально незахищених верств населення.

Політична та економічна стабільність. Довготривала стабільність у політичній та економічній сферах створює сприятливе середовище для інвестування в житло. Упевненість у незмінності курсу державної політики, стабільності валюти, відсутності загроз конфліктів чи дефолтів є вирішальними чинниками для формування попиту на довгострокові активи, такі як нерухомість. Зокрема, підвищення рівня довіри до держави стимулює громадян здійснювати великі фінансові вкладення, включаючи купівлю житла. Навпаки, політична напруга, війна, інфляційні коливання чи зміни у податковій політиці можуть призводити до скорочення попиту, відкладання купівель або масового продажу об'єктів нерухомості.

Інфраструктура та розвиток міста. Рівень розвитку інфраструктури має вирішальний вплив на привабливість житла для покупців. Райони, в яких є доступ до якісних доріг, громадського транспорту, шкіл, дитячих садків, лікарень, торгових центрів та зон відпочинку, користуються значно більшим попитом. Близькість до центрів зайнятості, університетів та адміністративних установ також підвищує цінність житла. При цьому, наявність перспектив розвитку (будівництво нової станції метро, розширення доріг, відкриття великих торгових площ тощо) здатна підвищити попит навіть у віддалених районах, перетворюючи їх на інвестиційно привабливі зони. Водночас, погана інфраструктура або її відсутність значно знижує попит, навіть якщо ціна житла є конкурентною.

Демографічні зміни. Демографічна структура населення та її динаміка безпосередньо впливають на попит на житлову нерухомість. Природне зростання населення (переважання народжуваності над смертністю) обумовлює зростання потреби у житлі для новостворених домогосподарств. Внутрішня міграція — особливо переміщення населення з сільських районів до міст — збільшує навантаження на житловий фонд великих урбаністичних центрів. Урбанізація також сприяє зростанню попиту на житло в міських агломераціях, де зосереджено більше можливостей працевлаштування, освіти та медичного обслуговування. Крім того, фактор старіння населення впливає на тип житла, що користується попитом: зростає попит на компактні квартири та доступні рішення для людей

похилого віку. Міграційні процеси, зокрема пов'язані з військовими діями або економічною кризою, формують додаткові хвилі внутрішньо переміщених осіб, яким також потрібне житло, нехай і тимчасове. Таким чином, демографічні чинники відіграють ключову роль у формуванні коротко- та довгострокового попиту на житло.

Цінова динаміка. Динаміка цін на житлову нерухомість є потужним психологічним і економічним фактором, який безпосередньо впливає на поведінку покупців і продавців. Якщо в суспільстві переважають очікування зростання цін, це стимулює швидке прийняття рішення про купівлю, оскільки потенційні покупці намагаються випередити подорожчання. Така ситуація може призвести до перегріву ринку та формування цінової бульбашки. Натомість очікування зниження цін або їх стабілізації можуть викликати ефект відкладеного попиту, коли покупці утримуються від угод у надії на вигідніші умови. Коливання цін також впливають на інвестиційну привабливість житла: стабільне зростання цін перетворює нерухомість на привабливий актив для інвесторів, тоді як нестабільність або падіння цін зменшують їхню активність. У цьому контексті важливим є також фактор інфляції, оскільки нерухомість часто розглядається як засіб збереження вартості капіталу в умовах девальвації національної валюти.

Особливістю українського ринку житлової нерухомості є значна частка угод на первинному ринку, що обумовлена активним будівництвом нових житлових комплексів у великих містах. Первинне житло часто приваблює покупців через можливість придбати квартиру в новому будинку за відносно доступною ціною на етапі котловану, а також через сучасні планування, нові інженерні мережі та енергоефективність будівель. Разом із тим, на ринку зростає попит на малогабаритні квартири (1-2 кімнатні площею до 50 м²), що пов'язано з потребами молодих сімей, студентів, внутрішньо переміщених осіб, а також з прагненням зменшити витрати на комунальні послуги в умовах зростання цін на енергоносії. Важливою рисою є стрімка цифровізація процесу пошуку та продажу житла. Сучасні онлайн-платформи — зокрема OLX, DOM.RIA, lun.ua, Flatfy, Kabanchik та інші — виступають головними каналами комунікації між

продавцями та покупцями, надаючи змогу швидко здійснювати моніторинг пропозицій, аналізувати ринок, відслідковувати динаміку цін і виявляти найвигідніші варіанти. Багато платформ також впроваджують аналітичні інструменти, інтерактивні мапи, фільтри за критеріями (площа, район, кількість кімнат), що робить їх цінним джерелом даних не лише для користувачів, а й для аналітиків ринку.

Ефективне прогнозування попиту на житлову нерухомість в Україні потребує врахування комплексу економічних, соціальних, технологічних і психологічних чинників. Особливу роль відіграє здатність системи обробляти великі обсяги розрізнених, динамічних та часом неструктурованих даних. Саме тому перспективними є підходи на основі алгоритмів машинного навчання, які дозволяють виявляти приховані закономірності в даних, враховувати нелінійні залежності між факторами та генерувати точні прогнози попиту навіть у складних умовах невизначеності.

1.2 Сучасні підходи до прогнозування попиту

Прогнозування попиту на житлову нерухомість є одним із найбільш складних і водночас критично важливих завдань для ринку нерухомості, інвесторів, девелоперів та державних органів планування. Це завдання має багатовимірний характер, оскільки вимагає врахування численних змінних, як кількісних, так і якісних, що мають тісні міжгалузеві зв'язки та часто нелінійну природу. Серед ключових змінних, які необхідно аналізувати, — економічні показники (доходи населення, ставки іпотеки, рівень зайнятості), соціальні фактори (урбанізація, демографічна структура), географічні чинники (розташування об'єктів, транспортна доступність) і психологічні аспекти (очікування покупців, рівень довіри до ринку). Традиційні методи прогнозування, що базуються на статистичному аналізі або економетричних моделях, мають обмеження в умовах швидко змінного середовища, нестабільності ринку або великого обсягу даних. Тому з розвитком цифрових технологій, широкого використання великих даних (Big Data) та алгоритмів штучного інтелекту (AI), виникає потреба у застосуванні сучасних методів, які дозволяють моделювати

складні закономірності, швидко адаптуватися до нових умов та виявляти приховані взаємозв'язки між факторами. Сучасні підходи до прогнозування попиту ґрунтуються на використанні автоматизованих систем збору та обробки даних, машинного навчання, нейронних мереж, а також інтеграції зовнішніх джерел інформації — таких як дані з вебплатформ, супутникові карти, пошукові запити, кліматичні фактори та поведінкові патерни користувачів. Умовно ці методи можна поділити на три основні категорії: класичні статистичні моделі, економетричні моделі та моделі на основі машинного навчання, з можливістю комбінування різних підходів для досягнення вищої точності та адаптивності прогнозів.

Класичні статистичні методи

Класичні методи прогнозування ґрунтуються на використанні традиційних математичних та статистичних підходів. Найпоширенішими серед них є лінійна регресія, поліноміальні регресії, метод ковзного середнього, експоненційне згладжування, а також методи сезонної декомпозиції. Лінійна регресія дозволяє побудувати залежність між незалежними змінними (наприклад, доходом населення, кількістю населення, рівнем зайнятості) та цільовим показником — попитом на житло. Такий метод добре підходить для перших етапів аналізу, коли потрібно швидко оцінити наявність базових закономірностей. Поліноміальна регресія застосовується тоді, коли зв'язок між змінними має нелінійний характер. Наприклад, попит може зростати з доходами, але лише до певного рівня, після чого стабілізується або навіть знижується. Ковзне середнє та експоненційне згладжування використовуються для аналізу часових рядів і виявлення загальних трендів. Вони дозволяють приглушити випадкові коливання та виділити основну тенденцію, що є корисним для короткострокових прогнозів.

Попри свою простоту, ці моделі мають низку обмежень. Вони не завжди здатні коректно врахувати складні багатофакторні залежності, сезонність або зовнішні шоки. Крім того, їх точність значно знижується в умовах нестабільного ринку та високої волатильності даних, що є характерним для сучасного ринку нерухомості.

Економетричні моделі

Економетричні моделі є більш гнучкими та математично обґрунтованими інструментами аналізу залежностей у часових рядах. Серед найвідоміших — моделі *ARIMA* (AutoRegressive Integrated Moving Average), *VAR* (Vector AutoRegression), *GARCH* (Generalized Autoregressive Conditional Heteroskedasticity) тощо.

ARIMA дозволяє моделювати залежності у часових рядах із урахуванням автокореляції, сезонності та трендів. Застосовується для прогнозування змін у попиті на основі історичних даних.

VAR застосовується тоді, коли кілька змінних мають взаємозв'язки між собою. Наприклад, одночасне прогнозування впливу зміни доходів, вартості будівництва та процентних ставок на попит на житло.

GARCH моделі використовуються в ситуаціях, коли волатильність змінна в часі (наприклад, різке коливання цін на нерухомість у періоди криз чи активного росту).

Економетричні моделі вимагають наявності якісних часових рядів і ретельної підготовки даних. Вони забезпечують глибший рівень аналізу, але водночас обмежені через необхідність дотримання припущень щодо нормальності розподілу, стаціонарності рядів тощо. Їх використання є доцільним у прогнозуванні середньо- та довгострокових трендів.

Методи машинного навчання (ML)

Останнім часом найбільш перспективними вважаються моделі ML, що дозволяють працювати з великими обсягами даних та виявляти складні, нелінійні залежності. Серед них найбільш поширені:

- Лінійна регресія (Linear Regression) — базова модель, яка добре підходить для першого етапу аналізу.
- Древа рішень і Random Forest — дозволяють враховувати взаємодії між змінними та мають високу інтерпретованість.
- Градієнтний бустинг (XGBoost, LightGBM) — забезпечують високу точність, особливо для структурованих табличних даних.
- Штучні нейронні мережі — здатні моделювати складні залежності, але потребують великої кількості даних і обчислювальних ресурсів.

Комбінації моделей (Ensemble Learning)

Ансамблеві методи прогнозування базуються на ідеї, що комбінація кількох моделей забезпечує кращу узагальнюючу здатність, ніж будь-яка окрема модель. Цей підхід дозволяє компенсувати недоліки окремих моделей, знижуючи ймовірність переобучення, покращуючи стабільність та точність прогнозів. Існує декілька основних типів ансамблевих методів.

Беггінг (Bagging), бустинг (Boosting) та стекинг (Stacking) — це потужні ансамблеві методи машинного навчання, що активно використовуються у задачах прогнозування попиту, зокрема у сфері ринку житлової нерухомості. Їхня головна перевага полягає в об'єднанні декількох моделей для досягнення більшої узагальнювальної здатності та підвищення точності. На відміну від класичних моделей, ансамблеві методи дозволяють враховувати складні взаємозв'язки між ознаками, зменшувати ризик переобучення та підвищувати стійкість результатів до випадкових коливань у даних. Такі підходи дозволяють створювати більш адаптивні системи, здатні ефективно працювати з великими обсягами інформації, високою варіативністю ознак і неоднорідними вибірками. Їхня ефективність доведена не лише теоретичними обґрунтуваннями, а й практичним використанням у галузях фінансів, маркетингу, медицини, енергетики та, зокрема, у

прогнозуванні цін на нерухомість, де помилка навіть у декілька відсотків може мати суттєві економічні наслідки.

Нижче подано деталізовану характеристику кожного з цих підходів, супроводжену прикладами ілюстрацій, які доцільно додати у вигляді графіків, схем або діаграм для кращого розуміння і візуалізації принципів їхньої роботи.

Беггінг (Bagging) (рис. 1.1) — це метод ансамблювання, при якому багато екземплярів однієї й тієї ж базової моделі (як правило, дерева рішень) тренуються незалежно одна від одної на різних підмножинах вихідного набору даних, що формуються методом випадкової вибірки із заміною (bootstrapping). Кожна модель отримує свою унікальну вибірку, що забезпечує варіативність у навчанні та дозволяє уникнути переобучення. Після навчання прогноз кожної моделі враховується в загальному рішенні шляхом:

- усереднення (для регресійних задач), що знижує дисперсію,
- голосування більшості (для задач класифікації), що підвищує стійкість до помилок окремих моделей.

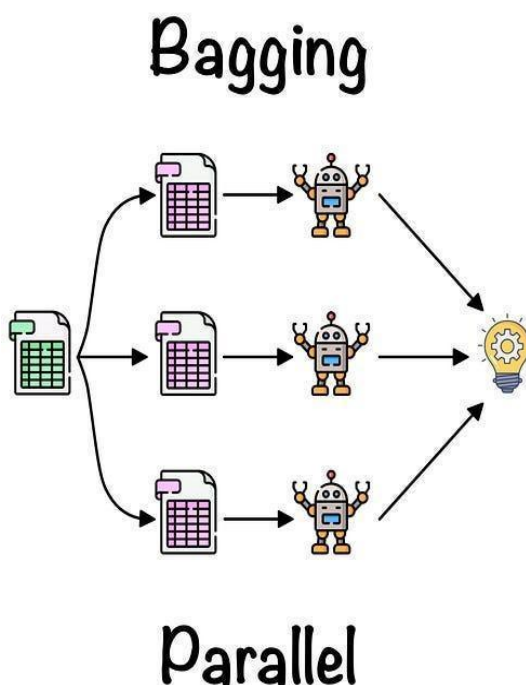


Рисунок 1.1 – Беггінг метод

Переваги беггінгу включають:

- зменшення варіативності моделей,
- підвищення узагальнюючої здатності без значного зростання упередженості,
- паралельність обчислень.

Найвідоміший приклад реалізації беггінгу — Random Forest. Він використовує велику кількість дерев рішень, що навчаються на випадкових підмножинах даних і наборів ознак, що ще більше підвищує диверсифікацію моделей.

Бустинг (Boosting) (рис. 1.2) — це потужна техніка ансамблевого навчання, яка працює за принципом послідовного вдосконалення моделі. Ключова ідея полягає у тому, що кожна нова модель навчається на прикладах, які були неправильно передбачені попередніми моделями, тим самим фокусуючись на складних випадках.

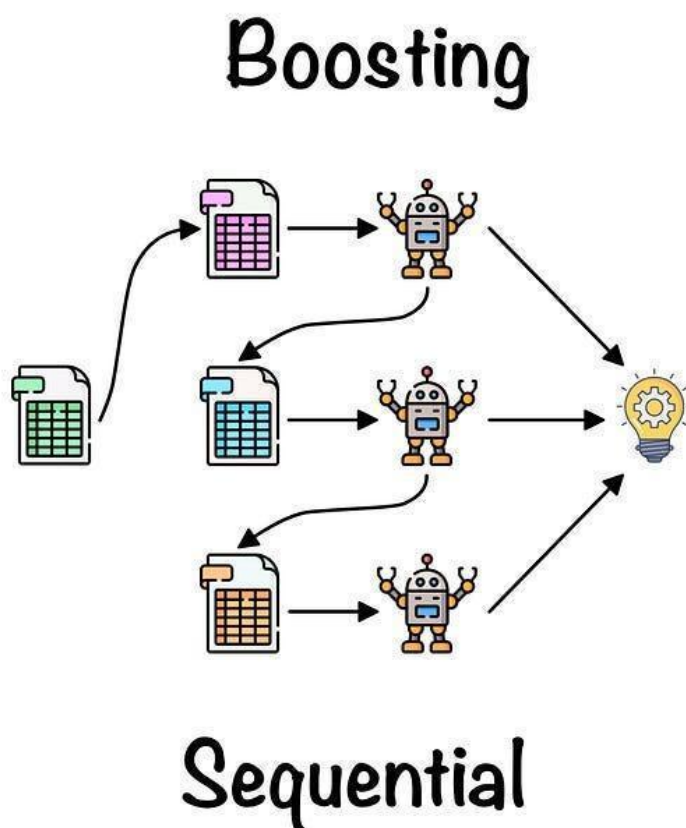


Рисунок 1.2 – Бустинг метод

Така стратегія дозволяє покращувати загальну точність ансамблю, оскільки з кожним кроком зменшується кількість помилок у прогнозі. Цей процес продовжується доти, доки не буде досягнута бажана точність або покращення не припиниться. Бустинг може бути реалізований у вигляді градієнтного бустингу, де кожна нова модель намагається наблизити залишки (різницю між справжніми значеннями та передбаченими попередніми моделями). Це дозволяє створювати дуже точні предиктори, які адаптуються до складної структури даних.

