

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ВЕТЕРИНАРНОЇ
МЕДИЦИНИ ТА БІОТЕХНОЛОГІЙ ІМЕНІ С.З. ГЖИЦЬКОГО
ФАКУЛЬТЕТ МЕХАНІКИ, ЕНЕРГЕТИКИ ТА ІНФОРМАЦІЙНИХ
ТЕХНОЛОГІЙ
КАФЕДРА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ**

КВАЛІФІКАЦІЙНА РОБОТА

другого (магістерського) рівня вищої освіти

на тему: **«Використання методів машинного навчання для аналізу
ефективності стратегій трейдингу на бінарних опціонах»**

Виконав: здобувач групи Іт-62

Спеціальності 126 «Інформаційні системи та
технології»

(шифр і назва)

Павловський Антон Олександрович

(Прізвище та ініціали)

Керівник: к.т.н., доцент Пташник В. В.

(Прізвище та ініціали)

Рецензент: к. т. н., доцент Сиротюк С. В.

(Прізвище та ініціали)

ДУБЛЯНИ-2025

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ПРИРОДОКОРИСТУВАННЯ
ФАКУЛЬТЕТ МЕХАНІКИ, ЕНЕРГЕТИКИ ТА ІНФОРМАЦІЙНИХ
ТЕХНОЛОГІЙ
КАФЕДРА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Другий (магістерський) рівень вищої освіти
Спеціальність 126 «Інформаційні системи та технології»

ЗАТВЕРДЖУЮ
Завідувач кафедри

(підпис)
д.т.н., професор, Тригуба А. М.
(вч. звання, прізвище, ініціали)
“ ” 202__ року

ЗАВДАННЯ
на кваліфікаційну роботу студенту

Павловського Антона Олександровича

(прізвище, ім'я, по батькові)

1. Тема роботи: «Використання методів машинного навчання для аналізу ефективності стратегій трейдингу на бінарних опціонах»
керівник роботи к.т.н., доцент Пташник Вадим Вікторович
(наук.ступінь, вч. звання, прізвище, ініціали)
затверджені наказом по університету від 28.02.2025 року № 140/к-с.

2. Строк подання студентом роботи 8 грудня 2025 року

3. Вихідні дані до роботи: Jupyter Notebook; історичні дані валютної пари у форматі CSV; мова програмування Python; бібліотеки чисельних обчислень та аналізу даних: numpy, pandas; бібліотеки для візуалізації даних: matplotlib, seaborn; бібліотека машинного навчання: scikit-learn; бібліотека градієнтного бустингу: xgboost; засоби збереження та відтворення моделей: joblib; середовище виконання та керування залежностями Anaconda.

4. Зміст розрахунково-пояснювальної записки (перелік питань, які необхідно розробити): _____

Вступ

1. Теоретичні основи використання методів машинного навчання для аналізу стратегій трейдингу на бінарних опціонах

2. Обґрунтування вибору методів машинного навчання

3. Практична реалізація аналізу ефективності стратегії трейдингу

4. Охорона праці та безпека надзвичайних ситуацій

5. Визначення ефективності

Висновки

Список використаних джерел

Додатки

5. Перелік графічного матеріалу: Ілюстративний матеріал подається у вигляді презентації

6. Консультанти з розділів:

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата		Відмітка про виконання
		завдання видав	завдання прийняв	
1, 2, 3, 5	Пташник В.В., к.т.н., доцент			
4	Городецький І.М., к.т.н., доцент			

7. Дата видачі завдання

3 березня 2025 р.

Календарний план

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів роботи	Відмітка про виконання
1	Теоретичні основи використання методів застосування машинного навчання для аналізу стратегій трейдингу на бінарні опціони	03.03.2025 – 15.04.2025	
2	Обґрунтування вибору методів машинного навчання	16.04.2025 – 31.05.2025	
3	Практична реалізація аналізу ефективності стратегій трейдингу	01.06.2025 – 30.09.2025	
4	Розгляд питань з охорони праці та безпеки у надзвичайних ситуаціях	01.10.2025 – 20.10.2025	
5	Визначення ефективності	21.10.2025 – 15.11.2025	
6	Завершення оформлення розрахунково-пояснювальної записки та презентаційного матеріалу	16.11.2025 – 30.11.2025	
7	Завершення роботи в цілому. Підготовка до захисту кваліфікаційної роботи	01.06.2025 – 08.12.2025	

Здобувач

(підпис) Павловський А. О
(прізвище та ініціали)

Керівник роботи

(підпис) Пташник В. В.
(прізвище та ініціали)

Використання методів машинного навчання для аналізу ефективності стратегій трейдингу на бінарних опціонів.

Павловський А.О. Кафедра інформаційних технологій – Дубляни, ЛНУВМБ імені С.З. Гжицького, 2025.

Кваліфікаційна робота: 66 с. текст. част., 32 рис., 3 табл., 24 джерел.

У ході дослідження проведено аналіз можливостей машинного навчання для оцінки ефективності стратегій трейдингу на бінарних опціонах. Обґрунтовано доцільність застосування ML-алгоритмів для прогнозування цінових рядів та підтримки прийняття рішень торгових стратегіях. Розглянуто та порівняно моделі лінійної регресії, SGD-регресора, Random Forest, Bagging Regressor та XGBoost щодо їх здатності прогнозувати рух ринку та знизити ризики.

Виконано розширене формування ознак, що дає змогу моделі враховувати зміщені значення цін та обсягів, ковзні середні, показники волатильності та доходності. Дані підготовлено шляхом очищення, нормалізації та поділу на навчальні та тестові вибірки. Проведено оцінювання моделей на основі статичних і фінансових метрик, включаючи кумулятивну дохідність.

Оцінено визначення ефективності для застосування розробленої системи з використанням прогнозів моделей машинного навчання. Розроблено рекомендації щодо використання моделей у реальних умовах трейдингу.

Ключова слова: машинне навчання, бінарні опціони, фінансове прогнозування, регресія, ансамблеві методи, технічний аналіз, часові ряди.

Key words: machine learning, binary options, financial forecasting, regression, ensemble methods, technical analysis, time series

ЗМІСТ

ВСТУП.....	6
РОЗДІЛ 1. ТЕОРЕТИЧНІ ОСНОВИ ВИКОРИСТАННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ АНАЛІЗУ СТРАТЕГІЙ ТРЕЙДИНГУ НА БІНАРНИХ ОПЦІОНАХ	8
1.1.Поняття та сутність бінарних опціонів.....	8
1.2.Проблеми ефективності та ризики в торгівлі на фінансових ринках..	10
1.3.Використання технічних індикаторів для аналітичних моделей.....	12
1.4.Використання машинного навчання на фінансових ринках	17
РОЗДІЛ 2. ОБГРУНТУВАННЯ ВИБОРУ МЕТОДІВ МАШИННОГО НАВЧАННЯ ТА МОДЕЛЕЙ ДОСЛІДЖЕННЯ.....	20
2.1.Постановка задачі прогнозування та особливості фінансових часових рядів	20
2.2.Вибір моделей машинного навчання	21
2.3.Вибір метрик оцінювання	22
2.4.Формування ознакового простору	24
2.5.Інструменти реалізації для побудови моделі машинного навчання....	25
РОЗДІЛ 3. ПРАКТИЧНА РЕАЛІЗАЦІЯ АНАЛІЗУ ЕФЕКТИВНОСТІ СТРАТЕГІЇ ТРЕЙДИНГУ	27
3.1.Опис вхідних даних для моделі машинного навчання.....	27
3.2.Попередня обробка та дослідницький аналіз даних.....	29
3.3.Реалізація розширеного ознакового опису фінансових даних.....	32
3.4.Реалізація та налаштування моделей машинного навчання.....	35
3.5.Порівняльний аналіз моделей машинного навчання.....	39
РОЗДІЛ 4. ОХОРОНА ПРАЦІ ТА БЕЗПЕКА У НАДЗВИЧАЙНИХ СИТУАЦІЯХ.....	43
4.1.Особливості безпеки під час роботи персональним комп'ютером	43
4.2.Вимоги безпеки під час роботи персональним комп'ютером	44
4.3.Безпека у надзвичайній ситуації.....	46
4.4.Напрями інформаційного забезпечення охорони праці.....	47
РОЗДІЛ 5. ВИЗНАЧЕННЯ ЕФЕКТИВНОСТІ	49
ВИСНОВКИ.....	53
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	55
ДОДАТКИ.....	57

ВСТУП

Стрімкий розвиток цифрових технологій та зростання доступності фінансових ринків призвели до появи нових форм інвестування, серед яких особливе місце займають бінарні опціони. Незважаючи на простоту механізму відкриття угод, ринок бінарних опціонів характеризується підвищеною волатильністю, високим рівнем невизначеності та значною чутливістю до ринкових шумів. У таких умовах традиційні методи аналізу, засновані на класичних технічних індикаторах або суб'єктивному досвіді трейдера, демонструють недостатню стабільність та точність.

Однією з найбільш перспективних стратегій для підвищення ефективності трейдингу є машинне навчання, яке дозволяє виявляти приховані закономірності у великих масивах фінансових даних, моделювання складні нелінійні залежності та здійснювати короткострокові прогнози з високим рівнем точності. На відміну від класичних статичних моделей, алгоритми машинного навчання здатні адаптовуватися до швидкозмінних ринкових умов, що є важливим у високому частотному та короткостроковому трейдингу, властивому ринку бінарних опціонів.

Актуальність теми магістерської роботи визначає тим, що бінарні опціони, недостатньо дослідженим сегментом у контексті наукового застосування алгоритмічного аналізу. Більшість робіт концентрується на прогнозуванні довгострокових трендів або класичному Forex.

Основною проблемою, що розглядається у роботі, є низька прогнозна здатність традиційних стратегій та обмежені можливості трейдерів щодо обробки великих обсягів ринкової інформації в режимі реального часу. У практиці трейдингу застосовують індикатори на кшталт ковзних середніх, осциляторів або графічних формацій, однак такі методи не враховують складну нелінійну природу рухів ціни та не адаптуються до швидких ринкових змін.

Метою даної роботи є розробка, навчання та оцінювання моделей машинного навчання для прогнозування короткострокового руху ціни та аналізу ефективності торгових стратегій на ринку бінарних опціонів. Для досягнення поставленої мети необхідно вирішити низку взаємопов'язаних завдань: зібрати

та обробити історичні дані, виконати статичних аналіз часових рядів, сформувані інформативні технічні ознаки, підібрати та навчити кілька моделей машинного навчання, здійснити їх порівняння за статичними та фінансовими метриками, а також визначити ефективну модель з точності та потенційної прибутковості.

У кваліфікаційній роботі, застосовують сучасні алгоритми машинного навчання. Оцінювання точності моделей здійснюється на основі стандартних статичних метрик, тоді економічна ефективність торгівлі оцінюється через фінансові метрики.

РОЗДІЛ 1.

ТЕОРЕТИЧНІ ОСНОВИ ВИКОРИСТАННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ ДЛЯ АНАЛІЗУ СТРАТЕГІЙ ТРЕЙДИНГУ НА БІНАРНИХ ОПЦІОНАХ

1.1. Поняття та сутність бінарних опціонів

Бінарні опціони – це опціони, які вимагають тільки передбачення напрямку ціни. Якщо передбачення трейдера є вірним у момент експірації (виконання опціону), він отримає фіксовану виплату. Якщо передбачення трейдера є хибним, трейдер втрачає усі витрачені на купівлю опціону грошові кошти. Бінарні опціони є одними з найпростіших для розуміння принципів їх роботи деривативів. Головною причиною такої доступності є гнучкість та простота даних фінансових інструментів, якими можна торгувати без розуміння постулатів традиційних опціонних стратегій торгівлі. Бінарні опціони зазвичай використовуються як короткостроковий інструмент спекуляції. З одного боку, даний від опціонів дозволяє отримати виграш у відносно короткий період часу, тому популярність бінарних опціонів постійно зростає у різних частинах світу. З іншого боку, їх широке поширення серед непрофесійних трейдерів створює підґрунтя для недоброчесних дій з боку онлайн-платформ, які надають можливості для торгівлі бінарними опціонами [0]. На рисунку 1.1 представлено основні варіанти прогнозування руху ціни базового активу: «In the money» – прогноз справдився; «At the money» - ціна активу впродовж торгів залишилась незмінною; «Out of the money» - прогноз не справдився.



Рисунок 1.1 – Варіанти справдження прогнозів по бінарних опціонах

У сучасних умовах бінарних опціонів розглядають не лише як форма швидких фінансових спекуляцій, але й як інструмент для тестування аналітичних і машинних моделей прогнозування ринкової динаміки. Вони дозволяють перевіряти ефективність різних торгових стратегій на коротких часових проміжках, що робить їх зручними для досліджень у сфері машинного навчання.

Також, бінарні опціони відіграють важливу роль у формуванні навичок управління ризиками. Хоча прибутковість таких операцій фіксована, кожна угода передбачає усвідомлене співвідношення ризику і винагороди. Це дозволяє трейдерам навчатися приймати рішення в умовах невизначеності, аналізувати власну поведінку на ринку та вдосконалювати підходи до прогнозування (Рис. 1.2).

Попри свою популярність, ринок бінарних опціонів є одним з суперечливих. З одного боку, він є доступний і простим способом входу у світ фінансів. З іншого – високий рівень ризику та ймовірність втрати всього капіталу вимагають дисципліни, психологічної стійкості та аналітичного мислення.