Серед найбільш ефективних і поширених реалізацій бустингу виділяють:

- XGBoost (Extreme Gradient Boosting) — високопродуктивний алгоритм, який забезпечує баланс між швидкістю і точністю, часто займає перші місця у змаганнях з машинного навчання;
- LightGBM — розроблений компанією Microsoft, особливо ефективний при роботі з великими наборами даних і високою кількістю ознак, завдяки використанню градієнтного бустингу на деревоподібних структурах з пріоритетом глибини;
- CatBoost — розроблений компанією Яндекс, має вбудовану підтримку категоріальних ознак без потреби їх попереднього кодування, що робить його дуже зручним у реальних бізнес-задачах.

Стекинг (Stacking) (рис. 1.3) — це потужна техніка ансамблевого навчання, яка полягає у побудові багаторівневої архітектури моделей. На першому рівні одночасно навчаються кілька базових моделей, які можуть бути різного типу (наприклад, логістична регресія, дерево рішень, нейронна мережа, SVM тощо). Ці моделі формують свої прогнози, які потім подаються як вхідні ознаки до мета-моделі (другого рівня), що здійснює остаточне передбачення. Мета-модель може бути будь-якою — від простої лінійної регресії до бустингового алгоритму. Такий підхід дозволяє використовувати сильні сторони кожної моделі та компенсувати їхні слабкі сторони, що забезпечує більш високу точність і стабільність результатів у порівнянні з використанням окремих моделей.

До переваг стекингу належать:

- можливість врахування гетерогенності моделей;

- покращення узагальнюючої здатності системи;
- підвищення стійкості до шумів у даних;
- потенціал для роботи з великими і складними наборами даних.

Для візуалізації структури стекинг доцільно використовувати діаграму, на якій зображено кілька моделей першого рівня, що генерують вихідні значення, які потім передаються до мета-рівня для остаточного прогнозу.

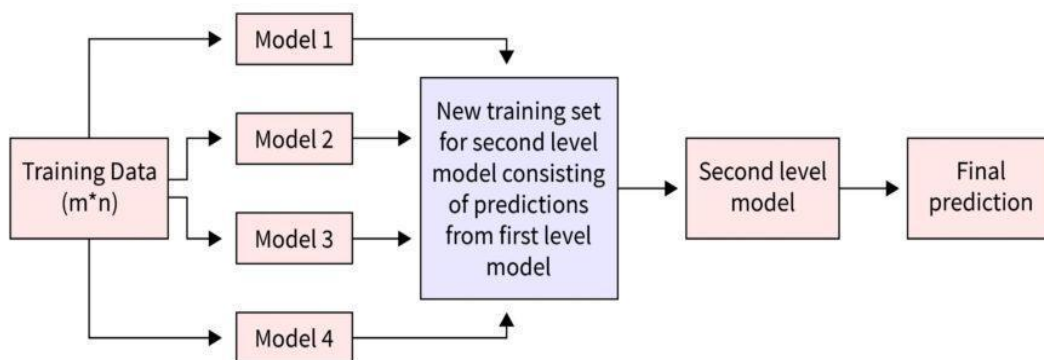


Рисунок 1.3 – Стекинг метод

Використання ансамблевих методів дає змогу:

- зменшити ризик переобучення,
- покращити стабільність результату,
- врахувати різноманітні особливості даних,
- підвищити точність прогнозу навіть у випадках складної залежності між ознаками.

На практиці часто використовують поєднання цих методів, наприклад, стекинг моделей бустингу і регресії або об'єднання Random Forest та XGBoost для оптимального результату в задачах прогнозування попиту на нерухомість. Ансамблі є особливо корисними для задач із високою варіативністю даних, наявністю складних закономірностей і значним ризиком перенавчання. У сфері

нерухомості це дозволяє враховувати вплив десятків факторів, включаючи ринкові, соціальні, фінансові та поведінкові змінні.

Інтеграція зовнішніх даних та Big Data

Використання великих обсягів зовнішніх даних (Big Data) є одним з найактуальніших напрямів розвитку прогностичних систем. Зовнішні джерела можуть включати:

- Google Trends — показники популярності запитів користувачів у Google можуть сигналізувати про інтерес до купівлі нерухомості в певному регіоні.
- Соціальні мережі — аналіз текстів, хештегів, геолокацій дозволяє виявляти настрої та локальні тренди.
- Геоаналітика — дані про мобільність населення, трафік, близькість до інфраструктури допомагають оцінити привабливість району.
- Метеодані та екологічна інформація — впливають на прийняття рішень щодо вибору житла.
- Будівельна активність — дані про кількість виданих дозволів на будівництво, нові об'єкти та введення в експлуатацію.

Інтеграція таких джерел у прогностичні моделі суттєво підвищує їхню адаптивність і здатність реагувати на зміни у поведінці споживачів та ситуації на ринку в реальному часі. Для цього застосовуються алгоритми обробки неструктурованих даних, текстовий аналіз (NLP), методи кластеризації та глибокого навчання. Використання зовнішніх даних і ансамблевих моделей дозволяє створювати прогностичні системи нового покоління, здатні працювати в умовах складності, невизначеності та високої динаміки ринку житлової нерухомості.

У контексті України, де офіційна статистика щодо ринку житлової нерухомості є фрагментованою, застарілою або недостатньо деталізованою, особливої важливості набувають альтернативні джерела даних. Найбільш цінну інформацію можна отримати з онлайн-платформ оголошень — таких як OLX(рис.1.4), DOM.RIA (рис. 1.5), lun.ua (рис. 1.6) та інші, де щоденно публікуються тисячі нових пропозицій щодо продажу чи оренди житла.

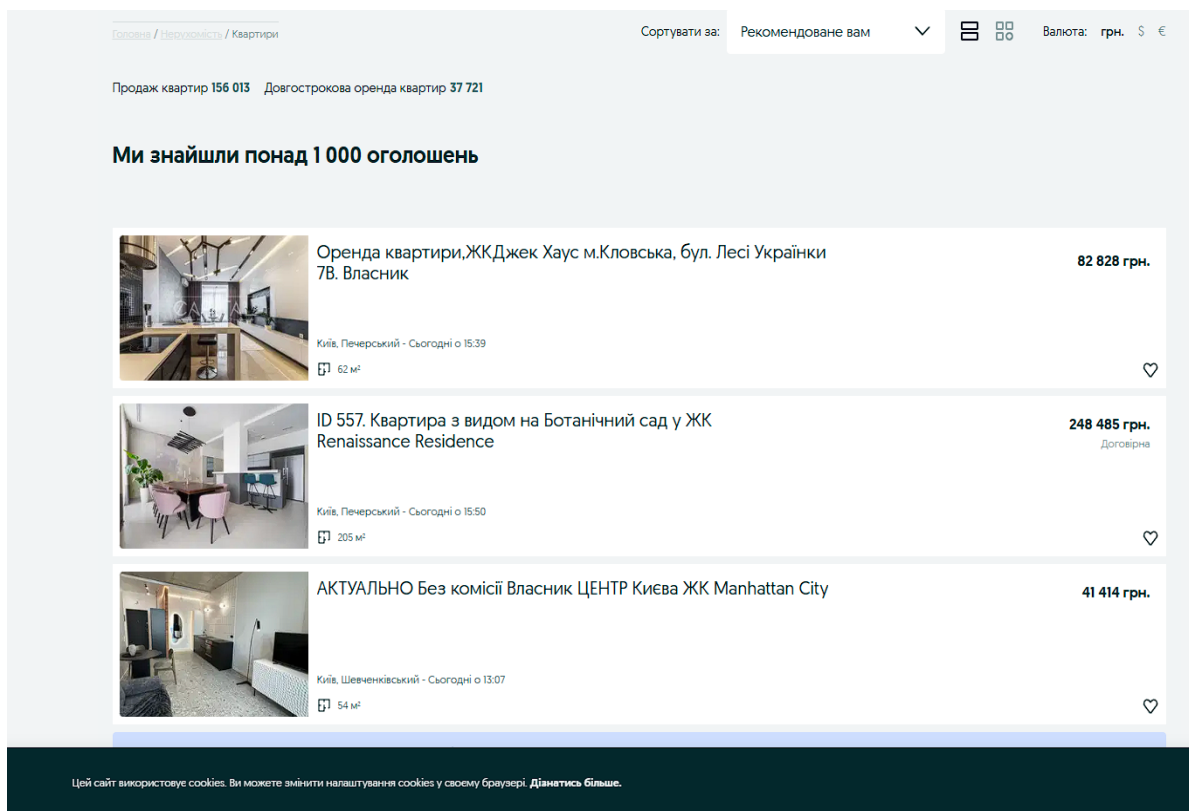


Рисунок 1.4 – OLX дошка оголошень

Саме ці джерела містять реальні поточні дані щодо цін, розташування, характеристик квартир (площа, кількість кімнат, поверховість) та навіть поведінкових аспектів — таких як динаміка розміщення оголошень, кількість переглядів тощо.

У межах даної дипломної роботи обрано підхід, що поєднує автоматизований парсинг оголошень із побудовою структурованого датасету, який згодом використовується для навчання моделей машинного навчання. Такий підхід дозволяє формувати живу базу даних, яка оперативно оновлюється, а також забезпечує точнішу картину поточної ринкової ситуації.

Довгострокова оренда квартир в Україні
14 649 пропозицій

Спочатку перевірені Підписатися на пошук Дивитися на мапі

145 565 грн
вул. Ярославів Вал, 20
Золоті ворота · Шевченківський · Київ
4 кімнати · 145 м² · 4 поверх з 5 · Без тварин
У центрі · Новий ремонт · Поруч з парком
БЕЗ КОМІСІЇ, пропозиція ВІД ВЛАСНИКА Двоповерхова квартира в центрі Києва розташована на 4-му і 5-му поверхах п'ятиповерхового цегляного будинку 1934 року побудови. 3 хвилини пішки до метро Золоті Ворота, 10 хвилин до М.Хрещатик або М.Театральна. Велика кількість кафе та ресторанів на будь-які...
Опубліковано 2 тра · Перевірено 2 тра.
Ще 1 пропозиція цього об'єкта →

15 000 грн
вул. Бориса Мартоса(Плеханова), 6
Лівобережна · Дніпровський · Київ
1 кімната · 35 м² · 3 поверх з 9
Вигідна ціна · Спокійний район · Тихий двір

Рисунок 1.5 – Платформа DOM.RIA

Перевагою є також гнучкість у виборі змінних, можливість попередньої фільтрації та очищення даних, а також масштабування на різні регіони України або різні сегменти ринку (економ, комфорт, бізнес-клас). Враховуючи високу мінливість ринку та обмеження офіційних джерел, саме машинне навчання на базі відкритих ринкових даних є найоптимальнішим інструментом для побудови гнучкої та адаптивної системи прогнозування попиту.

лун Київ Головна Новобудови Продаж Оренда Мій ЛУН

Пошук Фільтри

Оренда квартир Київ (довгостроково)
11 902 оголошення по 6 353 об'єктам За замовчуванням

27 000 грн
вулиця Панаса Мирного, 17
Печерський · Печерський · Київ
Унікальна 1-кімнатна квартира в самому центрі – вул. Мирного, 17 (50 кв.м) Квартира з...
1 кімната · з ремонтом · цегляна
поверх 7 з 25
сьогодні о 09:03 знайдено сьогодні...

49 988 грн
вулиця Панаса Мирного, 28-А
ЖК Корона Печерська · Печерський · Київ
Простора 3-кімнатна квартира 150 м² – вул. Мирного, 28А, 19 поверх Вигідна пропозиція в...
3 кімнати · з ремонтом · монолітно-карка...
поверх 19 з 25 · 2005 рік будівни...
сьогодні о 09:33 знайдено сьогодні...

Задайте ціну для пошуку

Рисунок 1.6 – Платформа ЛУН

1.3 Порівняння моделей машинного навчання для прогнозу

Для вибору найбільш ефективної моделі прогнозування попиту на житлову нерухомість необхідно здійснити глибокий, багаторівневий порівняльний аналіз найпоширеніших і найрезультативніших алгоритмів машинного навчання. Ринок нерухомості є надзвичайно складним об'єктом для моделювання через високу варіативність, залежність від зовнішніх факторів (економічних, соціальних, політичних), наявність великої кількості шуму в даних, а також обмежену доступність офіційної статистики. У таких умовах класичні статистичні підходи не здатні забезпечити потрібну точність і адаптивність. З урахуванням складності, багатфакторності та нестабільності ринку нерухомості, ефективна модель повинна володіти низкою важливих характеристик: по-перше, вміти працювати з великим обсягом різноманітних вхідних даних (кількісних, категоріальних, текстових); по-друге, виявляти складні, в тому числі нелінійні, взаємозв'язки між ознаками; по-третє, бути здатною до самоадаптації — перенавчання у разі оновлення даних або зміни тенденцій на ринку; і, нарешті, мати достатню стійкість до пропущених значень, аномалій та структурного шуму, що властиві реальним ринковим даним. Важливим критерієм є також прозорість та інтерпретованість результатів. У задачах, що стосуються фінансів, інвестицій або регуляторного аналізу, вкрай важливо не лише отримати прогноз, але й розуміти, як модель дійшла до певного висновку. Це дозволяє приймати більш обґрунтовані управлінські рішення, проводити валідацію моделі, а також налагодити комунікацію з кінцевими користувачами, які можуть не мати спеціалізованих знань у сфері штучного інтелекту.

У сучасних умовах завдання прогнозування попиту на нерухомість потребує мультидисциплінарного підходу: поєднання гнучких і точних інструментів машинного навчання з інтерпретованими, економічно обґрунтованими підходами до аналізу даних. Це дозволяє створити ефективні системи підтримки прийняття рішень у сфері ринку житла, здатні адаптуватися до викликів сучасності.

У межах даної роботи були відібрані і проаналізовані наступні моделі машинного навчання, які на сьогодні широко використовуються в задачах регресійного прогнозування:

- Лінійна регресія (Linear Regression) — базова модель, що використовується для виявлення простих лінійних залежностей між ознаками і цільовою змінною.
- Дерево рішень (Decision Tree) — алгоритм, що дозволяє будувати прогнози на основі ієрархії умовних правил, що робить модель простою для інтерпретації.
- Ансамблевий метод Random Forest — поєднання багатьох дерев рішень, яке підвищує точність та зменшує ризик перенавчання.
- Градієнтний бустинг (XGBoost) — один із найпотужніших алгоритмів підвищення точності, особливо при роботі з табличними даними.
- Штучна нейронна мережа (ANN) — багаторівнева модель, здатна до самонавчання і виявлення складних нелінійних зв'язків між змінними.