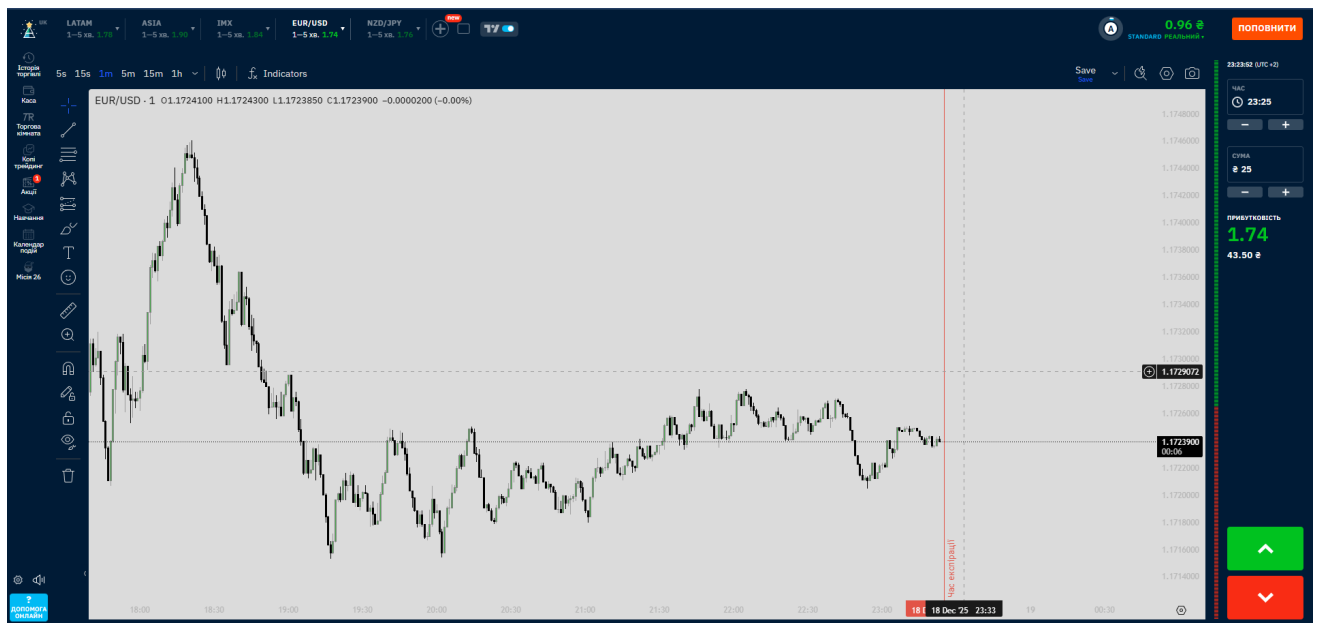


Рисунок 1.2 – Приклад фінансового інструменту платформи Binomo на активі валютного курсу EUR/USD у реальному часі

1.2. Проблеми ефективності та ризики в торгівлі на фінансових ринках

Фінансові ринки – це складна система, яка взаємодіє тисячі учасників: інвестори, трейдери, аналітики, банки, фонди, брокери. Їхні рішення формує ринкову динаміку, що постійно змінюється під впливом економічних, політичних, технологічних та психологічних факторів. Дана торгівля на ринках, поєднає у елементі аналізу, статистики, інтуїції та ризику.

Незважаючи на технологічний прогрес і доступ до величезних масивів даних, ефективність торгівлі залишається однією з головних проблем фінансової діяльності. Ефективність як співвідношення між отриманим результатом та використаними ресурсами. На практиці, досвідчені трейдери потрапляють з певними труднощами, які знижують стабільність результатів. Ознакою даного результату – природна властивість ринку та людський фактор.

З основних проблем ефективності на ринках є інформаційна асиметрія – ситуація, коли ринкова інформація розподілена нерівномірно. Хтось з учасників ринку отримує новини раніше, має доступ до потужних аналітичних інструментів, спираючись на основі публічних даних. В результаті одні трейдери опиняються у вигідному становищі, а решта – приймати рішення з обмеженими можливостями та знаннями.

Інша проблема – висока волатильність фінансових активів. Це інструменти (валюти, акції, індекси), котрі різко змінюють через передбачувані події: економічні звіти, політичні заяви, технічні фактори або чутки. Для трейдингу на бінарному опціоні, де прибуток залежить від точного прогнозу напрямку руху ціни у фіксований час.

Психологічний тиск є одним із найсильніших ворогів ефективності – емоційна нестійкість трейдерів. Кожен з трейдерів має відмінну стратегію, але певний страх втрат, бажання швидко відігратися або надмірна впевненість часто призводять до порушення правил.

Крім людського фактора існують системні недоліки трейдера. Тобто трейдери використовують популярні технічні індикатори, не розуміючи, що їх ефективність обмежена певними ринковими умовами.

Щодо відсутності адаптивності у більшості класичних торгових систем – ринок змінюється, тому з часом торгові системи перестають бути актуальними. Тому при використанні методів машинного навчання здатні підлаштовувати під нові ринкові умови.

Серед основних видів ризиків, класифікують:

- 1) Ринковий ризик – головний ризик, він полягає у можливості зміни ціни активу в несприятливому напрямку.
- 2) Операційний ризик – пов'язані з технічними аспектами: збоями у роботі торгової платформ, затримки котирувань або неправильно обробкою сигналів.
- 3) Психологічний ризик – невпевненість, страх пропустити можливість, надмірна впевненість у власних силах ведуть до порушення торгового плану.
- 4) Системний ризик – глобальні потрясіння, які впливають на весь фінансовий ринок: кризи, інфляційні сплески, геополітичні події.
- 5) Ризик ліквідності – виникає при неможливості відкрити чи закрити позицію за бажаною ціною через низьку активність на ринку.
- 6) Ризик помилки моделі – якщо алгоритм побудований на неповних або застарілих історичних даних, є здатність подання на хибні рішення.

Успішна торгівля неможлива без правильного балансу між прибутком і ризиком. Усі стратегії умовно розподілені на агресивні (з високим потенціалом прибутку, але і високими втратами) та консервативні (з меншою дохідністю, але більш стабільні).

У випадку бінарних опціонів дана проблема має особливу гостроту, оскільки структура виплат фіксована. Наприклад, якщо брокер виплачує у проміжку від 76% до 92% прибутку за виграшну угоду, то кожна помилка коштує 100% ставки. Таким чином, трейдеру необхідно виграти щонайменше 55-60% угод лише для того щоб залишитися в нулі.

Методи машинного навчання здатні виявляти приховані закономірності у великих обсягах історичних даних. За допомогою алгоритмів класифікації, навчають модель визначити ймовірність того, що наступний рух ціни буде вгору чи вниз. Нейронні мережі також запам'ятовують патерни поведінку ринку.

1.3. Використання технічних індикаторів для аналітичних моделей

Технічні індикатори – ключовий інструмент у процесі аналізу фінансових ринків, дозволяють формалізувати поведінку ціни, виявляти тенденції та закономірності, що не завжди помітні при візуальному спостереженні за графіком. В контексті машинного навчання, відіграють роль вхідних ознак, які використовують моделями для навчання прогнозувати напрям руху ціни чи ймовірність прибуткової угоди.

Традиційно технічні індикатори базують на статичних операціях над історичними даними, тобто ціною відкриття, закриття, максимумом, мінімумом та обсягом торгів. Основне завдання полягає у згладжуванні «шуму» ринку та виявленні ключових закономірностей, які можна використовувати для створення аналітичних або торгових стратегій.

У сучасних дослідженнях трейдингу, особливо при використанні методів машинного навчання, дані індикатори виконують функцію кількісного опису ринку, що дозволяє моделі формувати логічні зв'язки між ринковими умовами та результатами угод.

Індикатори тренду – допомагають визначити загальний напрям руху ціни (зростання, спад або боковий рух). З типів поділяють на:

- 1) Проста ковзна середня (Simple Moving Average) – показник відображає вартість активу за певний проміжок часу, демонструє зміни, що сталися (Рис. 1.3) [11].

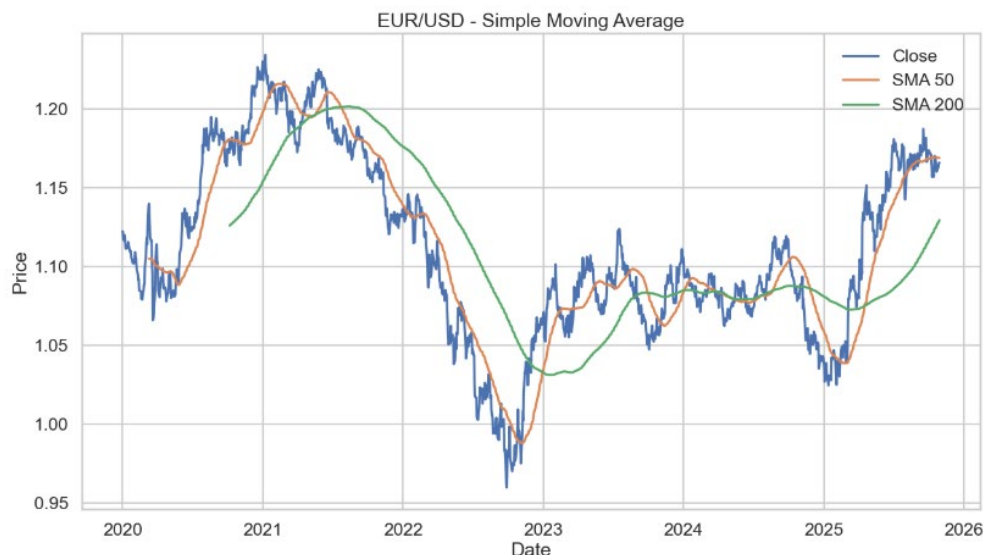


Рисунок 1.3 – Графік ціни EUR/USD з простою ковзною середньою

- 2) Експоненціальна ковзна середня (Exponential Moving Average) – приносить найбільшу користь при використанні короткострокових трейдингових стратегій, адже аналізує дані за останній час [11].
- 3) Ковергенція/дивергенція ковзних середніх (Moving Average Convergence Divergence) – аналізує ковзні середні та підказує, коли, найвірогідніше, почнеться новий тренд [11] (Рис 1.4).

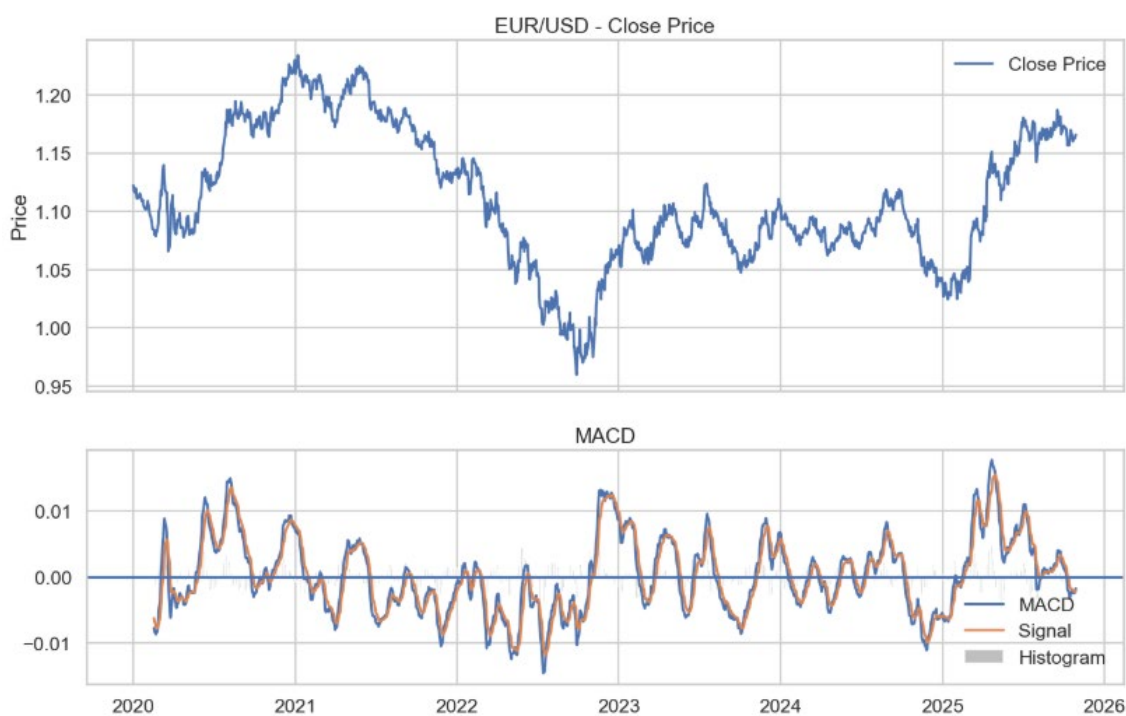


Рисунок 1.4 – Графік ціни з індикатором ковергенцією/дивергенцією ковзних середніх (MACD)

Індикатори імпульсу відображають швидкість зміни ціни та допомагають визначити можливі точки розвороту тренду або перекупленості/перепроданості активу:

- 1) Індекс відносної сили (Relative Strength Index) – показує співвідношення між середніми прибутками і збитками за певний період (Рис 1.5).