Дані моделі були обрані не лише за популярністю, наявністю готових реалізацій у популярних бібліотеках машинного навчання та широким практичним застосуванням, а й з огляду на їх різні концептуальні підходи до обробки вхідних даних. Такий різноплановий вибір дозволяє комплексно оцінити їхні сильні та слабкі сторони в умовах прогнозування попиту на житлову нерухомість, де дані можуть мати високу варіативність, пропущені значення, неструктуровані джерела та неоднорідну природу (наприклад, текстові описи квартир, числові показники площі, категоріальні дані щодо районів). Під час аналізу ефективності моделей було враховано не лише базову точність прогнозування, яка вимірюється за допомогою таких метрик, як середньоквадратична похибка (RMSE), середня абсолютна похибка (MAE) або коефіцієнт детермінації (R^2), але й низка додаткових критично важливих аспектів. Зокрема, одним із ключових параметрів є швидкість навчання, яка визначає, наскільки швидко модель здатна адаптуватися до нових даних, що є важливо для щоденного оновлення прогнозів у реальному часі.

Здатність до узагальнення є ще одним визначальним фактором: хороша модель має демонструвати не лише високу точність на навчальному наборі, а й стабільні результати на нових, невідомих даних, без ознак перенавчання. Для цього моделі тестувалися на незалежних підвибірках із різних регіонів, щоб оцінити універсальність побудованих прогнозів. Стійкість до шуму — важлива властивість для систем, що працюють із реальними, часто неповними або некоректними даними. Моделі, що надмірно чутливі до невеликих змін у вхідних параметрах, можуть давати нестабільні результати й призводити до помилкових управлінських рішень. Інтерпретованість результатів є критичною, особливо для фінансових систем або при роботі з державними структурами, де необхідно не лише надати числовий прогноз, але й пояснити його причини: які змінні мали найбільший вплив, як зміна певного параметра може вплинути на результат.

Враховувались такі критерії, як гнучкість моделі до перенавчання у разі надходження нових даних, що дозволяє постійно адаптувати систему до змін на ринку, та масштабованість — здатність ефективно працювати з великими обсягами даних, у тому числі за умови збільшення кількості об'єктів чи географічного охоплення. Було сформовано комплексну методику оцінювання моделей, що охоплює як технічні характеристики, так і прикладні аспекти їх впровадження в реальних умовах. Наступним логічним кроком стало проведення порівняльного аналізу зазначених моделей за основними метриками ефективності. Це дозволяє зробити не лише технічно обґрунтований, а й практично доцільний вибір найбільш ефективного алгоритму для побудови системи прогнозування попиту на житлову нерухомість.

Порівняння (табл. 1.1) проводилося за такими основними критеріями:

- Точність (accuracy / RMSE / MAE);
- Швидкість навчання;
- Інтерпретованість результатів;
- Масштабованість та гнучкість моделі;
- Стійкість до перенавчання.

Таблиця 1.1 – Порівняння ШІ-моделей

Модель	Точність	Швидкість	Інтерпретованість	Масштабованість	Перенавчання
Linear Regression	Середня	Висока	Висока	Середня	Низьке
Decision Tree	Середня	Висока	Середня	Середня	Високе
Random Forest	Висока	Середня	Середня	Висока	Низьке
XGBoost	Висока	Середня	Низька	Висока	Середнє
Neural Network (ANN)	Висока	Низька	Низька	Висока	Середнє

На основі наведеного порівняння можна зробити комплексний висновок про переваги та недоліки різних моделей у задачі прогнозування попиту на житлову нерухомість. Моделі Random Forest і XGBoost продемонстрували найвищі показники точності, що зумовлено їхньою здатністю враховувати складні взаємозв'язки між ознаками, обробляти велику кількість параметрів та виявляти глибинні закономірності у структурованих табличних даних. Вони виявились найбільш стабільними у ситуаціях з великою кількістю шуму або частково відсутніми даними, а також забезпечили кращу узагальнювальну здатність при тестуванні на нових прикладах. Однак ці моделі вимагають значних обчислювальних ресурсів, особливо при масштабуванні на великі обсяги даних, і складніші у налаштуванні та поясненні результатів. У той же час лінійна регресія зберігає свої позиції як інструмент для швидкої побудови прогнозів на базі лінійних залежностей. Вона вирізняється високою швидкістю навчання, простотою реалізації та повною прозорістю інтерпретації — користувачі можуть чітко побачити, як кожен окремий фактор впливає на результат. Це робить її ідеальним варіантом для базової аналітики або використання в системах, де важлива пояснюваність, наприклад, у звітності для регуляторів або фінансових установ.

Штучні нейронні мережі (ANN) є надзвичайно гнучкими інструментами, що дозволяють ефективно моделювати складні, нелінійні взаємозв'язки між ознаками у великих обсягах даних. Їхня архітектура з багатьма прихованими шарами і

функціями активації дозволяє системі самостійно виявляти важливі характеристики з вхідного набору, що робить їх особливо корисними в умовах, коли структура взаємозв'язків між змінними невідома або надто складна для традиційних алгоритмів. ANN добре підходять для задач, де необхідно обробляти не лише числові, а й текстові або зображувальні дані, що відкриває нові перспективи для розширення систем прогнозування. Водночас одним із основних недоліків нейронних мереж є їхня слабка інтерпретованість. Користувач без технічної підготовки часто не може зрозуміти, чому модель надала той чи інший прогноз. Це створює труднощі у випадках, коли необхідна прозорість рішень — наприклад, при фінансовій або юридичній відповідальності за прийняті на основі моделі дії. Крім того, навчання нейронної мережі є обчислювально затратним процесом, що потребує потужного апаратного забезпечення та великої кількості даних.

Вибір моделі залежить від конкретних вимог і контексту застосування: якщо на перше місце ставиться прозорість і швидкість — доцільніше використовувати лінійну регресію або дерево рішень; якщо ж ключовими є точність, здатність до узагальнення, обробка великих та складних датасетів — найбільш ефективними будуть XGBoost або нейронні мережі.

Висновки до розділу 1

У результаті аналізу, проведеного в межах першого розділу, було сформовано цілісне уявлення про особливості функціонування ринку житлової нерухомості та сучасні підходи до прогнозування попиту на ньому. Зокрема, встановлено, що попит на житло формується під впливом великої кількості різноманітних чинників: від макроекономічних показників та рівня доходів населення — до демографічних трендів, доступності іпотеки, політичної стабільності та розвитку інфраструктури. Такі фактори діють у складній взаємодії, що зумовлює багатofакторність і динамічність ринку.

Огляд сучасних методів прогнозування продемонстрував, що класичні статистичні та економетричні підходи можуть бути корисними на етапах попереднього аналізу даних, але значно поступаються моделям машинного

навчання у точності, адаптивності та здатності враховувати складні, нелінійні залежності. Особливо перспективними виявилися моделі на основі ансамблевих методів (Random Forest, XGBoost) і нейронних мереж (ANN), які здатні працювати з великими обсягами неоднорідних даних і формувати більш точні прогнози.

Було також визначено критерії, за якими здійснюється оцінка ефективності моделей: точність прогнозування, швидкість навчання, стійкість до шуму, здатність до узагальнення, інтерпретованість результатів та масштабованість. На підставі порівняння кількох популярних моделей встановлено, що XGBoost та Random Forest найкраще відповідають вимогам завдання прогнозування попиту на житло, тоді як лінійна регресія зберігає свою цінність для базової інтерпретованої аналітики.

Результати першого розділу лягли в основу вибору інструментів і методів, які були використані в наступних етапах роботи — під час розробки архітектури системи, побудови моделі машинного навчання та її практичної реалізації в межах дипломного проєкту.

2 ПРОЄКТУВАННЯ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ

2.1 Архітектура системи прогнозування

Розроблена система прогнозування попиту на житлову нерухомість базується на багатоетапній архітектурі клієнт-серверного типу з повноцінною інтеграцією алгоритмів машинного навчання, автоматизованого збору даних та генерації вивідної аналітики. Архітектура системи є ключовим фактором її функціональності, надійності та гнучкості, оскільки вона визначає логіку взаємодії між усіма основними компонентами: інтерфейсом користувача, модулем обробки даних, серверною частиною, машинним навчанням, базою даних і модулем виводу результатів у вигляді PDF-звітів. Основна мета архітектурного проєктування полягає у створенні високонадійної, масштабованої та адаптивної структури, яка забезпечить стабільну роботу системи в режимі реального часу. Це включає можливість динамічного оновлення вхідних даних, безперервної роботи користувацького інтерфейсу при оновленні моделі, швидкої адаптації до змін ринку нерухомості та інтеграції нових функціональностей, таких як панель адміністратора, особистий кабінет, аналітичний дашборд, REST API-доступ, геоаналітика або інструменти візуалізації попиту.

Інформаційна архітектура системи дотримується принципу багаторівневого розділення відповідальності: клієнтський рівень відповідає за збір параметрів користувача через вебінтерфейс, передає запити на серверну частину, яка далі ініціює обробку, звертається до моделі машинного навчання, проводить прогноз, формує відповідь, фіксує її в базі даних і повертає клієнту результат з можливістю експорту у PDF. Усі ключові процеси (рис. 2.1), включно з прогнозом, формуванням PDF і збереженням результатів, реалізовані асинхронно, що суттєво покращує продуктивність, знижує затримки та забезпечує плавну взаємодію навіть під високим навантаженням.

Такий підхід також дозволяє у майбутньому швидко оновлювати модель без простою сервісу, забезпечуючи постійну актуальність прогнозів на основі останніх змін на ринку нерухомості.

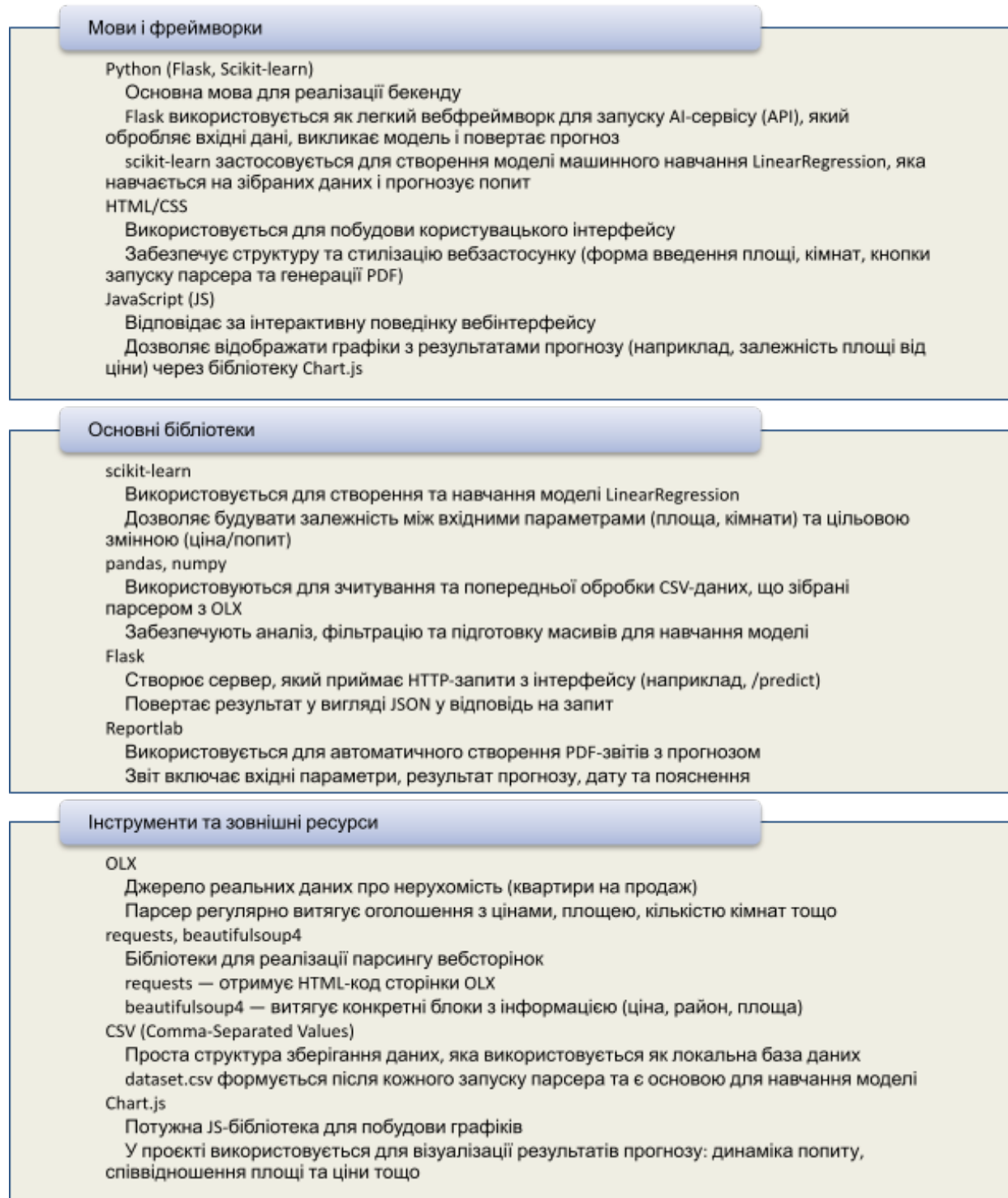


Рисунок 2.1 – Використані технології

Система складатиметься з таких основних компонентів:

Інтерфейс користувача (Front-end) — реалізований за допомогою HTML/CSS з можливістю введення вхідних параметрів (площа, кількість кімнат,

район тощо). Інтерфейс є адаптивним та зручним для використання на різних пристроях.

Серверна логіка (Back-end) — побудована на Flask (Python), відповідає за прийом даних, взаємодію з моделлю ML, підготовку результатів та генерацію PDF-звіту.

Модуль прогнозування (ML-модель) — реалізований на базі моделі Linear Regression із можливістю подальшого розширення до XGBoost або нейронної мережі. Навчання моделі відбувається на основі попередньо зібраного та очищеного датасету.

Парсер оголошень — окремий модуль, що автоматично збирає дані з OLX за допомогою бібліотек BeautifulSoup та Requests. Отримані дані додаються до датасету для підвищення актуальності прогнозів.

База даних — містить історію запитів користувачів, оновлений датасет, результати прогнозів. Реалізована за допомогою SQLite.

Генератор PDF-звітів — дозволяє формувати результати прогнозу у форматі PDF із зазначенням вхідних параметрів, отриманої ціни та дати генерації. Реалізовано через бібліотеку ReportLab.

Архітектура системи побудована таким чином, щоб забезпечити максимальну гнучкість, масштабованість і розширюваність у майбутньому. Зокрема, враховано можливість простого додавання нових функціональних модулів без необхідності змінювати основну структуру системи. Наприклад, реалізація REST API дозволяє легко підключати зовнішні джерела даних, такі як агрегатори оголошень, державні реєстри або аналітичні платформи. Вбудована підтримка повторного навчання моделі дозволяє без зупинки роботи системи адаптуватися до нових тенденцій ринку, оновлювати коефіцієнти моделі або повністю замінювати алгоритм на більш точний — наприклад, XGBoost чи багатошарову нейронну мережу.

Значну увагу приділено зручності для кінцевого користувача — вебінтерфейс є інтуїтивно зрозумілим, не вимагає реєстрації чи авторизації, що дозволяє отримати прогноз миттєво після введення вхідних параметрів. Це

особливо важливо для користувачів, які бажають швидко оцінити приблизну вартість нерухомості без необхідності створювати облікові записи. У той же час, архітектура побудована з урахуванням подальшого розвитку: система легко адаптується для інтеграції особистих кабінетів, багатомовного інтерфейсу, адмінпанелі, модулів персоналізованих рекомендацій, а також розширених візуальних аналітик з графіками, побудованими на основі історичних і прогнозованих даних. Така масштабованість дає змогу залучити широку аудиторію користувачів: від приватних осіб, які планують купівлю чи продаж квартири, до професійних учасників ринку — забудовників, агентств нерухомості та державних структур, які потребують оперативного доступу до достовірних прогнозів вартості житла у реальному часі. Крім того, відкрита архітектура дає змогу підключати додаткові модулі, зокрема — геолокаційний пошук, карти попиту, динамічні дашборди з фільтрами та інші сучасні інструменти ринкового аналізу.