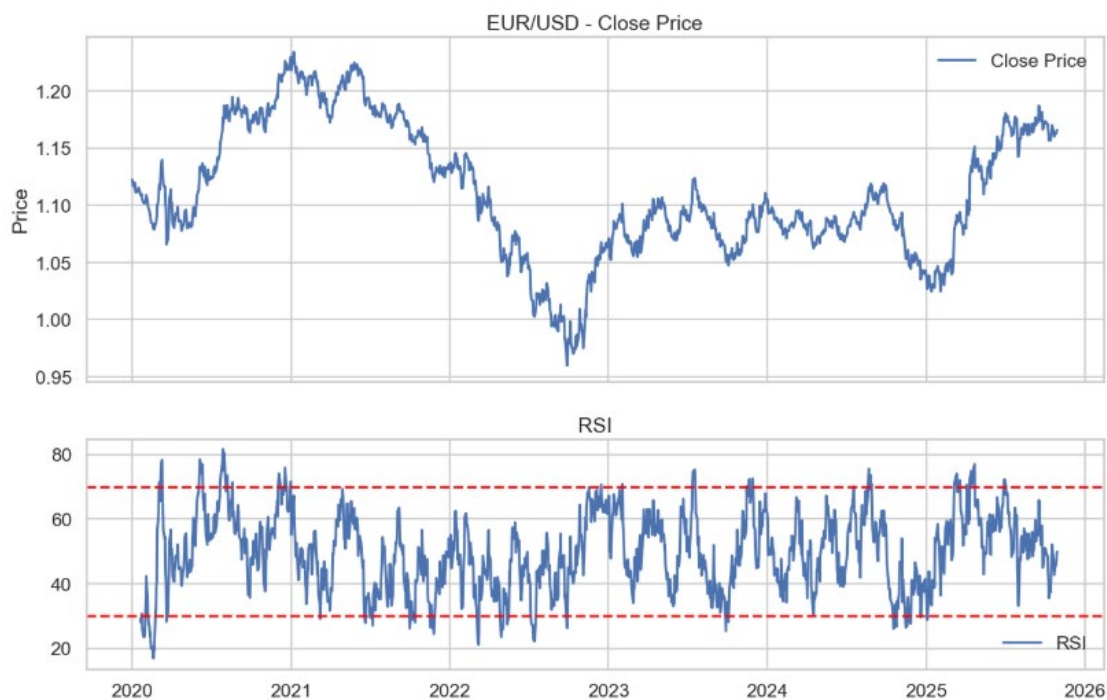


Рисунок 1.5 – Графік ціни EUR/USD з індексом відносної сили (RSI)

2) Стохастичний осцилятор (Stochastic Oscillator) – порівнює актуальну ціну вартістю за певний період. Допомогає відрізнити «бичачий» тренд від «ведмежого» (Рис 1.6) [11]

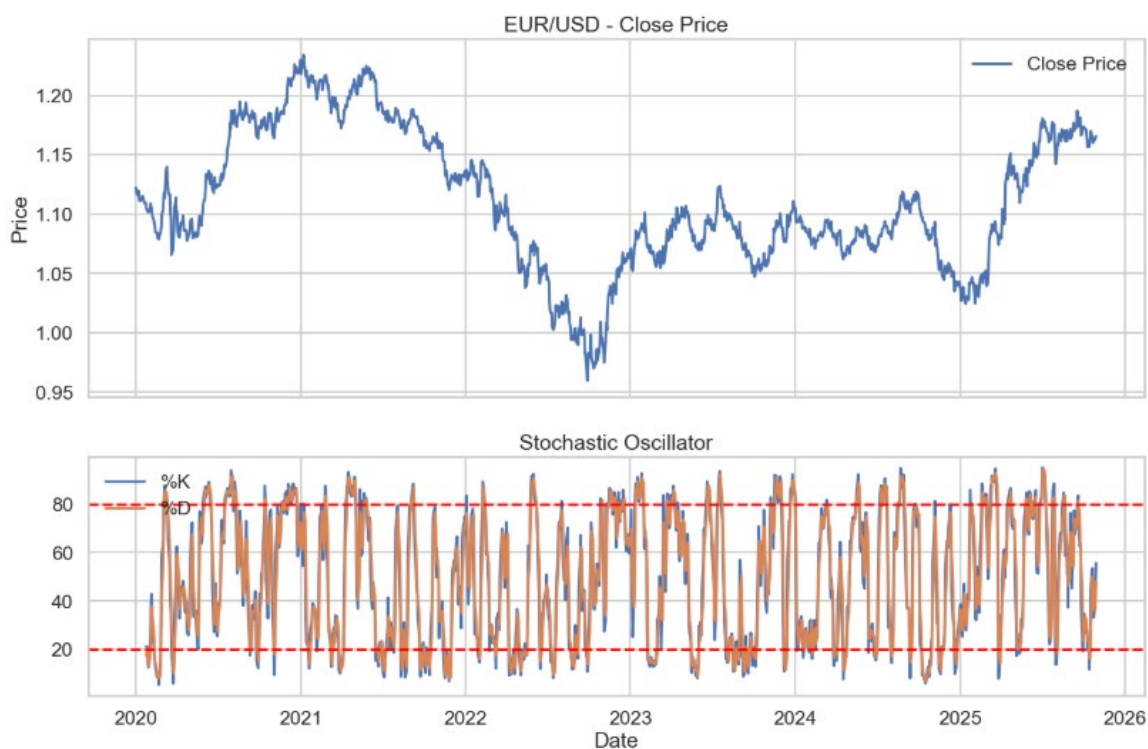


Рисунок 1.6 – Графік ціни EUR/USD з стохастичним осцилятором

3) Індикатор імпульсу (Momentum Indicator) – вимірює зміну ціни між поточним і попереднім періодами (Рис. 1.7).

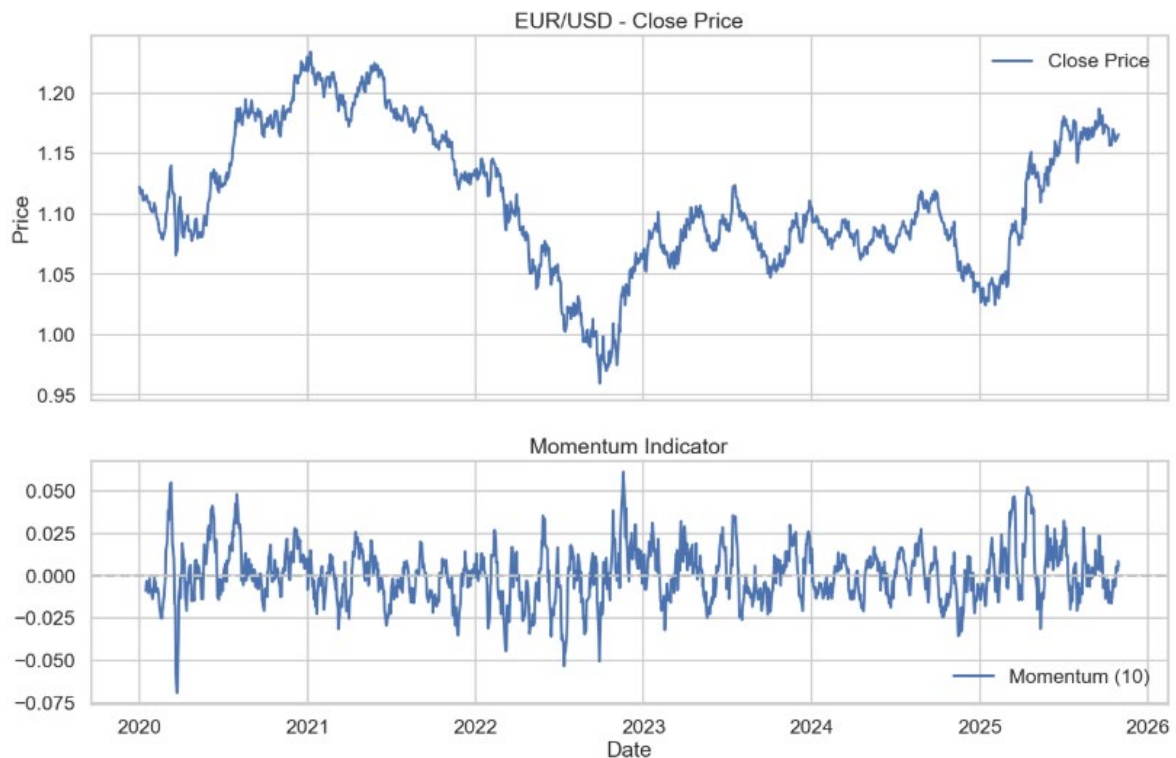


Рисунок 1.7 – Графік ціни EUR/USD з індикатором імпульсу (Momentum)

Індикатор волатильності характеризують ступінь коливання ціни й використовує для оцінки ризиків і налаштування параметрів торгових стратегій:

- 1) Смуги Боллінджера (Bollinger Bands) – формує динамічні межі навколо середньої ціни, відображаючи поточний рівень волатильності (Рис. 1.8).

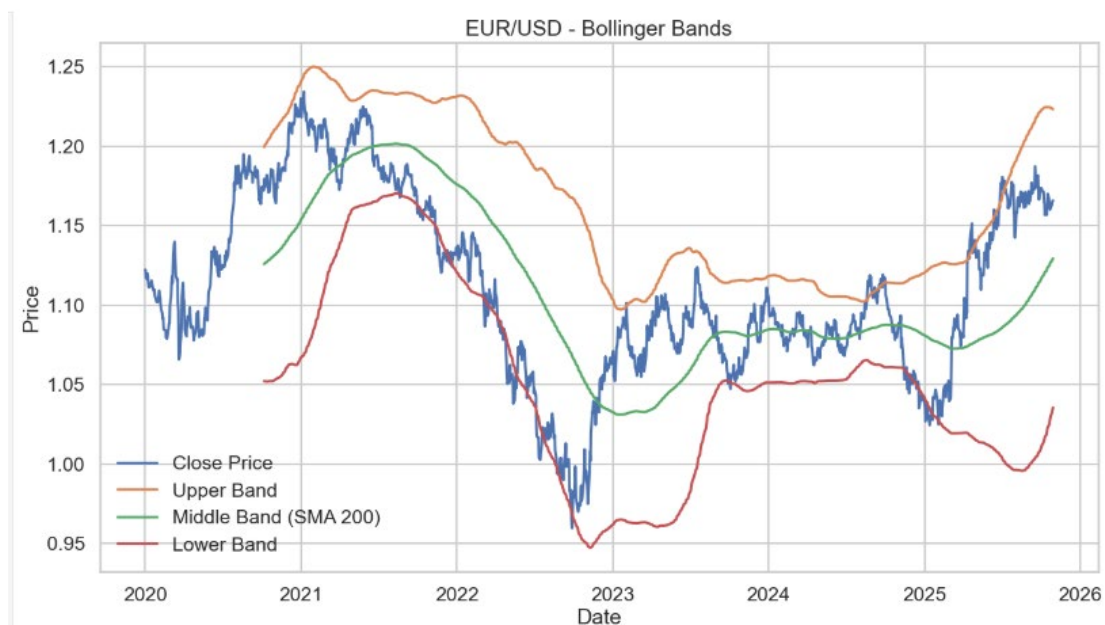


Рисунок 1.8 – Графік ціни EUR/USD з смугою Боллінджера

На рисунку 1.8, наведено індикатор Bollinger Bands, що складає з лінії закриття ціни (Close Price), середньої смуги (Middle Band), яка відповідає простій

ковзній середній, а також верхньої (Upper Band) та нижньої (Lower Band) смуг, що формують динамічний ціновий канал навколо ціни.

- 2) Індикатор волатильності (Average True Range) – показує середній діапазон коливань ціни за певний період, дозволяючи оцінити активність ринку (Рис. 1.9).

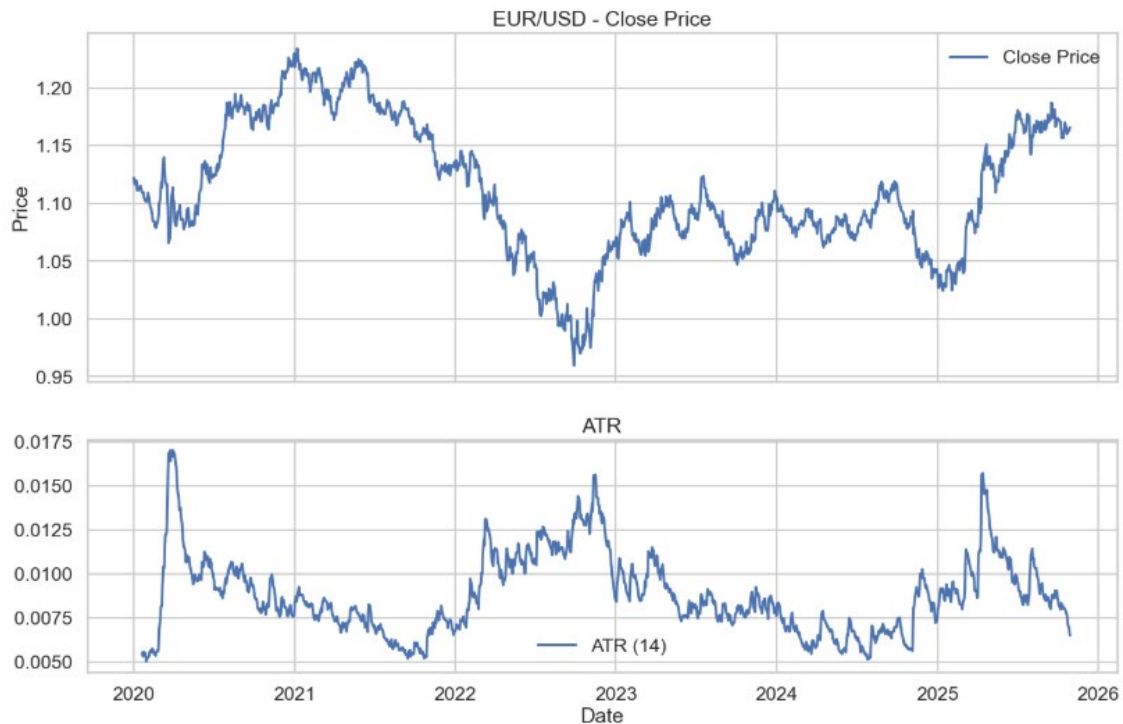


Рисунок 1.9 – Графік свічної ціни з індикатором волатильності (ATR)

Об'ємні індикатори – використовують для підтвердження сили тренду або визначення моментів, коли рух ціни підтримає збільшенням обсягів торгів:

- 1) Балансований обсяг (On-Balance Volume) – демонструє, наскільки зміна вартості активу співвідноситься з кількістю укладених угод, тобто індикатор показує, чи торгівля йде масово, чи вплив чинять епізодичні купівлі/продажі, як показано на рисунку 1.10 [11].
- 2) Обсяг (Volume) – показує, скільки одиниць того чи іншого активу було придбано/продано за певний відрізок часу [11].