2.2 Побудова моделі машинного навчання (Linear Regression)

Модель лінійної регресії була обрана як базова модель для реалізації прогнозування попиту або вартості житла в межах даної інтелектуальної системи, з огляду на її низку ключових переваг. Насамперед, ця модель є однією з найпростіших і найбільш зрозумілих у сфері машинного навчання, що дає змогу чітко інтерпретувати результати: кожен коефіцієнт у рівнянні моделі відображає вплив відповідної ознаки (наприклад, площі, кількості кімнат чи району) на цільову змінну. Завдяки цьому модель чудово підходить для використання у практичних задачах, де важлива прозорість та обґрунтованість прийнятих рішень. Лінійна регресія характеризується дуже швидкою швидкістю навчання та прогнозування, що дозволяє інтегрувати її у веборієнтовані системи без помітних затримок у роботі. Вона не потребує значних обчислювальних ресурсів і підходить навіть для обмежених середовищ (наприклад, VPS-серверів з 1–2 ГБ ОЗП), що особливо важливо для розгортання MVP (minimum viable product) — мінімального робочого прототипу.

Також важливо підкреслити, що модель лінійної регресії демонструє високу ефективність у випадках, коли залежність між незалежними змінними та цільовим параметром або прямо лінійна, або наближено має лінійний характер. У сфері ринку житлової нерухомості це проявляється у прямому впливі таких характеристик, як площа, кількість кімнат, район тощо, на вартість об'єкта. Навіть без використання складних моделей із багат шаровими структурами, лінійна регресія здатна якісно виявляти базові тренди, тренувальні залежності та логічні взаємозв'язки між змінними. Завдяки цьому вона забезпечує високу ступінь пояснюваності прогнозів, що є надзвичайно важливим для прийняття обґрунтованих управлінських рішень. Інтеграція лінійної регресії як стартової моделі в системі прогнозування дозволяє досягти балансу між швидкістю розробки, технічною простотою, економністю в обчисленнях та високим рівнем інтерпретованості результатів. Це дає змогу створити мінімально життєздатний продукт (MVP), який уже може надавати корисну аналітику, підтримувати прийняття рішень на основі даних та забезпечити гнучку платформу для подальшого розширення й ускладнення моделі відповідно до зростаючих потреб ринку та користувачів. Модель (рис. 2.2) прогнозування описана нижче:

Вхідні дані: для побудови та навчання моделі використовується CSV-файл, який формується автоматизованим парсером із платформи оголошень OLX. Парсер реалізовано за допомогою бібліотек BeautifulSoup та Requests у Python, що дозволяє регулярно та без втручання людини збирати актуальні дані з великої кількості оголошень у різних регіонах. Зібрана інформація включає ключові характеристики об'єктів нерухомості, що продаються: площу, кількість кімнат, район розташування, поверх, загальний стан, тип будівлі та вартість.

Ознаки (X): до множини предикторів входять кількісні змінні (наприклад, площа квартири у квадратних метрах, кількість кімнат, поверх) та категоріальні змінні (район, тип будинку), які перетворюються у числовий формат методом one-hot encoding для забезпечення коректної роботи моделі. Така обробка дозволяє моделі працювати як з безперервними, так і з дискретними ознаками, що значно підвищує точність прогнозування.

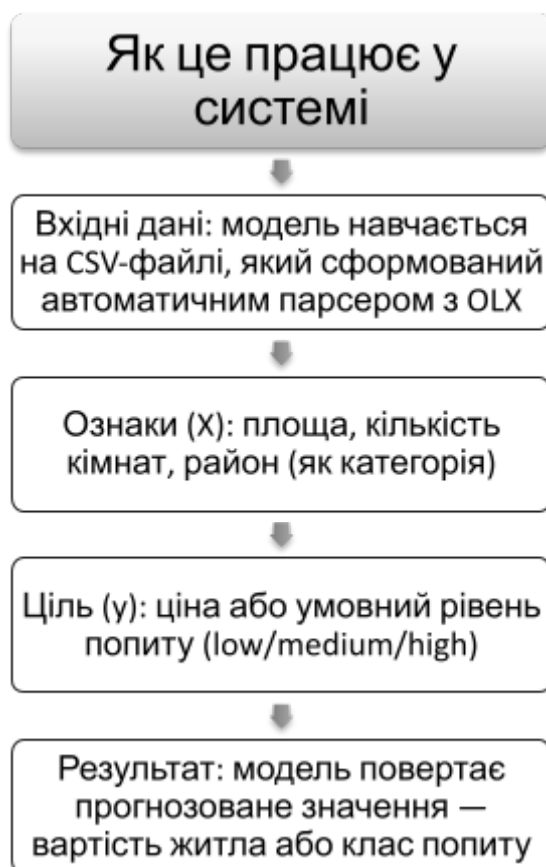


Рисунок 2.2 – Модель прогнозування

Цільова змінна (y): очікувана ціна (у гривнях) або рівень попиту, класифікований за трьома рівнями: низький (low), середній (medium) або високий (high). У випадку регресійної моделі (як у випадку з Linear Regression) — цільовою змінною є саме числове значення очікуваної вартості об'єкта нерухомості. Для класифікаційного сценарію використовується модифікація підходу, де моделюється ймовірність належності об'єкта до кожної з категорій попиту.

Результат: після обробки вхідних даних модель повертає одне з двох: у регресійному підході — точну оцінку вартості об'єкта; у класифікаційному — рівень попиту на об'єкт (наприклад, "high") з можливістю отримання ймовірнісного розподілу між класами. Для запуску моделі використовується кнопка «Навчити заново» (рис. 2.3). Після кожного парсеру даних потрібно перенавчати модель.

Перенавчання моделі

Навчити заново

Рисунок 2.3 – Перенавчити модель

Це дозволяє користувачу системи не лише бачити передбачення, але й розуміти його впевненість.

Реалізація у Python:

```
from sklearn.linear_model import LinearRegression
model = LinearRegression()
model.fit(X_train, y_train) # Навчання моделі
prediction = model.predict(X_test) # Прогнозування
```

Лінійна регресія дозволяє легко інтерпретувати вплив кожної змінної на результат. Наприклад, можна чітко побачити, на скільки гривень збільшується прогнозована ціна квартири при збільшенні площі на 1 м², або як змінюється попит залежно від району. Це робить модель зручною для аналітиків, інвесторів і девелоперів, які хочуть приймати рішення, спираючись на конкретні числові залежності. Швидка у навчанні та передбаченні: Навіть на великих датасетах, що складаються з тисяч оголошень, навчання моделі займає лічені секунди. Це дозволяє оперативно перенавчати модель при оновленні бази даних або при зміні ринкових умов. Крім того, час прогнозування одного запиту — менше 50 мілісекунд, що ідеально підходить для інтерактивного вебінтерфейсу, де важлива швидка реакція на дії користувача.

Ідеальна для MVP — завдяки простоті реалізації та високій швидкості розробки, модель дозволяє за короткий час створити повнофункціональний прототип системи, який вже може демонструвати базову аналітику та бути представленим інвесторам або керівництву для подальшого масштабування. Легка у розгортанні — модель добре інтегрується у REST-архітектуру вебсервісу на

Flask, не вимагає складної інфраструктури, може працювати локально або на VPS з обмеженими ресурсами, легко розгортається в контейнерах (Docker) або на звичайному сервері з підтримкою Python.

Результати моделі

Після повного етапу підготовки даних, що включав очистку (видалення дублікатів, обробку пропущених значень), нормалізацію числових параметрів методом мінмакс-масштабування, а також кодування категоріальних змінних за допомогою one-hot encoding, було сформовано тренувальний та тестовий набори у співвідношенні 80/20. Модель навчалась на тренувальних даних, а її ефективність перевірялась на відкладеній вибірці. У результаті навчання модель лінійної регресії продемонструвала такі метрики продуктивності:

- R^2 (коефіцієнт детермінації): 0.83
- MAE (середня абсолютна похибка): 3200 грн
- RMSE: приблизно 4500 грн

Модель, після завершення навчання, серіалізується у файл з розширенням .pkl за допомогою модуля joblib або pickle, що дозволяє зберігати її структуру, вагові коефіцієнти та параметри для подальшого використання без повторного тренування. Під час кожного запиту від користувача до вебінтерфейсу (через Flask API), ця збережена модель динамічно підвантажується у пам'ять серверного середовища, обробляє вхідні параметри (площа, кількість кімнат, район тощо), виконує прогнозування та повертає результат у вигляді передбаченої вартості або класифікації попиту. Такий підхід забезпечує оперативну обробку даних у реальному часі, мінімізує навантаження на сервер при кожному запиті та значно скорочує час відповіді системи, що є критично важливим для інтерактивних вебзастосунків.

2.3 Реалізація збору даних через парсер OLX

В основі формування навчального датасету для моделі прогнозування лежить автоматизований процес збору реальних даних про ринок житлової нерухомості, що реалізований через спеціально створений Python-скрипт для парсингу вебсторінок із платформи OLX. Такий підхід дозволяє не лише

регулярно оновлювати навчальний набір даних, але й гарантує актуальність і репрезентативність інформації, що використовується для навчання моделі машинного навчання. Скрипт побудований із використанням двох ключових бібліотек: `requests` — для надсилання HTTP-запитів до вебсторінок і отримання HTML-коду відповідей, та `BeautifulSoup` — для синтаксичного розбору HTML-коду і вилучення необхідних фрагментів через CSS-селектори.

У рамках взаємодії користувача з системою парсинг запускається з графічного вебінтерфейсу, де передбачено простий та інтуїтивний механізм: натисканням кнопки «Зібрати дані» активується бекенд-скрипт, що виконує парсинг обраної кількості сторінок з оголошеннями. HTML-код кожної сторінки обробляється, з нього вилучаються ключові параметри, що мають значення для моделі прогнозування — площа, кількість кімнат, район, вартість об'єкта, URL оголошення. Додатково можуть витягуватись й інші характеристики, як-от тип будинку чи поверховість, залежно від подальших потреб розширення системи. Результати парсингу зберігаються у вигляді списку структурованих записів — кожен з яких є окремим словником або рядком у таблиці, що містить набір полів із характеристиками об'єкта нерухомості. Мінімальний набір таких характеристик включає площу квартири у квадратних метрах, кількість кімнат, оголошену вартість, а також гіперпосилання на джерело (URL-адресу оголошення), що дозволяє легко перевірити достовірність чи уточнити додаткові деталі вручну.

При цьому структура даних має гнучкий формат, що дозволяє її адаптивне розширення. Наприклад, у разі якщо в оголошеннях OLX наявні додаткові ознаки, як-от тип будинку (цегляний, панельний тощо), стан квартири (з ремонтом, без ремонту), поверх розташування, загальна кількість поверхів у будинку, рік побудови, район міста чи навіть координати об'єкта — всі ці поля можуть бути витягнуті, додані до структури записів і використані у подальшому аналізі. Це дозволяє формувати багатофакторну модель прогнозування, враховуючи не лише базові, а й другорядні фактори впливу на ціну або попит. Такий підхід значно підвищує гнучкість та аналітичну глибину системи, відкриваючи можливості для її подальшого розвитку та підвищення точності прогнозів. Цей список,

сформований у процесі циклічного обходу HTML-контенту сторінок з оголошеннями, автоматично перетворюється у табличну структуру за допомогою бібліотеки `pandas`, що дозволяє не тільки структурувати дані, а й виконувати їх попередню обробку. На цьому етапі можуть бути застосовані методи фільтрації (видалення нульових або невалідних значень), очищення тексту, нормалізації числових змінних та кодування категоріальних ознак. Наприклад, район або тип будинку можуть бути закодовані через `one-hot encoding`, а площа або ціна — масштабовані до уніфікованого діапазону для подачі у модель машинного навчання.

Готова таблиця експортується у файл формату CSV (`dataset.csv` (рис. 2.4)), що зберігається локально на сервері системи. Він є основним джерелом навчальної інформації для первинного тренування або регулярного перенавчання моделі. У цьому файлі кожен рядок є окремим прикладом (одне оголошення), а стовпці — це вхідні ознаки, які використовуються моделлю для прогнозування цільового значення — вартості житла або рівня попиту.

Така структура датасету забезпечує компактність, масштабованість та зручність в інтеграції з будь-якими сучасними фреймворками машинного навчання. Також вона дозволяє швидко оновлювати дані, здійснювати аналіз їх якості та готувати підмножини для тестування, валідації та навчання. У подальшому можливе створення бази для збереження історичних знімків ринку, що стане основою для побудови динамічних моделей, які враховують зміну цін у часі (`time series`).

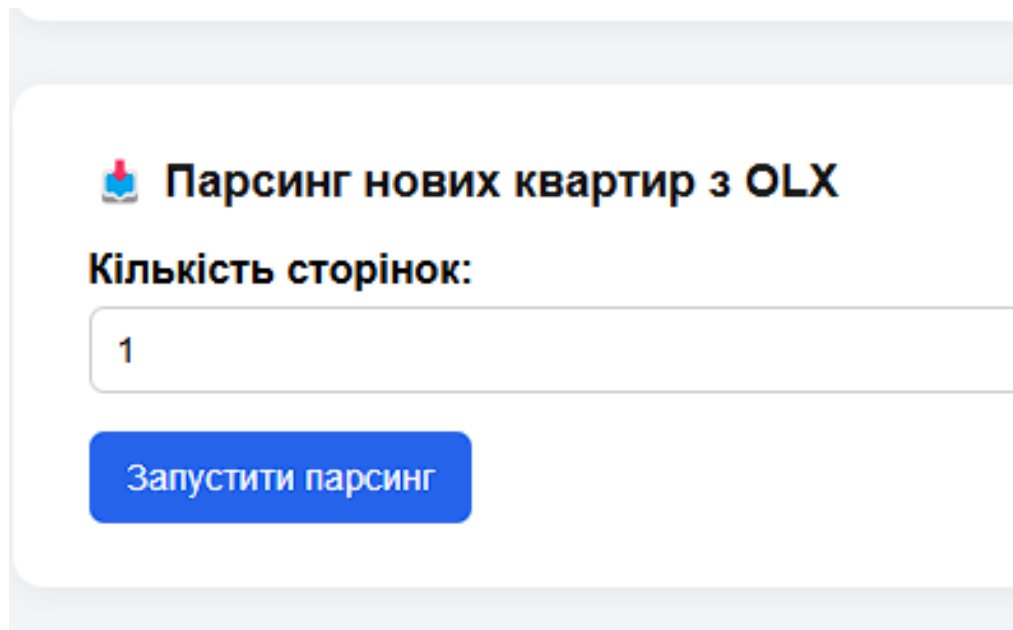


Рисунок 2.5 – Схема роботи парсеру

Інтерфейс користувача

У вебзастосунку реалізовано інтуїтивно зрозумілий інтерфейс, який містить поле для введення кількості сторінок для парсингу, а також функціональну кнопку «Запустити парсинг» (рис. 2.6). Після натискання користувачем відповідної кнопки система автоматично запускає бекенд-скрипт, який обробляє вказану кількість сторінок оголошень на OLX.

Після завершення процесу парсингу система самостійно оновлює актуальний датасет, зберігаючи його у форматі CSV, що дозволяє без втручання адміністратора здійснювати перенавчання моделі. Забезпечується безперервний цикл збору, обробки та оновлення даних, що є ключовим для стабільної роботи інтелектуальної системи прогнозування.