Поєднання ціннових та об'ємних характеристик дозволяє аналітичним моделям враховувати не лише напрям руху ринку, а «енергію», що стоїть за ним.

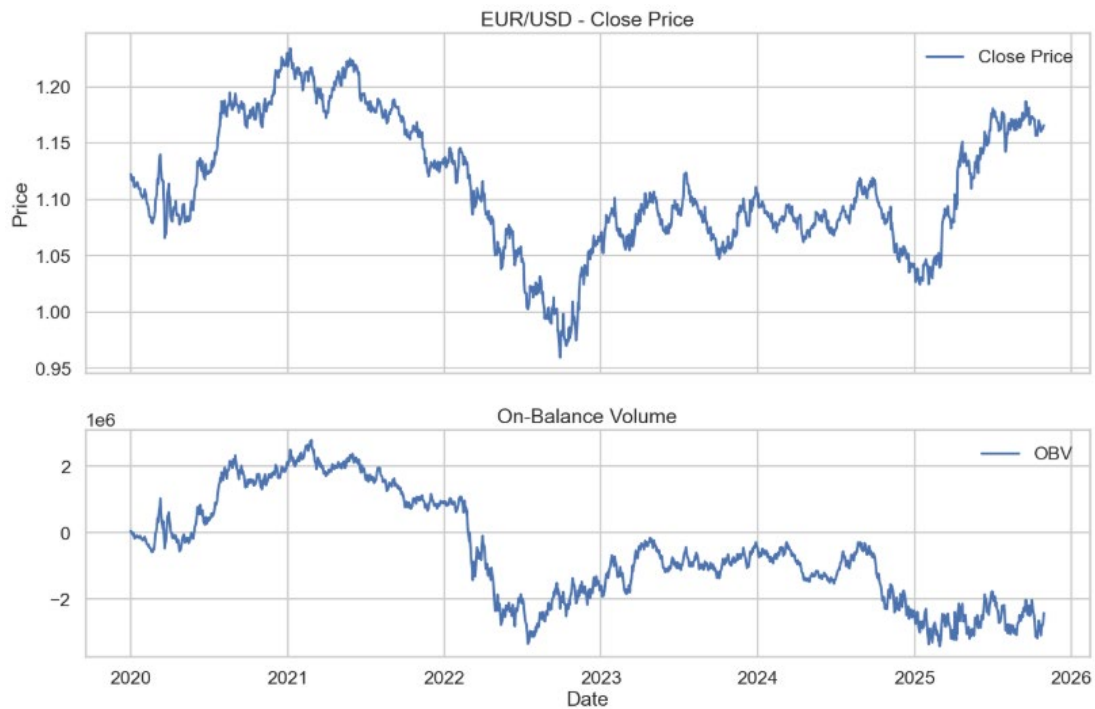


Рисунок 1.10 – Графік ціни EUR/USD з балансованим об’ємом (OBV)

Для підвищення якості моделей для побудови машинного навчання враховують на кореляцію між індикаторами. Надлишок схожих показників призведе до «зашумлення» даних і зниження точності прогнозів. Тому перед тренування моделі проводять аналіз важливості ознак (feature importance) або відбір ознак (feature selection).

1.4. Використання машинного навчання на фінансових ринках

У сучасних умовах розвитку цифрових технологій та зростання обсягів фінансової інформації методи машинного навчання стали важливим інструментом для аналізу, прогнозування та прийняття рішень на фінансових ринках [2].

Суть машинного навчання полягає у створенні моделей, здатних навчитися на основі історичних даних і прогнозувати щодо майбутніх подій. У фінансовій секторі це прогнози ціни активу, визначення моменту для відкриття або закриття позиції, оцінка ризику або ефективності торгової стратегії [3].

З основних методів машинного навчання (ML), що застосовуються у

фінансовому аналізу:

- 1) Лінійна та логістична регресія – використовується для прогнозу цін або дохідності активів, а логістична – для класифікації результатів угод (прибуткова/збиткова) [4].
- 2) Метод опорних векторів (Support Vector Machine, SVM) – використовує для розділення даних на класи. Цей метод ефективний у задачах з великою кількістю параметрів і може працювати навіть тоді, коли кількість спостережень невелика [5].
- 3) Рішення деревами та ансамблеві методи – дозволяють враховувати нелінійні залежності між факторами, працюють із шумними даними та забезпечують добру інтерпретованість [6].
- 4) Нейронні мережі – складають з шарів взаємопов'язаних нейронів і здатні моделювати складні нелінійні залежності [7].
- 5) Методи кластеризації – використовують для групування схожих за поведінкою фінансових активів або виявлення нетипових ринкових ситуацій [8].
- 6) Методи підкріплення (Reinforcement Learning) – цей підхід імітує процес навчання через досвід – система отримує винагороду за правильні рішення і «штрафи» за неправильні [9].

У контексті фінансових ринків, машинне навчання використовує для прогнозування цінових коливань, аналізу ринків та оптимізації торгових стратегій. Основні підходи машинного навчання за типом отримання зворотного зв'язку можна поділити на три категорії: контрольоване навчання, неконтрольоване навчання та навчання з підкріпленням [16].

Неконтрольоване навчання застосовує, коли дані не мають попередньо визначених міток або категорій. Мета даного методу – виявити приховані структури, згрупувати подібні об'єкти або виявити аномальні спостереження [7].

У фінансових ринках цей підхід використовує для кластеризації торгових днів, групування активів за поведінковими патернами або пошуку нетипових рухів ціни. Наприклад, алгоритм k-середніх розподілити дні торгівлі за рівнем волатильності, а метод PCA (Principal Component Analysis) – зменшити розмірність даних, спрощуючи аналіз головних факторів впливу на ціну.

Контрольоване навчання – це процес, у якому алгоритм навчає на основі прикладів із відомими даними та очікуваними результатами. Метою є створення функції, яка узагальнює закономірність між вхідними ознаками та цільовими значенням [7].

Навчання з підкріпленням – це підхід, заснований на принципі винагороди та покарання, де агент взаємодіє із середовищем і вчиться приймати рішення, що максимізує на винагороду.

У бінарних опціонах, метод використовують для побудови автономних торгових агентів, які самі визначають, коли входити або виходити з угоди. Специфіка алгоритму, навчає на основі симульованих торгів, тобто якщо його рішення приводить до прибутку – він отримує «винагороду», якщо до збитку – «програш»

До відомих методів навчання з підкріпленням належать Q-Learning, Deep Q-Network (DQN), Policy Gradient та Actor-Critic.

Таблиця 2.1 – Порівняльна характеристика тип навчання

Тип навчання	Основна ідея	Типові алгоритми	Застосування у трейдингу
Контрольоване	Навчання мічених даних	Лінійна/логістична регресія, Random Forest, XGBoost	Прогноз напрямку ціни, оцінка прибутковості стратегії
Неконтрольоване	Виявлення структур без міток	K-means, DBSCAN, PCA	Аналіз патернів, кластеризація ринку, виявлення аномалій
З підкріпленням	Навчання через винагороди та помилки	Q-learning, DQN, Policy Gradient	Розробка автономних торгових агентів оптимізація стратегій.

РОЗДІЛ 2.

ОБГРУНТУВАННЯ ВИБОРУ МЕТОДІВ МАШИННОГО НАВЧАННЯ ТА МОДЕЛЕЙ ДОСЛІДЖЕННЯ

2.1. Постановка задачі прогнозування та особливості фінансових часових рядів

У даному дослідженні розглядається задача прогнозування динаміки курсу валютної пари EUR/USD на основі історичних ринкових даних із застосуванням методів машинного навчання. З формальної точки зору дана задача належить до класу задач регресійного аналізу часових рядів, де цільовою змінною є значення ціни закриття фінансового інструменту, а вхідними даними – набір ознак, сформованих на основі попередніх значень цін та обсягів торгів.

Фінансові часові ряди мають низку специфічних властивостей, які ускладнюють процес прогнозування та зумовлюють необхідність використання сучасних методів аналізу даних. До таких властивостей, належать нестационарність, висока волатильність, наявність шуму, а також слабо виражені або нестабільні залежності між змінними [12].

Однією з ключових характеристик валютних ринків є їхня динамічність та залежність від великої кількості зовнішніх факторів, зокрема макроекономічних показників, монетарної політики центральних банків, геополітичних подій та поведінки учасників ринку. У результаті ці фактори відображаються у цінових даних у вигляді складних нелінійних закономірностей, які важко описати аналітично, але частково виявлені за допомогою методів машинного навчання [13].

У рамках даної роботи прогнозування розглядається інструменти підтримки прийняття торгових рішень. Якість прогнозу оцінюється у практичній корисності для формування торгових сигналів та отримання фінансового результату [10]. Такий підхід відповідає сучасним дослідженням у галузі алгоритмічної торгівлі, де акцент робиться на поєднанні прогнозної здатності і ризик-орієнтованого аналізу.

Важливою складовою постановки задачі є вибір горизонту прогнозування.

У даному дослідженні використовується коротко- та середньострокового горизонт, що дозволяє враховувати ринкову інформацію та зменшувати вплив довгострокових структурних змін. Такий підхід доцільним для валютних ринків, де довгострокові прогнози часто мають високу невизначеність [15].

Оскільки фінансові часові ряди зазвичай не відповідають припущенням нормальності та лінійності, у роботі застосовується підхід розширення ознакового простору. Він полягає у формуванні додаткових характеристик, що описують трендові, циклічні та волатильності ринку. З точки зору машинного навчання, це дозволяє перетворити вихідний часовий ряд у багатовимірний простір ознак, у якому моделі мають більше можливостей для виявлення прихованих закономірностей [10].

Задача прогнозування курсу валютної пари EUR/USD, у даній роботі формалізується задача побудови та порівняння декількох регресійних моделей машинного навчання, які навчаються на розширеному наборі ознак та оцінюються з використанням як статичних так і фінансових метрик.

2.2. Вибір моделей машинного навчання

Для розв'язання даної задачі прогнозування було використано декілька регресійних моделей, що належать до різних концептуальних підходів машинного навчання.

Лінійна регресія застосовується як базова модель, вона дозволяє оцінити лінійну залежність між сформованими ознаками та цільовою змінною і слугує еталоном для порівняння з більш складними методами. Вона добре підходить для аналізу базових лінійних залежностей та допомагає виявити, які ознаки мають найбільший вплив на прогнозовану ціну. Проте лінійна регресія обмежена у можливості моделювати складні нелінійні взаємозв'язки, які характерні для фінансових часових рядів.

Стохастичний градієнтний спуск (SGD Regressor) був обраний через його здатність ефективно працювати з великими наборами ознак та застосовувати регуляризацию. Регуляризація зменшує перенавчання, де ознаки часто

корельовано між собою. Також, вибрана модель є гнучким і швидким методом, здатним адаптуватися до потокових даних, що дозволяє моделі швидко оновлювати прогнози при надходженні нових ринкових даних.

Bagging Regressor використано через його здатність зменшити дисперсію прогнозів шляхом усереднення результатів кількох базових моделей. Метод Bagging дозволяє підвищити стійкість моделі та надійні результати.

Random Forest Regressor обрано через його здатність моделювати складні нелінійні залежності між ознаками та цільовою змінною. Фінансові ринки часто демонструють нелінійні взаємозв'язки, наприклад, вплив волатильності на короткострокові тренди або залежність від комплексних індикаторів технічного аналізу. Він стійкий до викидів і здатний працювати з різнорідними ознаками, а також автоматично оцінює їхню важливість, що допомагає виділити ключові фактори впливу на ціну.

XGBoost Regressor вибрано як потужний градієнтний ансамблевий метод, який здатний ефективно виявляти складні нелінійні закономірності. Дане використання обґрунтоване здатністю автоматично визначати важливість ознак і демонструвати високу точність у задачах фінансового прогнозування. Також поєднує переваги бустингу та регуляції, що дозволяє моделі адаптуватися до фінансових даних з шумом і високою волатильністю.

2.3. Вибір метрик оцінювання

Оцінювання ефективності моделей машинного навчання здійснюється з використанням як класичних статичних метрик точності прогнозування, так і фінансово-орієнтованих показників, які дозволяють оцінити практичну доцільність застосування моделей у торгових стратегіях.

До групи статичних метрик віднесено:

- 1) Середньоквадратичну помилку (RMSE), яка чутлива до великих відхилень прогнозу від фактичних та дозволяє виявити значні помилки у прогнозі ціни. Вона визначає за математичною формулою:

$$RMSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2. \quad (2.1)$$

- 2) Середню абсолютну помилку (MAE), що забезпечує інтерпретовану оцінку середньої похибки незалежно від напрямку відхилення. Вона визначає за математичною формулою:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|. \quad (2.2)$$

- 3) Коефіцієнт детермінації R^2 , який характеризує частку поясненої дисперсії залежної змінної, пояснену моделлю, та дозволяє оцінити загальну якість апроксимації. Вона визначається за математичною формулою:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}. \quad (2.3)$$

Використання зазначених метрик є стандартною практикою у задачах регресійного аналізу та дозволяє порівнювати різні моделі між собою за рівнем точності прогнозування. Проте у фінансових задачах мінімізація статичної помилки не завжди корелює з фактичною прибутковістю торгової стратегії, оскільки навіть незначні відхилення прогнозу призводять до хибних торгових рішень.