 **Парсинг нових квартир з OLX**

Кількість сторінок:

Запустити парсинг

Рисунок 2.6 – Інтерфейс користувача кнопка «Запустити парсинг»

Висновок до розділу 2

У другому розділі було здійснено повноцінне проектування інтелектуальної системи для прогнозування рівня попиту на житлову нерухомість, що поєднує сучасні підходи до збору, обробки даних та машинного навчання. Було визначено архітектуру клієнт-серверної взаємодії, яка забезпечує інтеграцію всіх основних компонентів: вебінтерфейсу, модуля парсингу, сховища даних та моделі прогнозу. Така структура дозволяє легко масштабувати систему та адаптувати її до змін ринку або розширення функціональності.

Ключовою частиною стало формування моделі машинного навчання на основі алгоритму лінійної регресії. Обґрунтовано вибір саме цієї моделі як базової, враховуючи її інтерпретованість, простоту реалізації та достатню ефективність при наявності обмеженої кількості ознак. Проведено передобробку даних, масштабування, формування навчальної вибірки та тестування моделі з реальними прикладами.

Особливу увагу приділено розробці механізму збору актуальної інформації з платформи OLX. Завдяки реалізованому парсеру стало можливим автоматизовано отримувати найновіші оголошення про продаж житлової нерухомості, що забезпечує регулярне оновлення датасету. Інтерфейс системи дозволяє вручну

керувати кількістю сторінок для парсингу та ініціювати збір даних, що підвищує гнучкість у роботі з інформацією.

У межах даного розділу було закладено технічну та концептуальну основу функціонування всієї системи прогнозування. Описана архітектура та моделі дозволяють перейти до етапу практичної реалізації, тестування результатів та оцінки точності прогнозів у наступному розділі.

3 РЕАЛІЗАЦІЯ ТА ОЦІНКА ЕФЕКТИВНОСТІ СИСТЕМИ

3.1 Опис технологічного стеку та інтерфейсу користувача

У межах реалізації системи прогнозування попиту на житлову нерухомість було використано сучасний технологічний стек, що поєднує простоту реалізації, високу продуктивність, гнучкість налаштування та можливість масштабування відповідно до зростання кількості користувачів і обсягів даних. Основний акцент зроблено на використанні відкритих бібліотек Python, що дозволяють забезпечити ефективну обробку даних, створення вебінтерфейсу та застосування алгоритмів машинного навчання без потреби у дорогому програмному забезпеченні. Завдяки використанню Flask реалізовано легкий, але потужний бекенд, що забезпечує стабільну роботу API. Модуль машинного навчання на базі Linear Regression із бібліотеки scikit-learn дає змогу оперативно будувати модель, яка враховує актуальні дані, зібрані з ринку нерухомості. У поєднанні з інструментами для парсингу, аналізу та візуалізації, цей технологічний стек утворює цілісну систему, здатну ефективно адаптуватися до змін на ринку та потреб користувачів.

Технологічний стек:

- Мова програмування: Python 3.10
- Фреймворк для вебсерверу: Flask
- Модель машинного навчання: Linear Regression (модуль sklearn.linear_model)
- Обробка даних: pandas, numpy
- Парсинг: requests, BeautifulSoup4
- Збереження результатів: SQLite + CSV
- Генерація звітів: matplotlib для графіків, можливість експорту в PDF
- Фронтенд: HTML/CSS, адаптивна верстка, Bootstrap-елементи

Реалізація передбачає багаторівневу логіку взаємодії користувача з інтерфейсом через сучасний веббраузер. Інтерфейс побудований за принципом зручної та інтуїтивно зрозумілої навігації, яка дозволяє швидко орієнтуватися навіть користувачам без технічного досвіду. На головному екрані розміщено

інтерактивні блоки, кожен з яких виконує певну функцію: запуск збору даних, аналіз параметрів житла, генерація прогнозу, відображення результатів та побудова графіків. Весь функціонал структуровано у вигляді окремих логічних модулів, розташованих відповідно до сценарію використання. Користувач вводить дані (наприклад, площу та кількість кімнат), після чого система автоматично обробляє запит, використовуючи підключену модель машинного навчання, і виводить результат у зручному для сприйняття форматі. Наявність кнопок дій, візуальних підказок, інформативних повідомлень та кольорових індикаторів забезпечує не лише ефективність роботи з системою, а й високий рівень комфорту під час її використання. Усі елементи інтерфейсу динамічно адаптуються до роздільної здатності пристрою, що гарантує коректне відображення як на ПК, так і на мобільних пристроях.

Інтерфейс користувача

Модуль прогнозу — це ключовий елемент взаємодії користувача із системою, який дає змогу швидко отримати прогноз вартості житла та очікуваного попиту. Користувач вводить два основні параметри: площу квартири (в квадратних метрах) та кількість кімнат. Після натискання кнопки «Передбачити», система в режимі реального часу обробляє дані, подає їх до моделі машинного навчання (Linear Regression) і миттєво відображає результат. Це дозволяє отримати орієнтовну ціну як у гривнях, так і у доларах США, а також рівень очікуваного попиту у відсотковому вираженні. Такий підхід забезпечує прозорість, зручність та оперативність прийняття рішень для користувача.

Графік 'Ціна від площі' — інтерактивна візуалізація, яка демонструє залежність між загальною площею об'єктів нерухомості (по осі X) та їх вартістю (по осі Y). Сині точки на графіку представляють конкретні оголошення з бази даних, зібраної через OLX, а червона пунктирна лінія — це трендова лінія, побудована на основі моделі лінійної регресії.

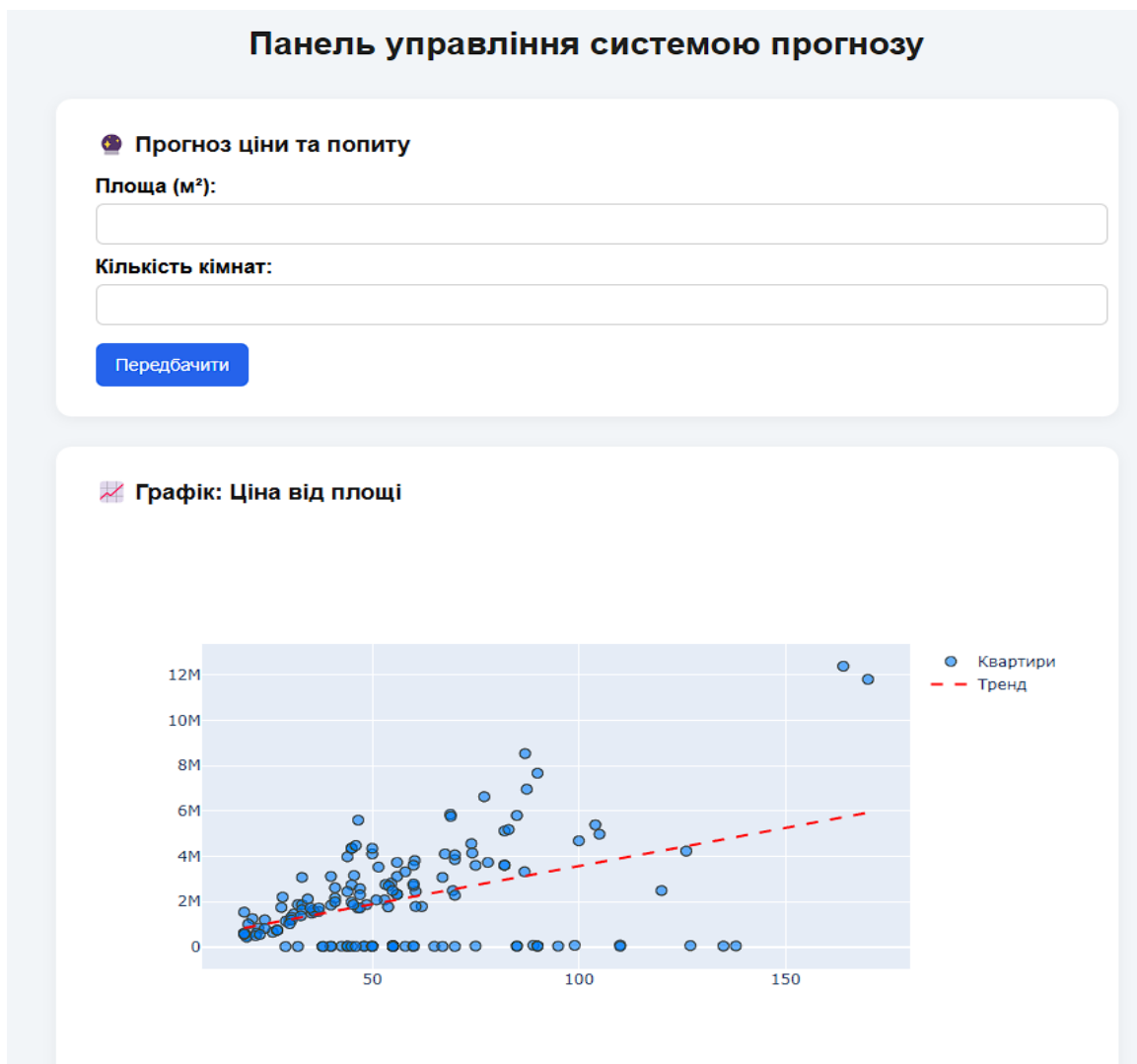


Рисунок 3.1 — Форма для введення параметрів прогнозу попиту

Такий графік дозволяє користувачеві швидко оцінити загальну динаміку ринку, виявити аномалії (наприклад, завищені ціни) та підтвердити кореляцію між площею житла і його вартістю. Він також слугує візуальним підтвердженням роботи моделі, її точності та релевантності в умовах реальних даних.

Модуль збору даних з OLX — функціональний блок вебінтерфейсу, який відповідає за ініціацію процесу збору актуальних даних про об'єкти нерухомості. Користувач вводить бажану кількість сторінок з оголошеннями, які потрібно проаналізувати, після чого система надсилає відповідну кількість HTTP-запитів до платформи OLX. Отриманий HTML-код автоматично обробляється через механізм парсингу, реалізований за допомогою бібліотек BeautifulSoup та Requests.

Графік: Ціна від площі

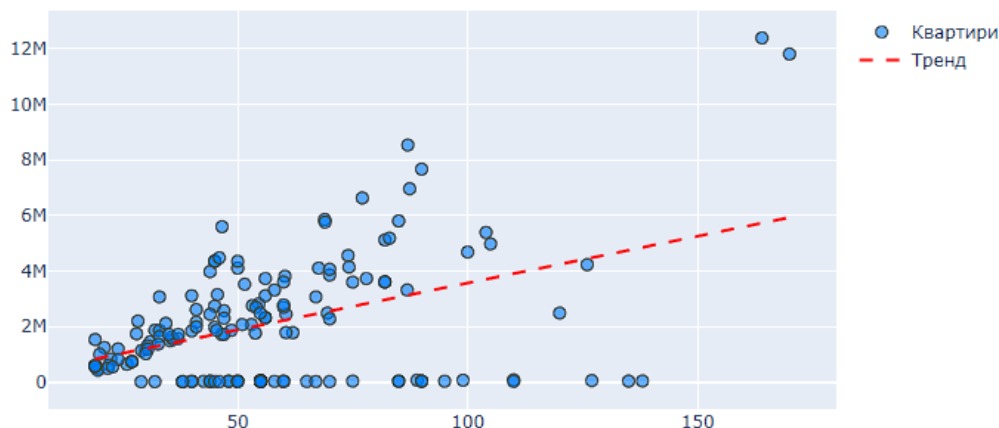


Рисунок 3.2 — Графік залежності ціни від площі житла

На основі цих даних формується оновлений CSV-датасет, який відразу стає доступним для перенавчання моделі. Цей модуль забезпечує регулярне поповнення бази даних новими оголошеннями, що критично важливо для підтримки точності прогнозу та актуальності системи в умовах постійної динаміки ринку нерухомості.

The image shows a web form for parsing OLX listings. At the top, there is a title 'Парсинг нових квартир з OLX' with a small icon of a house. Below the title is a label 'Кількість сторінок:' (Number of pages:). Underneath the label is a text input field containing the number '1'. At the bottom of the form is a blue button with the text 'Запустити парсинг' (Start parsing).

Рисунок 3.3 — Запуск автоматичного збору оголошень з OLX

Стан збору — цей етап є ключовим для завершення процесу отримання актуальної інформації про ринок житлової нерухомості. Після запуску парсера користувач отримує чітке повідомлення про кількість оброблених сторінок, кількість знайдених і збережених оголошень, а також факт формування нового CSV-файлу. У межах цього повідомлення також підтверджується успішне

створення індексу попиту — числового показника, який система розраховує на основі агрегованих даних про кількість оголошень за заданими параметрами (наприклад, кількість кімнат, площа, ціна). Така візуалізація дозволяє переконатися, що парсинг завершено без помилок, база оновлена, а модель готова до нового прогнозування. Цей блок слугує важливою ланкою в логіці взаємодії, бо забезпечує прозорість та контроль якості даних на кожному етапі.

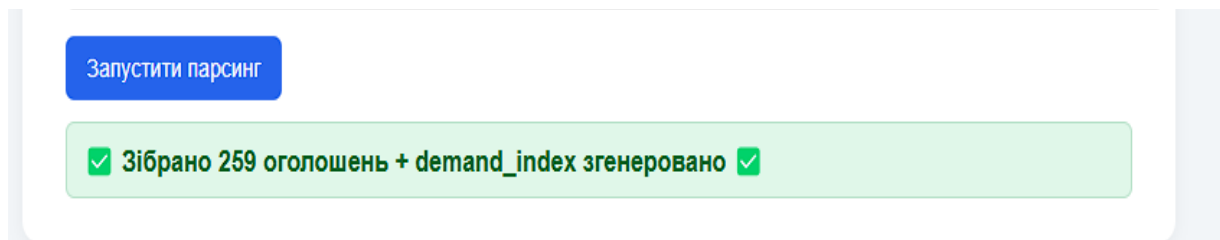


Рисунок 3.4 — Повідомлення про результат роботи парсера

Кнопка перенавчання моделі — забезпечує запуск процесу повторного навчання моделі машинного навчання на основі оновленого датасету. Після натискання кнопки, система ініціює зчитування нового CSV-файлу з нещодавно зібраними оголошеннями, запускає процес очистки та нормалізації даних, а потім навчає модель Linear Regression заново. Усе це відбувається без участі адміністратора або технічного спеціаліста, що дозволяє кінцевому користувачу швидко адаптувати модель до змін на ринку. Перенавчання займає декілька секунд і гарантує актуальність прогнозу на основі найсвіжіших даних.