Оскільки мінімізація статичної помилки не завжди корелює з торговою ефективністю, додатково застосовано фінансові метрики, що дозволяють оцінити ефективність моделей, такі як:

- 4) Коефіцієнт Шарпа (Sharpe Ratio), який використовується для оцінки співвідношення середньої доходності до рівня ризику та характеризує ефективність стратегії з урахуванням волатильності результатів. Вона визначає за такою формулою:

$$Sharpe = \frac{E(R_p - R_f)}{\sigma_p}. \quad (2.4)$$

- 5) Максимальну просадку (Max Drawdown) як індикатор ризиковості стратегії. Вона визначає за такою формулою:

$$Max Drawdown = \min_{t \in [1, N]} 1 \frac{V_t - \max(V_{1:t})}{\max(V_{1:t})}. \quad (2.5)$$

- 6) Частку прибуткових прогнозів (Win Rate), яка відображає відсоток успішних торгових рішень та є важливим показником стабільності моделі. Вона визначає за такою формулою:

$$Win Rate = \frac{N_{profit}}{N_{total}}. \quad (2.6)$$

- 7) Точність визначення напрямку руху ціни (Direction Accuracy), характеризує здатність моделі правильно передбачати напрям зміни ціни незалежно від абсолютної величини прогнозу.

$$Direction Accuracy = \frac{1}{N} \sum_{i=1}^N 1[sign(y_i - y_{i-1}) = sign(\bar{y}_i - \bar{y}_{i-1})]. \quad (2.7)$$

Поєднання статичних та фінансових метрик дозволяє здійснити комплексну оцінку моделей машинного навчання, враховуючи як точність прогнозування, так і практичну ефективність у контексті торгових стратегій.

2.4. Формування ознакового простору

Для підвищення інформативності вхідних даних застосовано підхід розширення ознакового простору шляхом генерації технічних та статичних характеристик на основі первинних ринкових даних. До набору ознак включено:

- 1) Лагові значення цін та обсягів.
- 2) Ковзні середні цін та обсягів на різних часових горизонтах.
- 3) Відносні співвідношення між коротко- та довгостроковими середніми.
- 4) Показники волатильності.
- 5) Дохідності за різні періоди та їхні усереднення значення.

Подальше формування ознакового простору ґрунтується на припущенні про наявність часової залежності між поточними та попередніми значеннями фінансових показників. Використання лагових змінних дає змогу моделі враховувати інерційність ринкових процесів. Завдяки цьому забезпечується

відображення короткострокової динаміки ринку у вхідних даних.

Ковзні середні, розраховані на різних часових горизонтах, використовуються для згладжування випадкових коливань та виділення трендових компонентів. Короткострокові ковзні середні відображають швидкі зміни ринкової кон'юнктури, тоді як довгострокові дозволяють оцінити загальний напрям руху ціни. Відносні співвідношення між середніми застосовуються для формалізації моментів зміни ринкових тенденцій, що підвищує чутливість моделей до структурних змін у часовому ряді.

Показники волатильності, стандартні відхилення цін за визначені часові інтервали, включено до ознакового простору з метою врахування рівня ринкового ризику. Волатильність є важливим фактором у фінансовому прогнозуванні, оскільки впливає на стабільність прогнозів та потенційну ефективність торгових стратегій.

Дохідності за різні періоди використовуються як альтернативне представлення цінових змін, що дозволяє зменшити вплив нестанціонарності фінансових часових рядів. Усереднені значення дохідностей сприяють зниженню рівня шуму у даних та покращують здатність моделей машинного навчання до узагальнення.

2.5. Інструменти реалізації для побудови моделі машинного навчання

Реалізація моделей машинного навчання потребує використання спеціалізованого програмного забезпечення та бібліотек, здатних забезпечити обробку великих обсягів фінансових даних, використання статичних операцій та навчання моделей різної складності. Сучасні дослідники у сфері фінансових технологій відзначають, що якість інструментів напряду впливає на точність моделі та швидкості її розробки та експериментування.

З платформ для наукових обчислень та машинного навчання є Python, який упродовж останніх років закріпився як стандарт у сфері фінансового аналізу та алгоритмічного трейдингу. Дана популярність обумовлена зручністю роботи з даними, широкою підтримкою у спільності та наявністю великої кількості

спеціалізованих бібліотек. Python дозволяє інтегрувати машинне навчання, візуалізацію, роботу з фінансовими API-запитами та моделювання торгових стратегій для проведення повноцінних експериментів.

Для аналітики фінансових часових рядів ключову роль відіграє бібліотека Pandas, яка забезпечує інструменти для роботи зі структурованими даними, включно з часовими індексами, очищенням даних, ресемплінгом, агрегацією та попередньою обробкою. Бібліотека Numpy призначена для виконання чисельних обчислень, лінійної алгебри та векторизованих операцій.

Для побудови та навчання моделей машинного навчання використовують Scikit-learn – це бібліотека, яка містить реалізації класичних алгоритмів, включаючи лінійну регресію, метод опорних векторів, дерева рішень, ансамблеві методи та метрики оцінювання [19].

На випадок побудови складних моделей використовують бібліотеки TensorFlow та PyTorch, які забезпечують високий рівень гнучкості й оптимізацію обчислень на графічних процесорах.

Для візуалізації результатів аналізу застосовують Matplotlib та Seaborn, що дає можливість будувати графіки часових рядів, теплові карти кореляцій, криві навчання моделей та порівняльні діаграми ефективності. Візуальна аналітика дозволяє інтерпретувати поведінку алгоритмів і виявляти особливості ринкової структури [20].

Програмне забезпечення Jupyter Noterbook забезпечує інтерактивний формат роботи, де код й пояснення, формули та результати досліджень можна об'єднати в єдиному середовищі. Для фінансового аналізу дозволяє швидко тестувати гіпотези, візуалізувати поведінку індикаторів і проводити експерименти з різними моделями [21].

РОЗДІЛ 3.

ПРАКТИЧНА РЕАЛІЗАЦІЯ АНАЛІЗУ ЕФЕКТИВНОСТІ СТРАТЕГІЇ ТРЕЙДИНГУ

3.1. Опис вхідних даних для моделі машинного навчання

У межах даного дослідження для побудови моделей машинного навчання використано історичні валютної пари EUR/USD, яка є однією з найбільш ліквідних та репрезентованих на світовому валютному ринку. Вибір саме цієї пари зумовлений високою стабільністю торгів, значними обсягами операцій та наявністю довготривалої історії котирувань.

Історичні дані валютної пари EUR/USD, було взято з відкритих фінансових ресурсів, де часто оновлюють щоденні котирування через форекс-біржах.

Часовий період дослідження охоплює з 2020 по 2025 року, що дозволяє врахувати стабільні ринкові умови. Дані збережено у форматі CSV для подальшої обробки та аналізу у середовищах машинного навчання, як показано на рисунку 3.1.

Додатково проведено первинний описовий аналіз даних, який включає побудову графіків динаміки цін, розрахунок основних статичних характеристик та візуалізацію розподілу показників. Даний підхід дозволив оцінити загальні тенденції ринку, виявити можливі аномалії та підтвердити доцільність подальшого застосування машинного навчання для прогнозування курсу валютної пари.

Структура набору даних включає п'ять основних фінансових показників:

- 1) Open – ціна відкриття.
- 2) High – максимальна ціна.
- 3) Low – мінімальна ціна.
- 4) Close – ціна закриття.
- 5) Volume – обсяг торгів.

```
[3]: eu = pd.read_csv('eu_1.csv', index_col=0, parse_dates=True, skipinitialspace=True)
     eu.drop('date', axis='columns', inplace=True)
     eu.head(10)
```

```
[3]:
```

	open	high	low	close	volume
2000-01-03	1.0073	1.0278	1.0054	1.0246	6642
2000-01-04	1.0243	1.0340	1.0213	1.0292	7339
2000-01-05	1.0293	1.0402	1.0284	1.0326	6570
2000-01-06	1.0325	1.0415	1.0272	1.0330	7223
2000-01-07	1.0329	1.0332	1.0260	1.0298	5689
2000-01-10	1.0286	1.0305	1.0211	1.0251	6197
2000-01-11	1.0253	1.0341	1.0246	1.0336	6637
2000-01-12	1.0337	1.0370	1.0261	1.0308	6877
2000-01-13	1.0309	1.0316	1.0237	1.0260	6176
2000-01-14	1.0261	1.0275	1.0114	1.0122	6176

Рисунок 3.1 – Код та виведення результату історичних даних

Частота спостережень – щоденна, що є оптимальною для середньострокового прогнозування та аналізу трендів.

Мета аналізу даних включав побудову графіка ціни закриття та розрахунок базових статичних характеристик. Побудова графіку ціни закриття валютної пари EUR/USD показано на рисунку 3.2.



Рисунок 3.2 – Графік фактичної ціни валютної пари EUR/USD

Для детального аналізу розподілу вхідних фінансових показників та виявлення можливих аномальних значень було застосовано метод box plot (діаграма розмаху). Даний метод дозволяє оцінити медіану, міжквартильний діапазон та наявність потенційних викидів у значеннях цін відкриття, максимуму, мінімуму та закриття валютної пари EUR/USD.

Побудовано box plot для показників Open, High, Low, Close, продемонстрували відносно симетричний розподіл цінових даних без критичних аномалій, про це показано на рисунку 3.3.

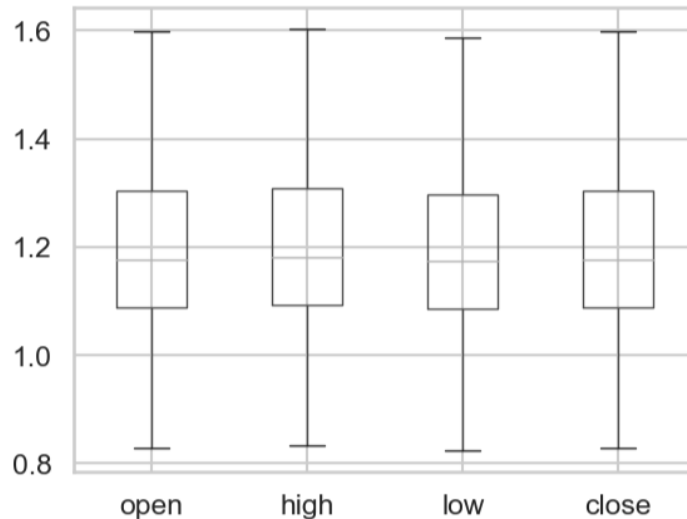


Рисунок 3.3 – Діаграма розмаху з ціновими значеннями

Також приділено аналізу показника обсягу торгів (Volume), який характеризується значною варіативністю та наявністю екстремальних значень. За результатами даного аналізу виявлено аномально великі та малі значення обсягів (Рис. 3.4).

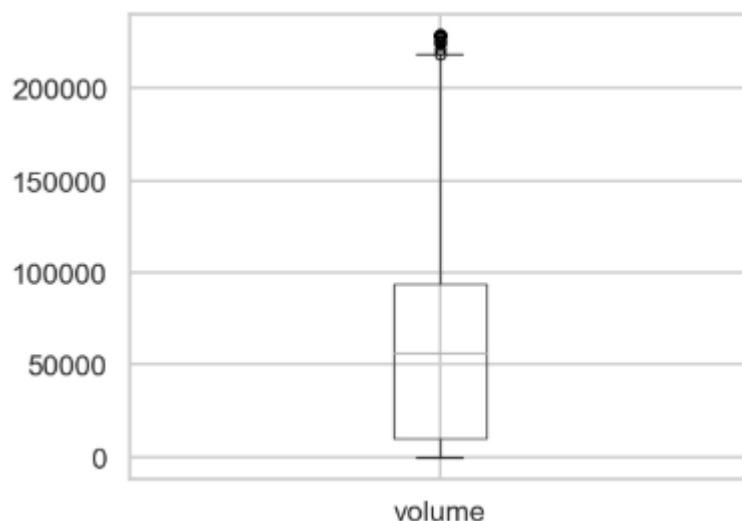


Рисунок 3.4 – Діаграма розмаху з об'ємним показником

3.2. Попередня обробка та дослідницький аналіз даних

Перед використанням даних у моделях машинного навчання було виконано комплекс процедур попередньої обробки, які спрямованих на підвищення якості

та стабільності навчання моделей.

На початковому етапі здійснено перевірку набору даних на наявність пропущених значень. Виявлені пропуски було усунено шляхом видалення відповідних спостережень або застосування методів імпутації, що дозволило мінімізувати ризик спотворення результатів навчання.

Далі проведено аналіз та усунення викидів для показника обсягу торгів. З цією метою застосовано метод обмеження екстремальних значень, який полягає у встановленні верхньої і нижньої порогів на основі статичних характеристик розподілу.

Наступним кроком виконано нормалізацію числових ознак із використанням стандартного масштабування за допомогою методу `StandardScaler`, що забезпечує приведення ознак до єдиного масштабу з нульовим середнім значенням та одиничним стандартним даного відхилення. Реалізацію даного етапу наведено на рисунку 3.6.

```
#initiating standard scaler
scaler = StandardScaler()

#fit the scaler in training features
scaler.fit(X_train)

#Rescale both sets using the trained scaler
X_scaled_train = scaler.fit_transform(X_train)
X_scaled_test = scaler.transform(X_test)
```

Рисунок 3.6 – Код нормалізації ознак за методом `StandardScaler`

Заключним етапом підготовки даних стало розділення вибірки на навчальну та тестову з урахуванням часової структури фінансових спостережень. Даний підхід виключає ефект витoku інформації та забезпечує об'єктивність оцінювання прогностичних моделей. Навчальна вибірка сформована на основі історичних даних за період з 1 січня 2020 року по 31 березня 2024 року, як показано на рисунку 3.7.

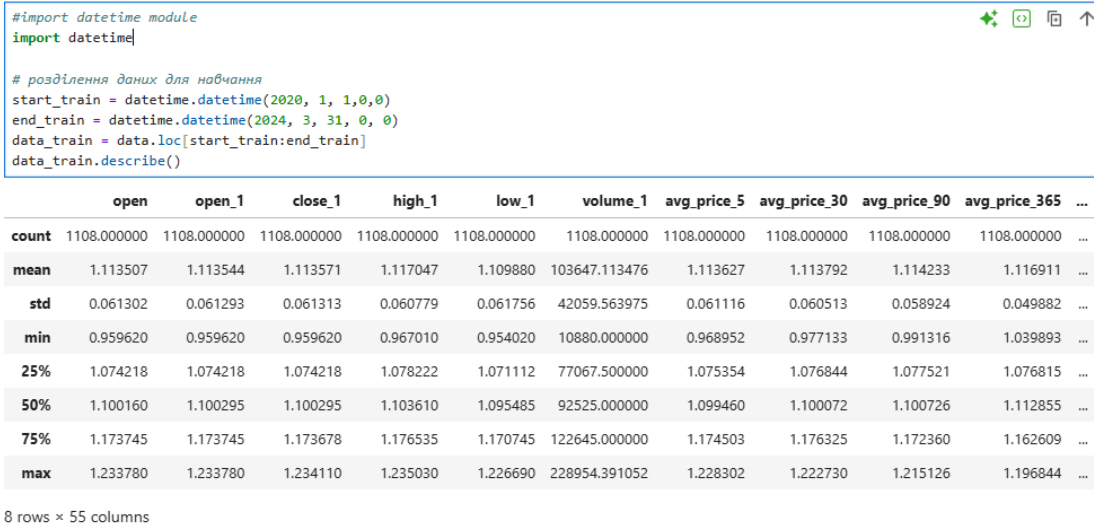


Рисунок 3.7 – Код та результат з навчальної вибірки даних

Використання індексації за часовими мітками, зазначено на рисунку 3.7, дозволяє коректно відібрати спостереження без порушення часової структури ряду.

Метод `describe()` застосовується для отримання описової статистики навчальної вибірки, що дозволяє оцінити розподіл ознак, їхні середні значення, дисперсію.

Тестова вибірка охоплює період з 1 квітня 2024 року по 29 жовтня 2025 року та використовується виключно для оцінювання якості прогнозування моделей (Рис. 3.8).

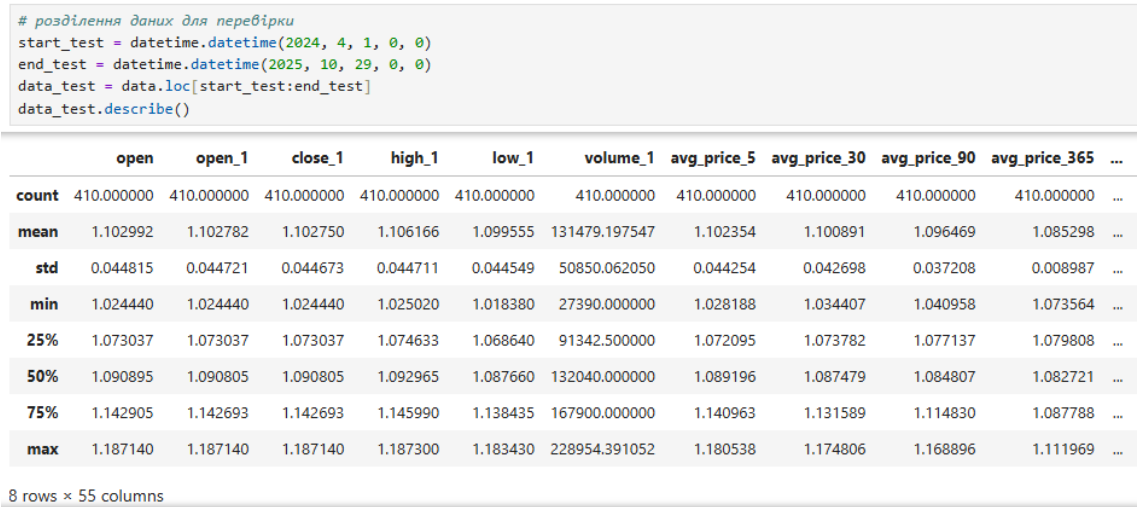


Рисунок 3.8 – Код та результат з тестової вибірки даних

3.3. Реалізація розширеного ознакового опису фінансових даних

Для реалізації розширеного ознакового простору використано Python та бібліотеки pandas, numpy, scikit-learn. У даній роботі для побудови ознакового простору було реалізовано функцію `generate_features`, яка автоматично генерує та похідні ознаки на основі історичних даних по валютній парі EUR/USD.

Першим кроком створено базові ознаки, що включають поточні значення ціни відкриття, а також лагові значення цін та обсягів торгів за попередній день. Лагові ознаки дозволяють моделі враховувати короткострокові зміни ринку та забезпечують основу обчислення більш складних характеристик. До цієї групи входять такі показники: ціна відкриття та закриття, максимум і мінімум, обсяг торгів попереднього дня, зазначено на рисунку 3.9.

```
df_new = pd.DataFrame()

# 6 оригінальних значень
df_new['open'] = df['open']
df_new['open_1'] = df['open'].shift(1)
df_new['close_1'] = df['close'].shift(1)
df_new['high_1'] = df['high'].shift(1)
df_new['low_1'] = df['low'].shift(1)
df_new['volume_1'] = df['volume'].shift(1)
```

Рисунок 3.9 – Код базових та лагових ознак за певними показниками

Для оцінки трендів ринку розраховано ковзні середні закритих цін за різні часові періоди – 5, 30, 90, 365 дні. Короткострокові середні дозволяють відслідковувати оперативні зміни, середньострокові та довгострокові – тенденції ринку на більш тривалих проміжках.

```
# 50 оригінальних значень
# ковзні середні цін
df_new['avg_price_5'] = df['close'].rolling(window=5).mean().shift(1)
df_new['avg_price_30'] = df['close'].rolling(window=30).mean().shift(1)
df_new['avg_price_90'] = df['close'].rolling(window=90).mean().shift(1)
df_new['avg_price_365'] = df['close'].rolling(window=365).mean().shift(1)
```

Рисунок 3.10 – Код до ковзні середні цін

На основі цих середніх додатково сформовано співвідношення між ними, наприклад, відношення середньої ціни за 5 днів до середньої ціни за 30 днів.

```

# average price ratio
df_new['ratio_avg_price_5_30'] = df_new['avg_price_5'] / df_new['avg_price_30']
df_new['ratio_avg_price_905_'] = df_new['avg_price_5'] / df_new['avg_price_90']
df_new['ratio_avg_price_5_365'] = df_new['avg_price_5'] / df_new['avg_price_365']
df_new['ratio_avg_price_30_90'] = df_new['avg_price_30'] / df_new['avg_price_90']
df_new['ratio_avg_price_30_365'] = df_new['avg_price_30'] / df_new['avg_price_365']
df_new['ratio_avg_price_90_365'] = df_new['avg_price_90'] / df_new['avg_price_365']

```

Рисунок 3.11 – Код до співвідношення середніх цін

Для оцінки активності ринку використано ковзні обсяги торгів за ті ж періоди. Агреговані показники обсягів дозволяють моделі враховувати зміну інтересу учасників ринку. Додатково сформовано співвідношення середніх обсягів між різними періодами, що дозволяє виділяти періоди підвищеної або зниженої ліквідності (Рис. 3.12).

```

# ковзні середні обсяги
df_new['avg_volume_5'] = df['volume'].rolling(window=5).mean().shift(1)
df_new['avg_volume_30'] = df['volume'].rolling(window=30).mean().shift(1)
df_new['avg_volume_90'] = df['volume'].rolling(window=90).mean().shift(1)
df_new['avg_volume_365'] = df['volume'].rolling(window=365).mean().shift(1)

# співвідношення середніх обсягів
df_new['ratio_avg_volume_5_30'] = df_new['avg_volume_5'] / df_new['avg_volume_30']
df_new['ratio_avg_volumee_5_90'] = df_new['avg_volume_5'] / df_new['avg_volume_90']
df_new['ratio_avg_volume_5_365'] = df_new['avg_volume_5'] / df_new['avg_volume_365']
df_new['ratio_avg_volume_30_90'] = df_new['avg_volume_30'] / df_new['avg_volume_90']
df_new['ratio_avg_volume_30_365'] = df_new['avg_volume_30'] / df_new['avg_volume_365']
df_new['ratio_avg_volume_90_365'] = df_new['avg_volume_90'] / df_new['avg_volume_365']

```

Рисунок 3.12 – Код до ковзних обсягів торгів та їх співвідношення

Для оцінки коливань ринку обчислено стандартне відхилення закритих цін та обсягів за аналогічні періоди. Стандартне відхилення цін відображає волатильність ринку, а стандартне відхилення обсягів – коливання активності інвесторів. Також було сформовано співвідношення стандартних відхилень між різними періодами, що дозволяє моделі оцінювати відносну волатильність у коротко- та довгостроковій перспективі (Рис 1.13).

```
# стандартне відхилення цін
df_new['std_price_5'] = df['close'].rolling(window=5).std().shift(1)
df_new['std_price_30'] = df['close'].rolling(window=30).std().shift(1)
df_new['std_price_90'] = df['close'].rolling(window=90).std().shift(1)
df_new['std_price_365'] = df['close'].rolling(window=365).std().shift(1)

# співвідношення стандартного відхилення цін
df_new['ratio_std_price_5_30'] = df_new['std_price_5'] / df_new['std_price_30']
df_new['ratio_std_price_5_90'] = df_new['std_price_5'] / df_new['std_price_90']
df_new['ratio_std_price_5_365'] = df_new['std_price_5'] / df_new['std_price_365']
df_new['ratio_std_price_30_90'] = df_new['std_price_30'] / df_new['std_price_90']
df_new['ratio_std_price_30_365'] = df_new['std_price_30'] / df_new['std_price_365']
df_new['ratio_std_price_90_365'] = df_new['std_price_90'] / df_new['std_price_365']
```

Рисунок 3.13 – Код до стандартного відхилень цін та їх співвідношення

Для кожного торгового дня було розраховано дохідність на основі ціни закриття за різні періоди.

Даний розрахунок дохідності виконувався як відносна зміна ціни закриття поточного дня порівняно з відповідною ціною у минулому. Для уникнення витоку інформації всі значення дохідності було зсунуто на один часовий крок, що забезпечує коректність використання ознак у задачах прогнозування (Рис. 3.14).

```
# дохідність
df_new['return_1'] = ((df['close'] - df['close'].shift(1)) / df['close'].shift(1)).shift(1)
df_new['return_5'] = ((df['close'] - df['close'].shift(5)) / df['close'].shift(5)).shift(1)
df_new['return_30'] = ((df['close'] - df['close'].shift(21)) / df['close'].shift(21)).shift(1)
df_new['return_90'] = ((df['close'] - df['close'].shift(63)) / df['close'].shift(63)).shift(1)
df_new['return_365'] = ((df['close'] - df['close'].shift(252)) / df['close'].shift(252)).shift(1)
```

Рисунок 3.14 – Код до розрахунку дохідності

Для згладжування короткострокових коливань та виявлення загальної динаміки змін було додатково обчислено ковзні середні значення одноденної дохідності. Було сформовано ковзні середні дохідності за різні періоди: короткостроковий період, середньостроковий період, довгостроковий період.

```
# ковзні середні дохідності
df_new['moving_avg_5'] = df_new['return_1'].rolling(window=5).mean()
df_new['moving_avg_30'] = df_new['return_1'].rolling(window=30).mean()
df_new['moving_avg_90'] = df_new['return_1'].rolling(window=90).mean()
df_new['moving_avg_365'] = df_new['return_1'].rolling(window=365).mean()
```

Рисунок 1.15 – Код до ковзні середні дохідності

На завершальному етапі практичної реалізації до сформованого набору ознак додано цільову змінну, яка відповідає значенню ціни закриття фінансового інструменту.

Після формування всіх ознак виконано очищення даних шляхом видалення рядків із пропущеними значеннями, які виникли внаслідок використання ковзних вікон та лагових операцій.

Для перевірки правильності формування ознак проаналізовано перші рядки отриманого набору даних на рисунку 3.16. Перші рядки датафрейму містять повний набір базових, похідних та цільових змінних.

	open	open_1	close_1	high_1	low_1	volume_1	avg_price_5	avg_price_30	avg_price_90	avg_price_365
2000-12-21	0.9094	0.8952	0.9094	0.9100	0.8947	8345.0	0.89786	0.876148	0.864083	0.923226
2000-12-22	0.9164	0.9094	0.9165	0.9171	0.9059	8706.0	0.90262	0.879548	0.864751	0.922779
2000-12-25	0.9230	0.9164	0.9232	0.9262	0.9143	8181.0	0.90812	0.883533	0.865386	0.922345
2000-12-26	0.9255	0.9230	0.9260	0.9280	0.9197	2356.0	0.91414	0.887719	0.866062	0.921920
2000-12-27	0.9308	0.9255	0.9306	0.9327	0.9234	6139.0	0.92114	0.891414	0.866879	0.921527

5 rows × 55 columns

Рисунок 3.16 – Виведення результату для перевірки правильності формування ознак

3.4. Реалізація та налаштування моделей машинного навчання

У практичної реалізації було використано декілька моделей машинного навчання, що представляють різні підходи до регресійного аналізу. Для дослідження включено: базові лінійні моделі, ансамблеві методи та градієнтний бустинг. Основною метою цього етапу досягнення високої точності прогнозування та оцінка практичної придатності кожної моделі для використання у торгових стратегіях.