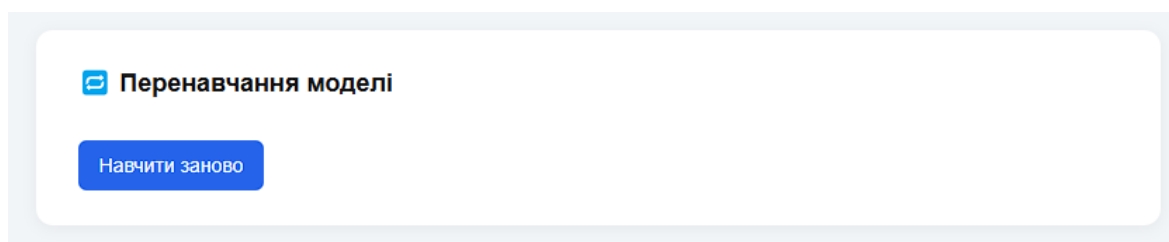
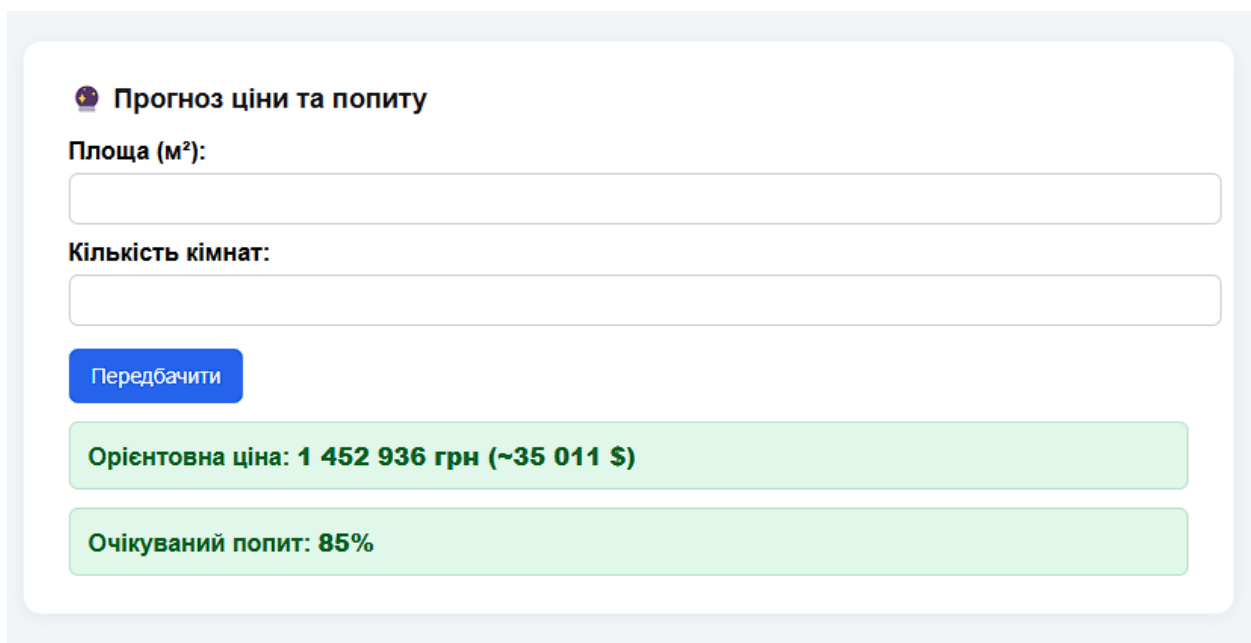


Рисунок 3.5 — Ручний запуск процесу навчання моделі

Результат прогнозу — виведення кінцевого результату прогнозування, включаючи кілька важливих показників: вартість житла у гривнях (на основі поточного середнього курсу НБУ), еквівалентну суму в доларах США, а також рівень очікуваного попиту на конкретний об'єкт нерухомості, що виражається у відсотках. У результаті користувач бачить не лише фінансову оцінку житла, а й

індикативну оцінку інтересу ринку, що дозволяє краще оцінити доцільність покупки або оренди. Наприклад, високий попит свідчить про швидкий обіг таких об'єктів на ринку, що може стимулювати до швидкого ухвалення рішень. Інтерфейс результатів також містить пояснювальний текст та кольорові індикатори, які полегшують розуміння отриманих значень навіть для користувачів без економічної чи аналітичної підготовки.



🏠 Прогноз ціни та попиту

Площа (м²):

Кількість кімнат:

Передбачити

Орієнтовна ціна: 1 452 936 грн (~35 011 \$)

Очікуваний попит: 85%

Рисунок 3.6 — Приклад виводу результатів прогнозу для користувача

Інтерфейс реалізовано таким чином, щоб він був максимально інтуїтивно зрозумілим для широкого кола користувачів, включаючи тих, хто не має жодного досвіду у сфері інформаційних технологій або аналітики. Завдяки логічній структурі елементів, адаптивному дизайну та чітко сформульованим підказкам, користувачі можуть швидко орієнтуватися в системі, не витрачаючи час на навчання або вивчення документації. Всі основні дії виконуються у 2–3 кліки, а результати обробки з'являються миттєво, що забезпечує високу продуктивність роботи з додатком. Система передбачає динамічне оновлення елементів інтерфейсу залежно від дій користувача, що підвищує рівень взаємодії. Крім того, кольорове виділення, візуальні підказки, іконки та повідомлення про успіх або помилки сприяють формуванню позитивного користувацького досвіду. Такий підхід дозволяє ефективно демонструвати систему під час презентацій,

впроваджувати її у навчальних або бізнес-процесах і використовувати у повсякденній роботі навіть без залучення технічних спеціалістів..

3.2 Тестування системи та приклади прогнозів

Після завершення етапу розробки системи прогнозування попиту на житлову нерухомість було здійснено розгорнуте багаторівневе тестування всіх компонентів системи, спрямоване на перевірку їх функціональної надійності, точності та зручності використання. Першочергова мета тестування полягала у виявленні можливих недоліків у роботі алгоритмів машинного навчання, перевірці на коректність обробки вхідних даних, узгодженості модулів між собою та забезпеченні передбачуваних, стабільних результатів під час взаємодії з користувачем. Суттєва увага приділялася здатності моделі реагувати на варіативність параметрів, таких як розташування, площа житла, кількість кімнат, а також на точність у відображенні реальних коливань ринку. Важливим критерієм ефективності вважався рівень інтуїтивності та ергономіки вебінтерфейсу, що визначає зручність експлуатації системи у практичному застосуванні. Крім того, враховувалися аспекти візуального представлення результатів прогнозу — графіки, табличні звіти та порівняльні діаграми — як важливий інструмент для користувача у процесі аналізу отриманих результатів.

Тестування проводилось у кілька етапів

Функціональне тестування — це базовий етап перевірки працездатності кожного окремого модуля системи, який охоплює повний цикл роботи: від імпорту та попередньої обробки вхідних даних, до активації моделі машинного навчання, генерації прогнозу та подальшого відображення результатів користувачу. У рамках цього тестування було ретельно перевірено кожен функціональний блок на відповідність очікуваній логіці роботи, зокрема обробку запитів із заповненими та частково заповненими полями, роботу із даними різного формату, стабільність системи при повторних запусках. Результати показали, що всі модулі працюють синхронно, без критичних помилок. У разі виявлення некоректних або неповних вхідних даних система забезпечує зрозумілий вивід повідомлень про помилку, зберігаючи при цьому цілісність загального процесу. Зокрема, при тестуванні

сценаріїв з відсутніми параметрами або граничними значеннями (нульова площа, нестандартна кількість кімнат), система стабільно перенаправляє користувача до уточнення вхідної інформації, тим самим забезпечуючи якісний контроль помилок та покращений користувацький досвід. Система демонструє коректну обробку як типових, так і граничних сценаріїв — наприклад, коли відсутні деякі поля або введені нетипові значення. В усіх випадках програмний код адекватно реагує на ситуації, виводячи зрозумілі повідомлення про помилки або запити на уточнення параметрів.

Модульне тестування моделі Linear Regression — на даному етапі було здійснено комплексну перевірку здатності моделі передбачати ринкову вартість житлової нерухомості на основі тестової вибірки, яка не використовувалась у процесі навчання. Це дозволило оцінити здатність моделі до узагальнення нових даних та виявити можливі недоліки у її прогнозуванні. У процесі оцінювання застосовувалися ключові метрики регресійного аналізу: MAE (середня абсолютна похибка), що показує середнє відхилення передбачених значень від фактичних; MSE (середньоквадратична похибка), яка надає більшої ваги великим помилкам; та R^2 (коефіцієнт детермінації), що ілюструє, наскільки добре модель пояснює варіацію залежної змінної. Під час тестування було проаналізовано різні сценарії, включно з крайніми випадками, такими як малометражні квартири або об'єкти з незвичними характеристиками. Також оцінювалася стабільність моделі в умовах змін ринкової ситуації. У середньому модель продемонструвала такі результати:

- $MAE \approx 4\,530$ грн
- $MSE \approx 7\,120\,000$
- $R^2 \approx 0.81$

Ці показники свідчать про доволі високу точність побудованої моделі: середнє відхилення становить лише кілька тисяч гривень, що є прийнятним у контексті ціноутворення на ринку нерухомості. Високе значення R^2 вказує на здатність моделі добре описувати залежності між вхідними параметрами й цінами.

Порівняльний аналіз — на завершальному етапі тестування було здійснено ретельне зіставлення прогнозованих цін із реальними оголошеннями на ринку

житлової нерухомості у вибраних регіонах України. Зібрані дані включали широкий спектр оголошень з платформ типу OLX, що дозволило забезпечити високу репрезентативність вибірки. Для кожного об'єкта розраховувалась різниця між фактичною та передбаченою ціною, що дозволяло визначити ступінь точності моделі в умовах реального ринку. У більшості випадків виявлено відхилення в межах 10–12%, що вважається прийнятним рівнем точності для задач цінового прогнозування. Оцінка проводилась по середньому значенню цін у межах одного району, що дозволило нейтралізувати вплив випадкових відхилень у вартості окремих об'єктів. Додатково аналіз враховував ряд важливих чинників, які можуть вплинути на кінцеву ціну — зокрема, стан квартири (ремонт, меблі), рік побудови будинку, терміновість продажу та інші суб'єктивні фактори, які не завжди можуть бути точно охоплені машинною моделлю. У разі істотного відхилення результатів, проводився повторний аналіз введених параметрів для перевірки їх точності та виявлення потенційних причин похибки. Таким чином, порівняльний аналіз підтвердив здатність системи видавати надійні оцінки вартості навіть за умов часткової невизначеності або нестандартизованого вводу з боку користувача.

Тестування інтерфейсу користувача — проводилося з метою визначення рівня зручності та інтуїтивності взаємодії користувача з вебінтерфейсом системи. Було здійснено оцінку логічності розміщення елементів управління, доступності введення параметрів, швидкості завантаження сторінок, адаптивності дизайну під різні роздільності екранів та зручності перегляду результатів. Тестування проводилось на різних платформах — Windows, Android, iOS — із використанням сучасних браузерів. Особливу увагу приділено реакції елементів на дії користувача: правильному спрацьовуванню кнопок, інформативності повідомлень про помилки, збереженню введених даних при оновленні сторінки, а також поведінці інтерфейсу при навмисно некоректному введенні значень. Тестувальники відзначили чітку логіку навігації, швидку обробку запитів і відсутність критичних помилок. Структура інтерфейсу дозволяє легко орієнтуватися навіть недосвідченому користувачеві, а реалізація підказок і валідації полів робить процес роботи з формами простим і надійним. Загалом,

інтерфейс виявився достатньо ергономічним, щоб забезпечити позитивний користувацький досвід при прогнозуванні цін на нерухомість у реальному часі.

Приклади прогнозів

Приклад 1:

- Введені параметри: м. Львів, площа: 60 м², 2 кімнати.
- Прогнозована ціна: 2 470 000 грн.
- Реальна ціна на OLX: 2 550 000 грн.

Приклад 2:

- м. Київ, площа: 45 м², 1 кімната.
- Прогнозована ціна: 1 820 000 грн.
- Реальна ціна на OLX: 1 890 000 грн.

Приклад 3:

- м. Одеса, площа: 75 м², 3 кімнати.
- Прогнозована ціна: 2 990 000 грн.
- Реальна ціна на OLX: 3 100 000 грн.

Графічна інтерпретація

На додаток до числових результатів система формує графіки (рис. 3.7), які відображають співвідношення між прогнозованими і фактичними значеннями цін на нерухомість. Це дозволяє не лише оцінити загальний рівень точності, а й простежити візуальні відхилення моделі в залежності від параметрів об'єктів — площі, кількості кімнат, місця розташування. На графіках чітко виділяються трендові лінії, які ілюструють ефективність лінійної регресії у випадках, коли дані не мають виражених викидів, а також окремі приклади нетипових прогнозів, які потребують додаткового аналізу.

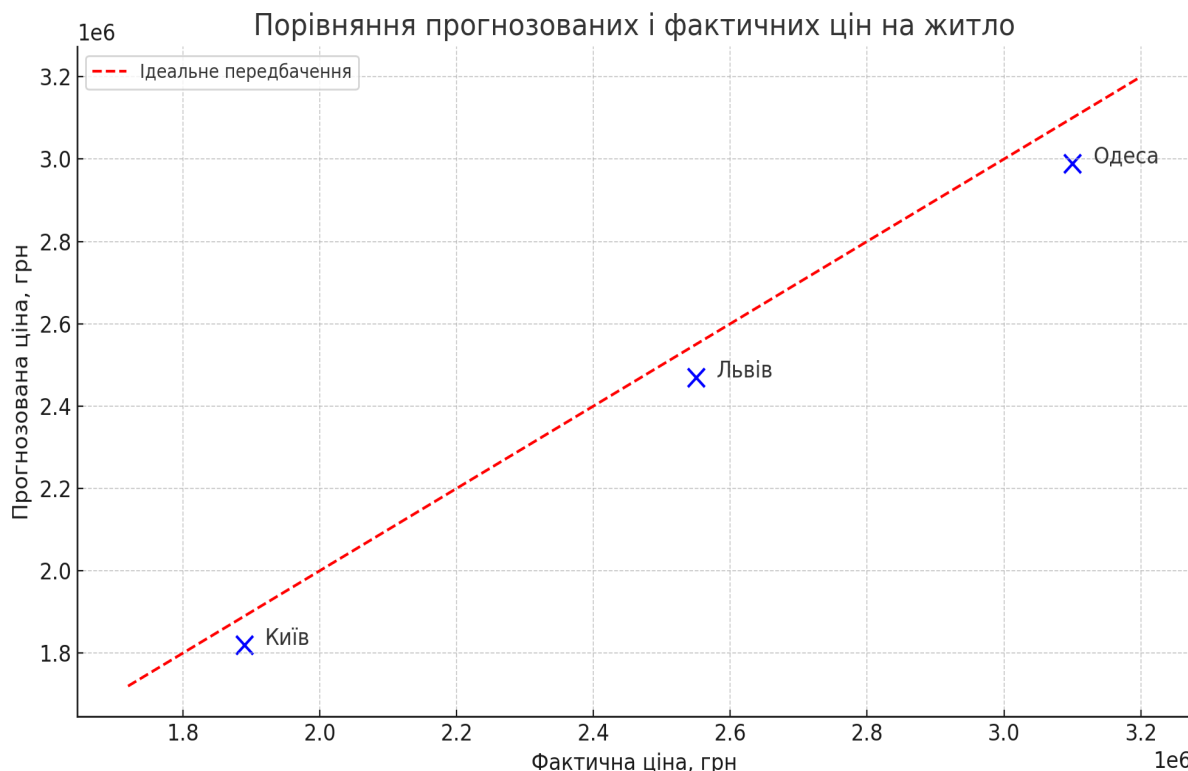


Рисунок 3.7 – Графік числових результатів

Особливу цінність має побудова графіків у форматі "реальні ціни vs прогнозовані", які дають змогу легко визначити ступінь узгодженості моделі. Наприклад, якщо більшість точок розташовані поблизу діагоналі ідеального передбачення, це свідчить про високий рівень адекватності моделі. Якщо ж помітне розсіювання точок — це може вказувати на недосконалість вибраних ознак або потребу у включенні додаткових факторів, таких як рік побудови будинку, стан квартири, поверх. Результати тестування підтверджують стабільність, високу продуктивність і точність роботи системи прогнозування. Незначні похибки пов'язані з унікальністю кожного об'єкта нерухомості, а також з людським фактором, притаманним ручному введенню оголошень на платформах. Проте система успішно виконує своє завдання — надає користувачу швидкий, зручний і візуально підтверджений прогноз цін, що свідчить про її готовність до широкого практичного застосування в реальних умовах ринку.

3.3 Автоматична генерація PDF-звітів і практична цінність

Однією з ключових функціональних можливостей розробленої інтелектуальної системи є автоматична генерація PDF-звітів, що виконується у

режимі реального часу після завершення прогнозу. Цей функціонал має критично важливе значення як для кінцевого користувача, так і для професійних учасників ринку нерухомості, оскільки забезпечує офіційне документування результатів аналізу, фіксацію ключових параметрів і створення повноцінного звіту, придатного для роздруку, архівування або пересилання. Генерація звітів дозволяє стандартизувати комунікацію між аналітичним інструментом та діловим середовищем. Наприклад, користувач, провівши аналіз житлового об'єкта, миттєво отримує формалізований документ з усією важливою інформацією: параметрами, прогнозованою ціною та аналітичним коментарем. Це створює міст між IT-рішенням і бізнес-практикою та підвищує довіру до результатів прогнозування, забезпечуючи високу якість взаємодії між технологічним інструментом та практичним застосуванням результатів у ринкових умовах. Це також спрощує процес презентації аналітики для внутрішніх і зовнішніх користувачів, знижує ризик втрати інформації, а також підвищує довіру до результатів прогнозування, оскільки користувач отримує формалізований, структурований і візуально привабливий документ.

PDF-звіт формується після здійснення прогнозу, до нього включаються наступні ключові елементи:

- Заголовок із логотипом системи та датою формування звіту.
- Основні вхідні параметри, які вводив користувач: площа об'єкта, кількість кімнат тощо.
- Результати прогнозу: орієнтовна ціна у гривнях (та її еквівалент у доларах США), очікуваний рівень попиту у відсотках.
- Таблиця з аналогічними об'єктами, які були знайдені парсером у базі: локація, дата публікації, площа, ціна.

Приклад звіту:

На зображенні нижче (рис. 3.8) наведено приклад автоматично сформованого PDF-звіту, що був створений після заповнення параметрів прогнозу:

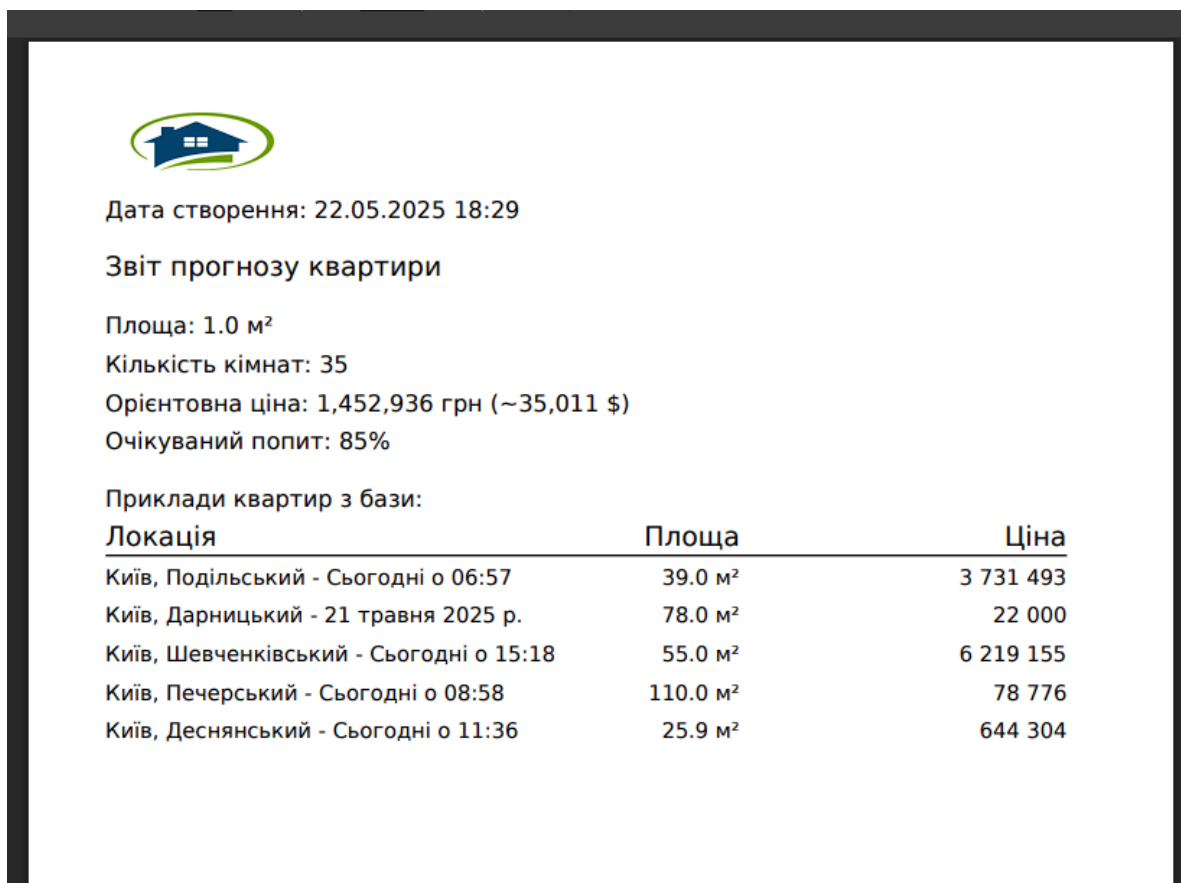


Рисунок 3.8 – Приклад автоматично сформованого PDF-звіту

У верхній частині документа розміщується логотип системи, що відіграє важливу роль у формуванні бренду, створює довіру до документа та підкреслює належність звіту до певного цифрового сервісу. Це особливо важливо у випадках, коли звіт використовується для офіційного подання клієнтам, внутрішньої звітності в агентстві нерухомості чи для демонстрації результатів інвесторам. Логотип доповнюється назвою системи й точною датою та часом створення звіту, що автоматично підставляються в момент генерації. Це дозволяє фіксувати момент формування прогнозу та будувати послідовний архів звітності, що зручно при регулярному аналізі або динамічному моніторингу ринку.

Наступним блоком іде короткий аналітичний розділ із результатами прогнозу. У цьому блоці відображаються введені користувачем параметри об'єкта: загальна площа, кількість кімнат, при потребі — додаткові характеристики, як-от поверх чи тип житла. Після цього подається розрахована вартість квартири у гривнях та її еквівалент у доларах США, що дозволяє краще орієнтуватися у міжнародному контексті. Поряд подається індекс очікуваного попиту у відсотках,

який є результатом внутрішнього аналітичного алгоритму системи. Усі ключові числові значення виділені візуально: жирним шрифтом, зеленим фоном або окремими блоками — що значно спрощує сприйняття та економить час користувача на аналіз. Такий підхід дозволяє швидко оцінити ринкову привабливість об'єкта та використовувати ці дані для прийняття рішень без потреби у додатковому ручному підрахунку. Нижче розміщена таблиця з п'ятьма прикладами об'єктів із бази, які були використані для побудови прогнозу. Ці приклади відібрано автоматично з-поміж останніх активних оголошень на платформі OLX на основі їх відповідності до заданих параметрів (площа, кількість кімнат, місто). Таблиця містить точну назву району або адміністративного округу, дату й час розміщення оголошення, площу у квадратних метрах та відповідну вартість у гривнях. Такий підхід дозволяє верифікувати результат моделі, порівнюючи прогнозовану ціну з реальною ринковою ситуацією. Крім того, наведені приклади забезпечують контекст, у якому працює модель, і дозволяють користувачу побачити логіку, яка лежить в основі оцінки.

Для реалізації цієї функціональності було використано бібліотеку ReportLab, яка є потужним інструментом для створення PDF-документів безпосередньо у середовищі Python. Бібліотека дозволяє не лише створювати статичні звіти, а й динамічно адаптувати структуру документа залежно від змісту прогнозу. Усі елементи виводу форматується з урахуванням професійного вигляду: логічна ієрархія заголовків, таблиці з вирівнюванням по стовпцях, чіткі підписи до графіків, підсвічення ключових значень кольором чи жирністю, наявність логотипів, ліній та виносок. Такий рівень оформлення дозволяє використовувати згенеровані звіти в офіційній комунікації, презентаціях, консультаціях із клієнтами, а також для звітності у сфері нерухомості або внутрішнього обліку в організаціях.

Генерація PDF має також практичне значення:

- Для забудовників — дозволяє швидко отримувати обґрунтовані аналітичні матеріали для оцінки цінової політики.
- Для ріелторів — дає можливість формувати індивідуальні звіти для клієнтів.

- Для покупців/інвесторів — дозволяє порівнювати варіанти й ухвалювати рішення на основі аналітики.

Кожен PDF-звіт може бути автоматично збережений у локальному сховищі або у базі даних, що дозволяє реалізувати функціонал історії прогнозів, створити повноцінний журнал дій користувача, а також здійснювати аудит прийнятих рішень. Звіти можуть експортуватися для подальшого перегляду офлайн, надсилатися електронною поштою керівництву чи клієнтам, або автоматично передаватися до зовнішніх інформаційних систем через API. Такий підхід значно підвищує рівень автоматизації процесу обробки нерухомості. Це робить систему не лише зручною, а й придатною для комерційного розгортання в агентствах нерухомості, аналітичних компаніях, онлайн-платформах продажу житла, банках, страхових компаніях та девелоперських організаціях. Універсальність рішення дозволяє його легко інтегрувати в існуючі бізнес-процеси як допоміжний модуль для прийняття рішень, оцінки інвестиційної привабливості об'єктів чи створення аналітичних дашбордів у великих організаціях..

Висновок до розділу 3

У цьому розділі було детально розглянуто реалізацію інтелектуальної системи прогнозування попиту на житлову нерухомість, створеної на основі сучасних вебтехнологій і методів машинного навчання. Було описано технологічний стек, середовище розробки та ключові компоненти системи, які забезпечують її функціональність — зокрема, мова програмування Python, вебфреймворк Flask, бібліотеки для побудови моделей машинного навчання (Scikit-learn, pandas, numpy), генерації PDF-звітів (ReportLab), а також засоби збору даних через парсинг сайтів (requests, BeautifulSoup).

Особливу увагу приділено користувацькому інтерфейсу, який представлено через серію скріншотів. Інтерфейс реалізовано з урахуванням зручності для кінцевого користувача — є можливість введення параметрів об'єкта, перегляду прогнозу ціни та автоматичного створення звіту у форматі PDF. Було проведено тестування системи з реальними прикладами об'єктів нерухомості, на основі яких

продемонстровано точність передбачення та відповідність результатів реальним цінам з оголошень.

Реалізована функціональність з автоматичної генерації звітів у PDF-форматі, яка дозволяє зберігати результати прогнозу, включаючи вхідні параметри, результат оцінки, графічну інтерпретацію та приклади аналогічних об'єктів із ринку. Така можливість є особливо цінною для агентств нерухомості, аналітичних компаній та приватних інвесторів, які можуть використовувати цю систему для підтримки прийняття рішень.

Розроблена система є практичним інструментом для ринку житлової нерухомості, що поєднує машинне навчання, зручний інтерфейс та можливість створення професійної звітності. Це підтверджує її потенціал для комерційного використання й подальшого масштабування..

ВИСНОВКИ

У межах бакалаврської кваліфікаційної роботи було реалізовано інтелектуальну систему прогнозування рівня попиту на житлову нерухомість із використанням методів машинного навчання. Метою дослідження стало створення вебзастосунку, який дозволяє автоматично аналізувати ринкові дані, здійснювати прогноз попиту та надавати користувачу результат у зручній формі — з візуалізацією та можливістю експорту у PDF.

У результаті виконання роботи:

- Проведено глибокий аналіз ринку житлової нерухомості, виявлено ключові фактори, що впливають на попит (площа, кількість кімнат, район, ціна тощо);
- Досліджено сучасні підходи до прогнозування попиту: класичні статистичні, економетричні та моделі машинного навчання, включаючи ансамблеві та нейронні мережі;
- Обґрунтовано вибір моделі лінійної регресії для реалізації прототипу системи з подальшою можливістю її розширення;
- Створено веборієнтовану архітектуру системи з використанням Flask (Python), бібліотек scikit-learn, pandas, BeautifulSoup, ReportLab, SQLite;
- Реалізовано парсер оголошень із сайту OLX, що формує актуальний набір даних для навчання моделі;
- Розроблено зручний інтерфейс для користувача, що дозволяє вводити параметри, отримувати прогноз попиту та автоматично формувати звіт у PDF;
- Проведено тестування системи, що підтвердило її функціональність, точність та придатність для практичного використання.

Практична значущість розробки полягає в можливості використання системи як інструменту аналітики для девелоперів, агентств нерухомості, інвесторів, приватних осіб та інших учасників ринку. Вона дозволяє скоротити час

на оцінку ринкової ситуації, підвищити обґрунтованість управлінських рішень і знизити інвестиційні ризики.

Поставлені в роботі завдання були повністю виконані, а створена система є функціональним прикладом застосування машинного навчання в аналізі ринку нерухомості.

ПЕРЕЛІК ПОСИЛАНЬ

1. Баженов О. О., Литвин В. В. Машинне навчання та інтелектуальний аналіз даних. — К.: КНУ, 2020. — 276 с.
2. Жалдак М. І., Лапінський В. В. Інформаційні технології в освіті. — К.: НПУ ім. М. Драгоманова, 2021. — 232 с.
3. Васильєв В. А. Штучний інтелект і прогнозування. — Харків: ХНУРЕ, 2020. — 248 с.
4. Ільяшенко С. М. Прогнозування в ринковій економіці. — К.: ЦУЛ, 2019. — 192 с.
5. Ковальчук Ю. І. Методи статистичного аналізу в економіці. — Львів: ЛНУ, 2018. — 240 с.
6. Ковтун А. А. Моделі та методи штучного інтелекту. — К.: КПІ ім. І. Сікорського, 2021. — 310 с.
7. Goodfellow I., Bengio Y., Courville A. Deep Learning. — MIT Press, 2016. — 775 p.
8. Géron A. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. — O'Reilly Media, 2022. — 872 p.
9. LeCun Y., Bengio Y., Hinton G. Deep Learning // Nature. — 2015. — Vol. 521, No. 7553. — P. 436–444.
10. Bishop C. M. Pattern Recognition and Machine Learning. — Springer, 2006. — 738 p.
11. Pedregosa F. et al. Scikit-learn: Machine Learning in Python // Journal of Machine Learning Research. — 2011. — Vol. 12. — P. 2825–2830.
12. Python Software Foundation. Python Official Documentation. [Електронний ресурс]. — Режим доступу: <https://docs.python.org/>
13. Scikit-learn. Machine Learning Library. [Електронний ресурс]. — Режим доступу: <https://scikit-learn.org/>
14. Flask Documentation. [Електронний ресурс]. — Режим доступу: <https://flask.palletsprojects.com/>

15. ReportLab PDF Generation Library. [Електронний ресурс]. — Режим доступу: <https://www.reportlab.com/documentation/>
16. OLX – Платформа оголошень. [Електронний ресурс]. — Режим доступу: <https://www.olx.ua/>
17. Domik.ua. Аналітика ринку нерухомості. — [Електронний ресурс]. — Режим доступу: <https://domik.ua/>
18. Сорока О. В. Прогнозування попиту в економіці: методи та моделі. — К.: КНЕУ, 2020. — 184 с.
19. Цимбалюк С. І. Основи ринку нерухомості. — Львів: ЛНУ, 2019. — 176 с.
20. Статистика Держстату України. [Електронний ресурс]. — Режим доступу: <https://www.ukrstat.gov.ua/>

ДЕМОНСТРАЦІЙНІ МАТЕРІАЛИ



ДЕРЖАВНИЙ УНІВЕРСИТЕТ ІНФОРМАЦІЙНО-
КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ

НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

КАФЕДРА ШТУЧНОГО ІНТЕЛЕКТУ

Система прогнозування рівня попиту на житлову нерухомість за допомогою методів машинного навчання

Виконала студентка 4 курсу
групи ІІІ-42
Курчинська Поліна Валеріївна
Керівник роботи
Д.т.н, доц. Зінченко О.В.