Усі моделі реалізовано в середовищі Jupyter Notebook з мовою програмування Python з використанням бібліотек scikit-learn та xgboost. Навчання

здійснювалося на передньому навчальному наборі даних, а оцінювання – на відкладеній тестовій вибірці з дотриманням часової послідовності спостережень.

Першою реалізованою моделлю машинного навчання є лінійна регресія. Вона передбачала на апроксимацію залежності між вхідними ознаками та цільовою змінною у вигляді лінійної комбінації.

Для реалізації лінійної моделі використано клас `LinearRegression` з бібліотеки `sklearn`, зазначено у Додатку В.

Перед навчанням моделі вхідні ознаки були нормалізовані, що дозволяє уникнути домінування окремих змінних та покращує числову стабільність алгоритму. Навчання даної моделі здійснювалося на масштабному навчальному наборі. Після навчання було отримано прогнози на тестовій вибірці.

Візуалізовано за допомогою бібліотеки `matplotlib`, графік фактичної ціни та прогнозованої ціни за даною моделлю (Рис. 3.17).

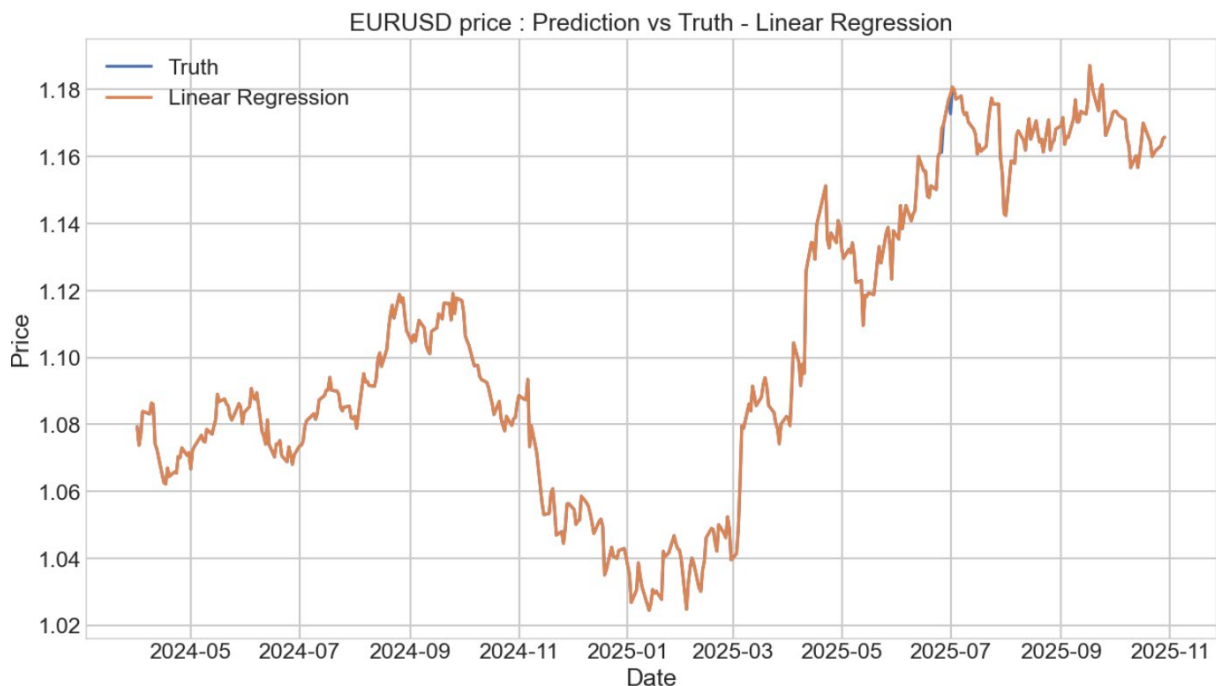


Рисунок 3.17 – Графік прогнозу та поточної ціни на основі лінійної регресії

Наступним моделі реалізації – Стохастичний градієнтний спуск. Особливістю алгоритмів на основі `SGD Regressor` є висока чутливість до масштабів ознак.

Для підвищення точності прогнозування застосовано `GridSearchCV`, що дозволяє автоматично підібрати оптимальну комбінацію гіперпараметрів шляхом крос-валідації.

У процесі оптимізації використовувалися такі параметри:

- `penalty` – тип регуляризації;
- `alpha` – коефіцієнт регуляризації, який контролює складність моделі;
- `eta0` – початковий темп навчання.

Результатом крос-валідації визначено оптимальний набір параметрів.

Після навчання оптимізованої моделі отримано прогнози на тестовій вибірці. Візуалізовано на основі прогнозу ціни за допомогою моделю SGD Regressor (Рис 3.18).

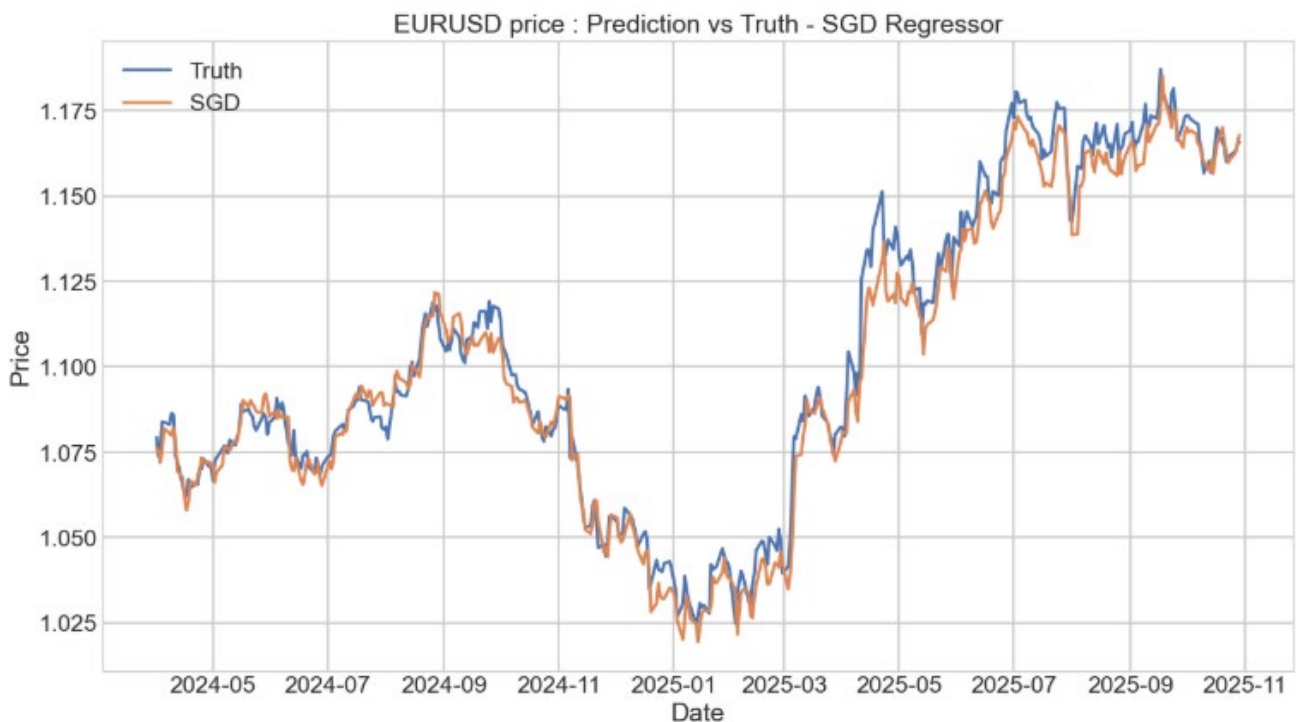


Рисунок 3.18 – Графік прогнозу та поточної ціни на основі стохастичного градієнтного спуску

Найбільш складною та потужною моделлю – XGBoost Regressor. На відміну від лінійних моделей, дана модель автоматично враховує взаємозв'язки між ознаками.

Для побудови моделі використано клас `XGBRegressor`. Дане навчання здійснювалося на навчальній вибірці без додаткового масштабування, оскільки деревоподібні алгоритми не чутливі до масштабу ознак. Після визначення оптимальних параметрів сформовано модель. Також модель навчалася на тестовій вибірці, з подальшою на отримання прогнози для тестового набору даних, зазначено у Додатку В.

Візуалізовано на основі фактичної ціни та прогнозу моделі XGBoost,

представлено на рисунку 3.19.

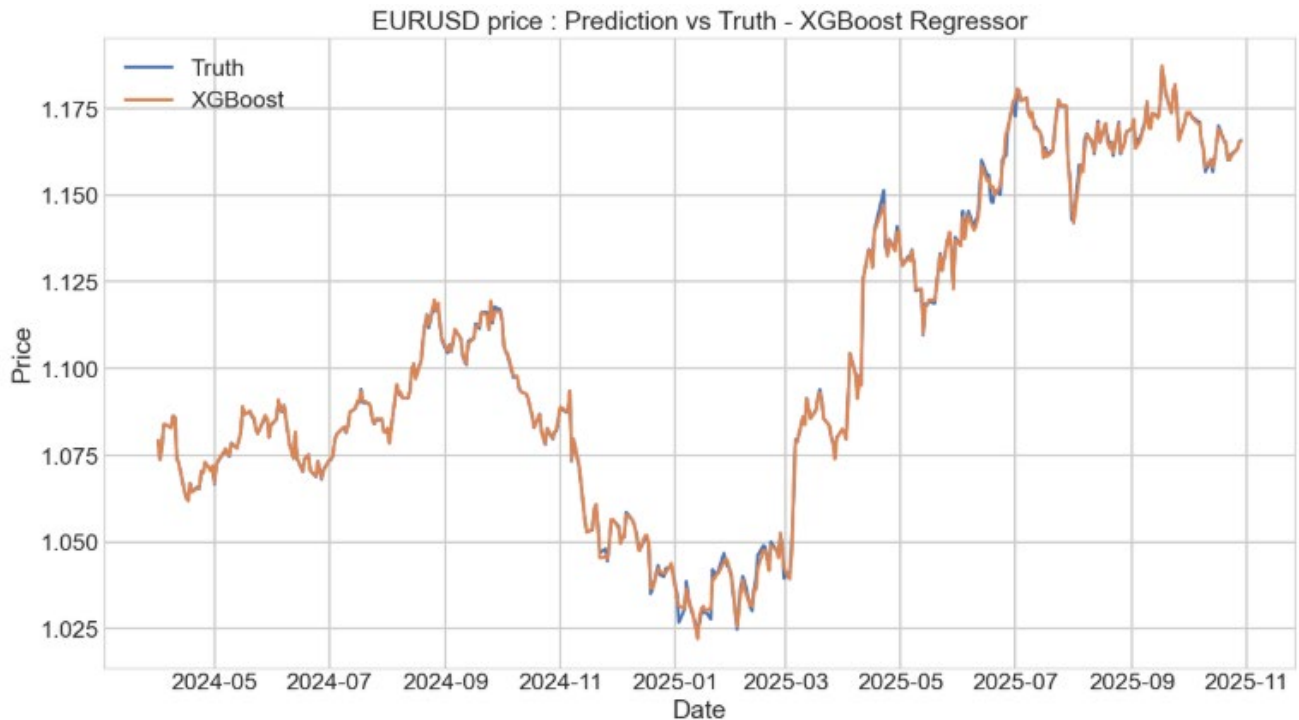


Рисунок 3.19 – Графік прогнозу та поточної ціни на основі XGBoost

Для підвищення стабільності прогнозів та зменшення впливу шуму фінансового ринку в роботі застосовано ансамблеві методи машинного навчання – Bagging Regressor та Random Forest Regressor.

Bagging Regressor реалізовано з використанням лінійної регресії. Модель була ініціалізована з такими параметрами:

- `estimator=lin` – базова модель лінійної регресії;
- `n_estimators=100` – кількість базових моделей в ансамблі;
- `oob_score=True` – використання out-of-bag оцінки;
- `n_jobs=-1` – паралельне використання обчислень;

Навчання моделі здійснювалося на масштабованих даних з використанням 100 базових оцінювачів. Додатково застосовано механізм Out-of-Bag оцінювання, який дозволив оцінити узагальнювальну здатність моделі без виділення окремої валідаційної вибірки.

Однією з основних ансамблевих моделей – Random Forest Regressor. У реалізації для даної моделі виконано підбір основних гіперпараметрів за допомогою GridSearchCV. Оптимізувалися наступними гіперпараметри:

- `max_depth` – максимальна глибина дерева, що контролює складність

моделі;

- `min_sample_split` – мінімальна кількість спостережень для поділу вузла.

Аналогічним чином, після завершення процедури оптимізації обрано найкращу модель. Модель навчалася на тренувальній вибірці та використовувалася для прогнозування значень цільової змінної на тестових даних, зазначено у Додатку В.

Візуалізовано двох прогнозованих цін моделей, а саме Bagging Regressor та Random Forest Regressor та поточних значень цін на основі історичних даних валютного курсу EUR/USD, представлено на рисунку 3.20.