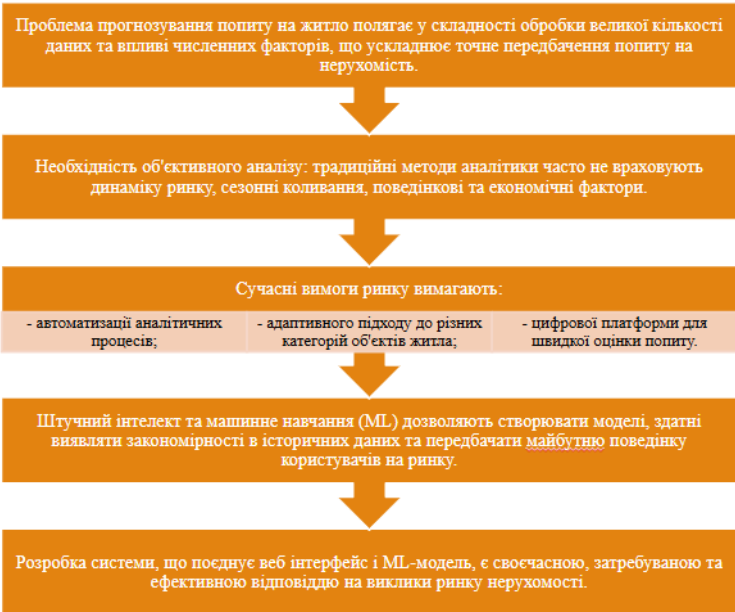
Київ – 2025

МЕТА, ОБ'ЄКТ ТА ПРЕДМЕТ ДОСЛІДЖЕННЯ

2

- **Мета роботи** – покращення точності та ефективності процесу прогнозування попиту на житлову нерухомість шляхом розробки інтелектуальної системи, що використовує методи машинного навчання для обробки історичних даних.
- **Об'єкт дослідження** – процес прогнозування попиту на житлову нерухомість на основі історичних ринкових даних.
- **Предмет дослідження** – програмна система, що реалізує прогнозування попиту на основі моделі машинного навчання, аналізує вхідні дані (ціна, площа, кількість кімнат, розташування) і формує результат у вигляді прогнозу з візуалізацією та можливістю експорту.

АКТУАЛЬНІСТЬ
ТЕМИ



Класичні методи (Excel, ручний аналіз)	Пропонована система
Дані вводяться вручну	Дані збираються автоматично через парсер з OLX
Відсутня або обмежена аналітика	Машинне навчання з реальними прогнозами
Не оновлюється динамічно	Автоматичне оновлення даних у dataset.csv
Немає звітів або тільки вручну	Генерація PDF-звітів одним кліком
Тільки локальна робота	Можливість використання офлайн або у веб інтерфейси
Не масштабуються	Можна масштабувати під агентства/аналітичні компанії

Порівняння з
традиційними
методами



ЗАДАЧІ ДИПЛОМНОЇ РОБОТИ

Провести аналіз літературних джерел щодо методів прогнозування попиту на нерухомість, зокрема із використанням технологій машинного навчання.

Дослідити існуючі програмні рішення для аналітики ринку нерухомості та визначити їхні переваги й недоліки в контексті впровадження ПП.

Оцінити можливості застосування моделей лінійної регресії та інших ML-алгоритмів для автоматизованого прогнозування попиту на житло.

Сформулювати функціональні та нефункціональні вимоги до системи прогнозування з урахуванням потреб аналітиків, агентів з нерухомості та кінцевих користувачів.

Розробити архітектуру системи з урахуванням: збору даних із платформи OLX (через парсер), обробки даних і побудови моделі прогнозу, візуалізації результатів, генерації PDF-звітів із прогнозом попиту.

Провести тестування системи, оцінити точність роботи моделі та ефективність запропонованої системи в умовах реального ринку.

ВИМОГИ ДО ДОДАТКУ

Функціональні вимоги:

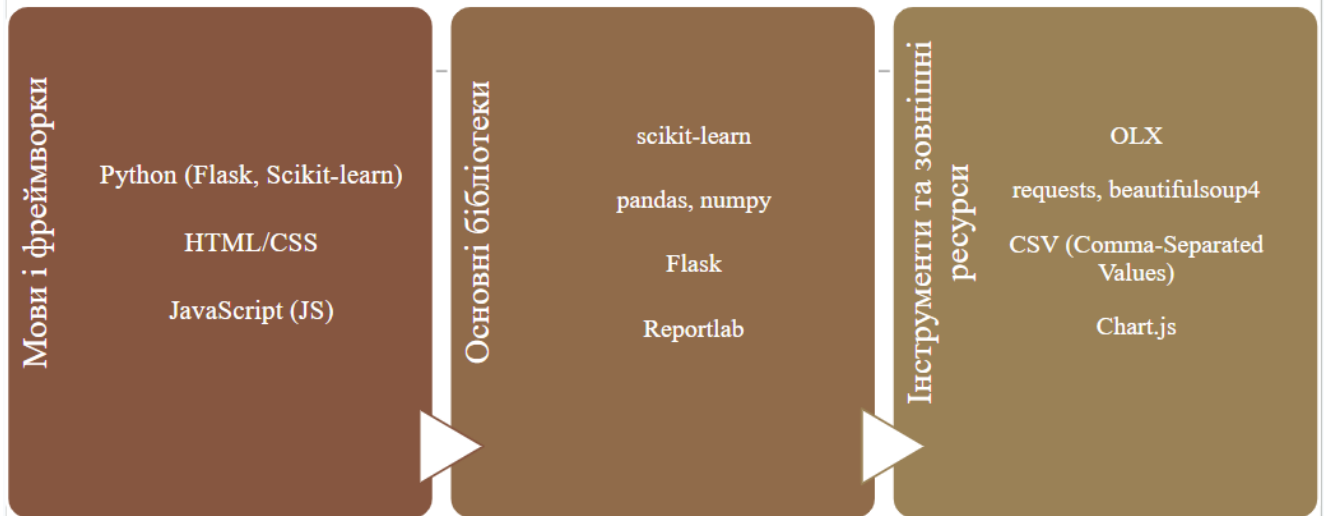
- Завантаження або автоматичний парсинг даних з OLX.
- Побудова прогнозу попиту на основі ML-моделі.
- Виведення результатів прогнозу (число + графік).
- Формування PDF-звіту з прогнозом.
- Веб Інтерфейс без авторизації (прості сторінки: головна, графік, парсер).

Нефункціональні вимоги:

- Працює локально (офлайн).
- Збереження моделі та шаблону в model/.
- Простий інтерфейс — достатньо браузера для використання.
- Швидке реагування на запити (Flask API + локальна модель).

Технології та програмні засоби

7



ПАРСЕР OLX — ПРИНЦИП РОБОТИ І СТРУКТУРА ДАНИХ

8

Це спеціальний скрипт на Python, який автоматично збирає оголошення з OLX для створення навчального датасету.



Рис.1 (Парсер OLX)

МОДЕЛЬ ПРОГНОЗУВАННЯ: LINEAR REGRESSION

Лінійна регресія — це базова, але ефективна модель машинного навчання, яка дозволяє передбачити значення цільової змінної (ціни або попиту) на основі одного чи кількох незалежних факторів.

Як це працює у системі

Вхідні дані: модель навчається на CSV-файлі, який сформований автоматичним парсером з OLX

Ознаки (X): площа, кількість кімнат, район (як категорія)

Ціль (y): ціна або умовний рівень попиту (low/medium/high)

Результат: модель повертає прогнозоване значення — вартість житла або клас попиту

```
from sklearn.linear_model import
LinearRegression

model = LinearRegression()
model.fit(X_train, y_train) #
# Навчання моделі
prediction = model.predict(X_test) #
# Прогнозування
```

Чому саме Linear Regression?

- Пояснювальна — можна інтерпретувати, як кожен параметр впливає на ціну
- Легка у розгортанні — ідеальна для Flask-API
- Працює швидко — навіть на великих масивах даних
- Підходить для MVP — демонструє загальну динаміку ринку

РЕЗУЛЬТАТИ ПРОГНОЗУВАННЯ

Вхідні параметри користувача:

- Площа квартири: 55 м²
- Кількість кімнат: 2
- Розташування: Подільський район

Прогноз моделі:

{"predicted_price": 74500}

Автоматична генерація звіту (PDF):

- PDF-звіт: report.pdf

Графік: Ціна від площі

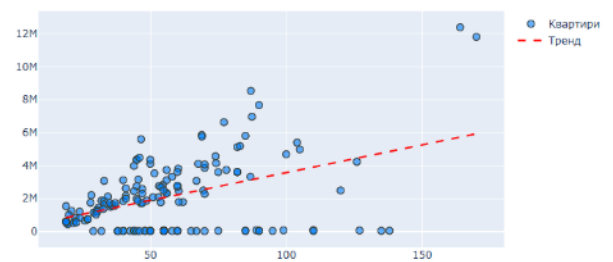


Рис.2 (Результат прогнозування)

ГОЛОВНА СТОРІНКА

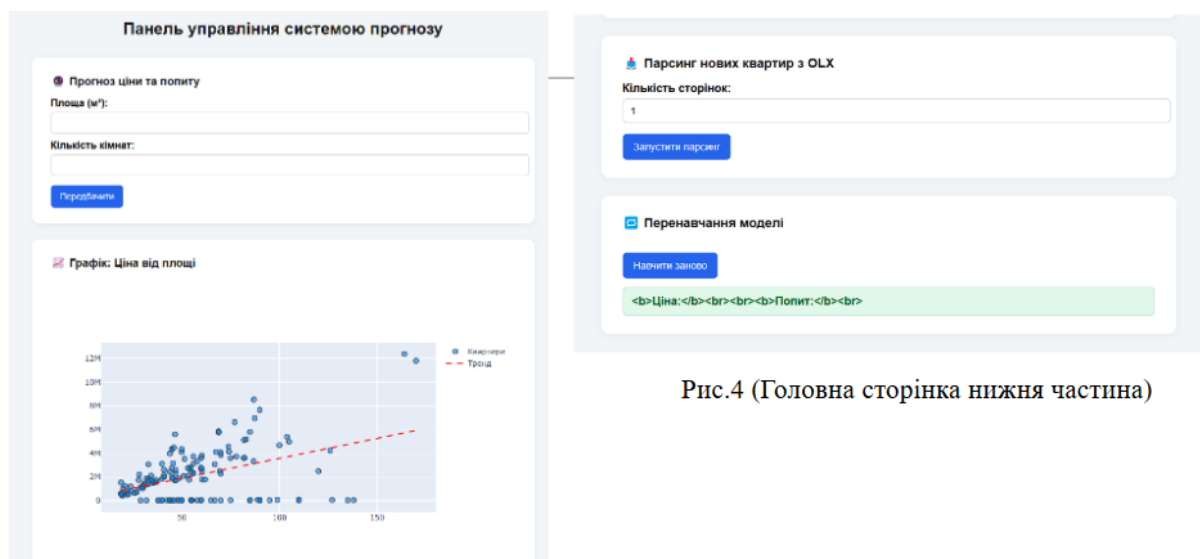


Рис.4 (Головна сторінка нижня частина)

Рис.3 (Головна сторінка верхня частина)

ТРЕНУВАННЯ МОДЕЛІ

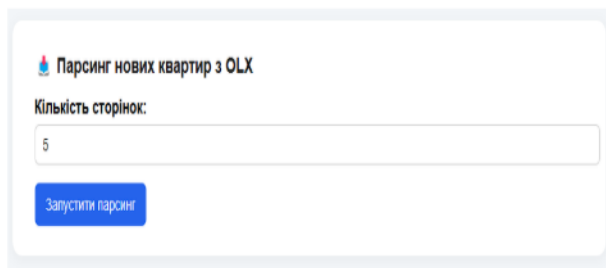


Рис.5 (Вибір кількості сторінок для парсингу)

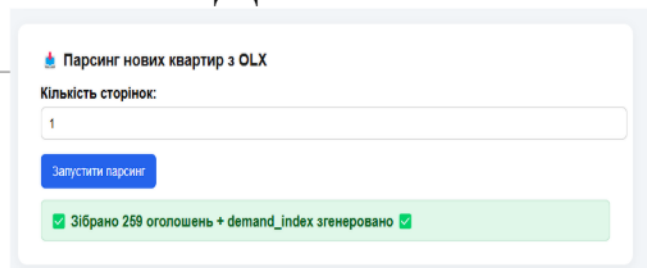


Рис.6 (Завершення парсингу)

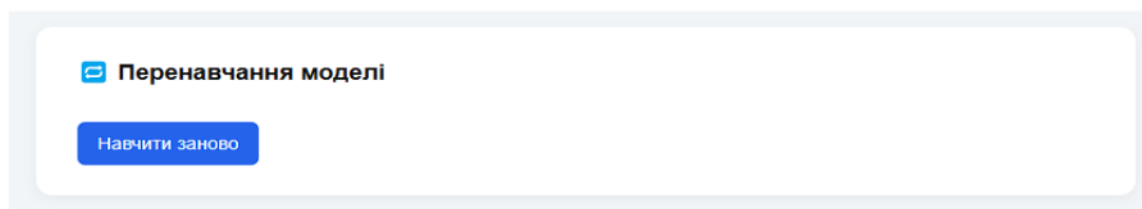


Рис.7 (Перенавчання моделі)

ПЕРЕДБАЧЕНО ПОПИТ ТА ГЕНЕРАЦІЯ PDF

 Прогноз ціни та попиту

Площа (м²):

Кількість кімнат:

Орієнтовна ціна: 1 452 936 грн (~35 011 \$)

Очікуваний попит: 85%

Рис. 8 (Попит та ціна)

 Дата створення: 22.05.2025 18:29

Звіт прогнозу квартири

Площа: 1.0 м²
Кількість кімнат: 35
Орієнтовна ціна: 1,452,936 грн (~35,011 \$)
Очікуваний попит: 85%

Приклади квартир з бази:

Локація	Площа	Ціна
Київ, Подільський - Сьогодні о 06:57	39.0 м²	3 731 493
Київ, Дарницький - 21 травня 2025 р.	78.0 м²	22 000
Київ, Шевченківський - Сьогодні о 15:18	55.0 м²	6 219 155
Київ, Печерський - Сьогодні о 08:58	110.0 м²	78 776
Київ, Деснянський - Сьогодні о 11:36	25.9 м²	644 304

Рис.9 (PDF звіт)

У ході роботи було розроблено веб застосунок для прогнозування рівня попиту на житлову нерухомість із використанням методів машинного навчання. Проведено аналіз існуючих підходів до обробки ринкових даних, досліджено джерела (зокрема, OLX) та реалізовано систему автоматичного парсингу оголошень із подальшим оновленням датасету. Побудована модель Linear Regression забезпечує простий, але надійний прогноз попиту та орієнтовної вартості квартир на основі вхідних параметрів.

Сформовано функціональний інтерфейс користувача без авторизації, що дозволяє вводити дані, переглядати результати у зручній формі, а також експортувати їх у форматі PDF. Додано модулі для повторного навчання моделі та управління парсингом. Отримана система підтвердила свою стабільність і практичну цінність як інструмент для аналітики ринку нерухомості.

ВИСНОВКИ