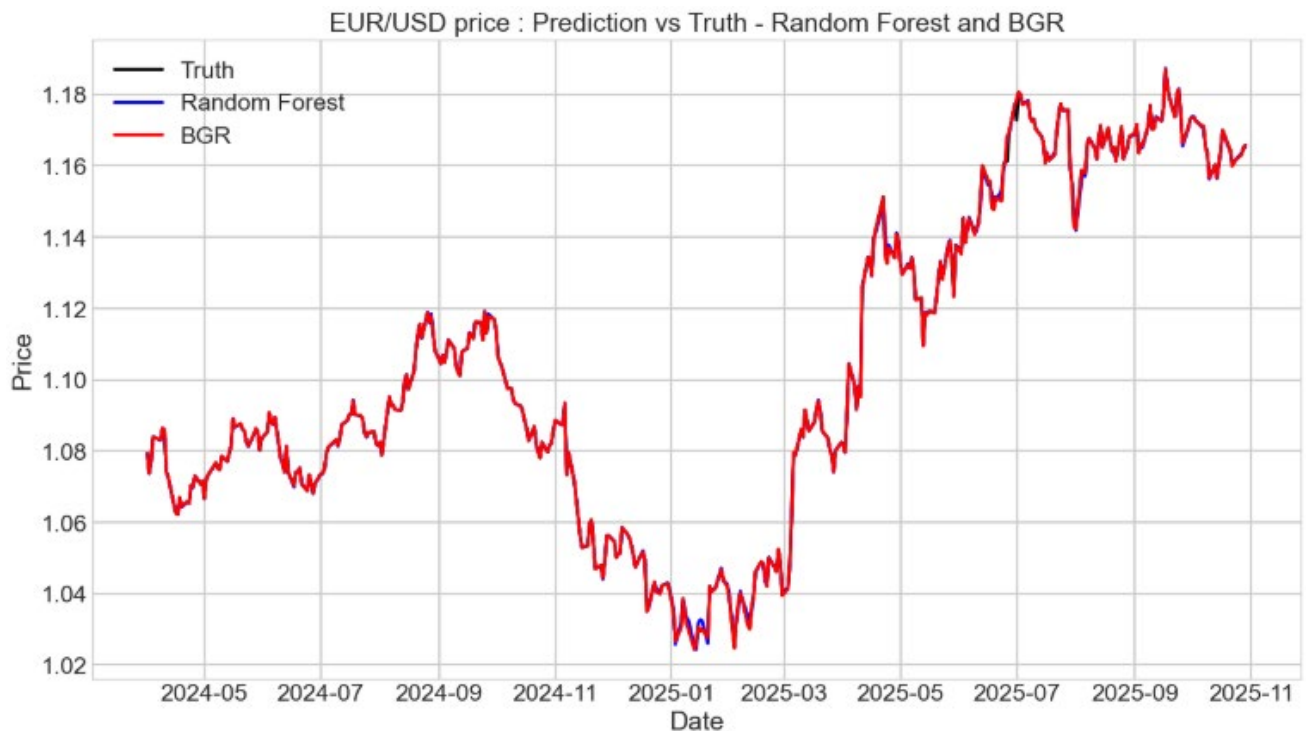


Рисунок 3.20 – Графік прогнозу та поточної ціни на основі Bagging Regressor та Random Forest

3.5. Порівняльний аналіз моделей машинного навчання

На завершальному етапі дослідження виконано порівняльний аналіз розроблених моделей машинного навчання з метою визначення найбільш ефективного підходу для прогнозування цінових рухів валютної пари EUR/USD. Дане порівняння здійснювалося на основі результатів тестової вибірки з

використанням як статичних, так і фінансових метрик.

До певного аналізу включено такі моделі: лінійна регресія, стохастичний градієнтний спуск (SGD Regerssor), Bagging Regressor, Random Forest Regressor та XGBoost Regressor.

Для об'єктивного порівняння ефективності побудованих моделей машинного навчання використано статичні метрики регресії, а саме середню абсолютну похибку (MAE), середньоквадратичну похибку (RMSE) та коефіцієнт детермінації R^2 .

У таблиці 3.1 наведено результати оцінювання моделей за основними регресійними метриками:

Таблиця 3.1 – Порівняльні результати моделей за статичними метриками

Модель	MAE	RMSE	R^2
Linear Regression	0,000096	0,00046	0,999894
SGD Regressor	0,005278	0,006779	0,977016
Bagging Regressor	0,00009	0,000459	0,999894
Random Forest Regressor	0,000372	0,000882	0,999611
XGBoost Regressor	0,000626	0,001067	0,999430

Аналіз результатів статичних метриків наведених у таблиці 3.1, свідчить про суттєву відмінність якості прогнозування між досліджуваними моделями. Найвищу точність відтворення фактичної динаміки курсу EUR/USD продемонстрували моделі Linear Regression та Bagging Regressor, яких зафіксовано мінімальні значення середньої абсолютної та середньоквадратичних похибок.

Моделі Random Forest Regressor та XGBoost Regressor, незважаючи на складну архітектуру та здатність виявляти нелінійні залежності, продемонстрували вищі значення похибок. Така поведінка пов'язана з тим, що для фінансових часових рядів із домінуванням трендових компонентів не часто призводить до підвищення точності прогнозу.

Найнижчу якість прогнозування серед усіх досліджуваних алгоритмів показав SGD Regressor. Отримані результати свідчать про його чутливість до шуму та особливостей масштабування ознак (Рис. 3.21).

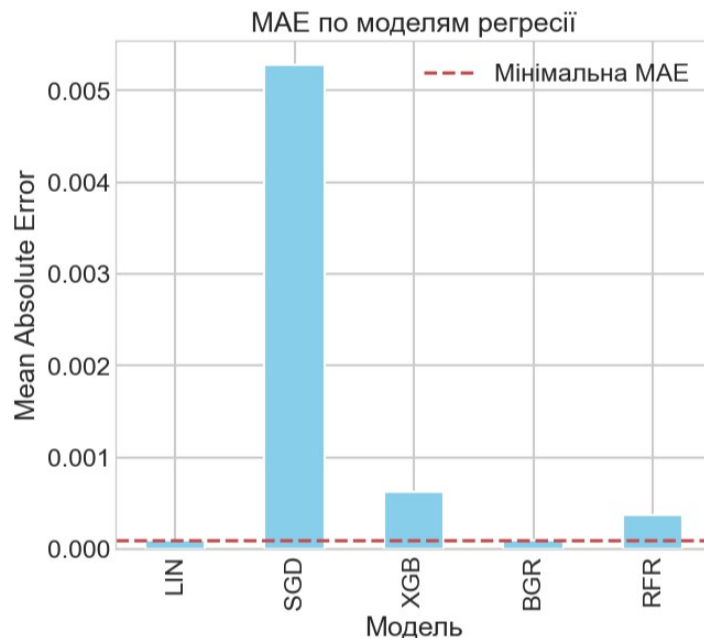


Рисунок 3.21 – Графік моделі за середньої абсолютної похибки

Водночас слід враховувати, що низькі значення похибок не завжди відображають ефективність моделі в умовах реальної торгівлі. Тому додатково, виконано оцінювання моделей за фінансовими метриками, що дозволяє визначити їхню практичну придатність для використання у торгових стратегіях.

У таблиці 3.2 наведено наступні результати оцінювання моделей за фінансовими метриками:

Таблиця 3.2 – Порівняльні результати моделей за фінансовими метриками

Модель	Sharpe Ratio	Max Drawdown	Win Rate	Direction Accuracy
Linear Regression	1,80165	-0.084602	0,987805	0,990220
SGD Regressor	1,80165	-0.084602	0,575610	0,990220

Bagging Regressor	1,80165	-0.084602	0,987805	0,990220
Random Forest Regressor	1,80165	-0.084602	0,970732	0,973105
XGBoost Regressor	1,80165	-0.084602	0,907317	0,909535

На основі отриманих значень коефіцієнта Sharpe Ratio є однаковими для всіх моделей, які становлять 1,80165. В межах реалізованої торгової логіки та однакових умов формування сигналів співвідношення дохідності до ризику залишалося стабільним. Аналогічна ситуація як у показника Max Drawdown, що вказує на однаковий рівень максимальних просадок капіталу для всіх моделей.

Водночас основні відмінності між моделями виявляться у показниках Win Rate і Direction Accuracy. Найвищі значення цих метрик продемонстрували Linear Regression та Bagging Regressor. Частка прибуткових прогнозів перевищує 98%, а точність визначення напрямку руху ціни становить понад 99%.

Модель Random Forest Regressor показала стабільні результати, що підтверджує її здатність формувати коректні торгові сигнали. У XGBoost Regressor характеризується помірним зниженням Win Rate та Direction Accuracy. Найменшу ефективність за показником Win Rate став SGD Regressor.

Таким чином, щодо вибору якнайкращих серед визначених моделей на основі результатів статичних та фінансових метриках, можна сказати, що модель Bagging Regressor стала ефективним для торгової стратегії. Вибраний модель показала найвищу точність прогнозування та забезпечила оптимальне співвідношення прибутку і ризику у тестовому періоді (Рис. 3.22).

```

Модель з найменшою середньою абсолютною похибкою:
BGR    0.00009
dtype: float64
Модель з найменшою середньоквадратичною похибкою:
BGR    0.000459
dtype: float64
Модель з найбільшою прохідністю:
LIN    0.995122
BGR    0.995122
dtype: float64

```

Рисунок 3.22 – Визначення вибору моделей за певними метриками

РОЗДІЛ 4.

ОХОРОНА ПРАЦІ ТА БЕЗПЕКА У НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1. Особливості безпеки під час роботи персональним комп'ютером

Сучасні комп'ютерні кабінети можуть бути обладнанні не лише персональними комп'ютерами і ноутбуками, але й оснащені інтерактивною дошкою і мультимедійним проектором. Для безпечної роботи з такою технікою кожна особа необхідно знати інструкції з охорони праці [22].

Під час роботи на комп'ютерах можуть діяти такі небезпечні та шкідливі фактори, як-от: фізичні, психологічні, біологічні.

Основним обладнанням робочого місця користувача комп'ютера є монітор, системний блок, клавіатура, принтери, БФП, комп'ютерна миша, і так далі.

Робочі місця мають бути розташовані на відстані не менше 1,5 м від стіни з вікнами, від інших стін на відстані 1м, між собою на відстані не менше 1,5 м. Відносно вікон робоче місце доцільно розташовувати таким чином, щоб природне світло падало на нього збоку, переважно зліва [22].

Робочі місця слід розташовувати так, щоб уникнути попадання в очі прямого світла. Джерела освітлення рекомендується розташовувати з обох боків екрану паралельно напрямку погляду. Для уникнення світлових відблисків екрану, клавіатури в напрямку очей користувача, від світильників загального освітлення або сонячних променів, необхідно використовувати антиполюскові сітки, спеціальні фільтри для екранів, захисні козирки, на вікнах - жалюзі. Фільтри з металевої або нейлонової сітки використовувати не рекомендується, тому що сітка спотворює зображення через інтерференцію світла. Найкращу якість зображення забезпечують скляні поляризаційні фільтри. Вони усувають практично всі відблиски, роблять зображення чітким і контрастним.

Монітор повинен бути розташований на робочому місці так, щоб поверхня екрана знаходилася в центрі поля зору на відстані 400-700 мм від очей користувача. Рекомендується розміщувати елементи робочого місця так, щоб витримувалася однакова відстань очей від екрана, клавіатури, тексту [22].

Зручна робоча поза при роботі з комп'ютером забезпечується регулюванням

висоти робочого столу, крісла та підставки для ніг. Раціональною робочою позою може вважатися таке положення, при якому ступні працівника розташовані горизонтально на підлозі або підставці для ніг, стегна зорієнтовані у горизонтальній площині, верхні частини рук - вертикальні. Кут ліктьового суглоба коливається в межах 70-90°, зап'ястя зігнуті під кутом не більше ніж 20°, нахил голови 15-20° [22].

Згідно зі статтею 18 Закону України «Про охорону праці» працівник зобов'язаний:

а) знати і виконувати вимоги нормативних актів про охорону праці, правила поведінки з устаткуванням та іншими засобами виробництва, користуватися засобами колективного та індивідуального захисту;

б) дотримуватись зобов'язань щодо охорони праці, передбачених колективним договором та правилами внутрішнього трудового розпорядку підприємства;

в) співробітничати з власником у справі організації безпечних і нешкідливих умов праці, особисто вживати посилюючих заходів щодо усунення будь-якої виробничої ситуації, яка створює загрозу його життю чи здоров'ю, або людей, які його оточують, повідомляти про небезпеку свого безпосереднього керівника або іншу посадову особу.

4.2. Вимоги безпеки під час роботи персональним комп'ютером

Безпечна робота за персональним комп'ютером передбачає дотримання певних правил як до початку роботи, так і після її завершення. Дана робота дозволяє запобігти технічним несправностям, мінімізувати ризики для здоров'я та забезпечити комфортні умови праці. Оскільки неправильна організація робочого місця або недбале користування обладнанням призведуть до перевтоми, проблем із зором, болем у спині, та іншим негативним наслідкам.

Для вимог безпеки перед початком роботи з ПК, розглядають:

- увімкнути систему кондиціонування в приміщенні;
- перевірити надійність встановлення апаратури на робочому столі:

- повернути монітор так, щоб було зручно дивитися на екран - під прямим кутом (а не збоку) і трохи зверху вниз, при цьому екран має бути трохи нахиленим, нижній його край ближче до оператора;
- перевірити загальний стан апаратури, перевірити справність електропроводки, з'єднувальних шнурів, штепсельних вилок, розеток, заземлення захисного екрана;
- відрегулювати освітленість робочого місця;
- відрегулювати та зафіксувати висоту крісла, зручний для користувача нахил його спинки;
- приєднати до системного блоку необхідну апаратуру. Усі кабелі, що з'єднують системний блок з іншими пристроями, слід вставляти та виймати при вимкненому комп'ютері;
- ввімкнути апаратуру комп'ютера вимикачами на корпусах в послідовності: монітор, системний блок, принтер (якщо передбачається друкування);
- відрегулювати яскравість свічення монітора, мінімальний розмір світної точки, фокусування, контрастність. Не слід робити зображення надто яскравим, щоб не втомлювати очей.

Рекомендовано:

- відношення яскравості монітора до яскравості оточуючих його поверхонь в робочій зоні - не більше 3:1;
- мінімальний розмір точки свічення не більше 0,4 мм для монохромного монітора і не менше 0,6 мм для кольорового, контрастність зображення знаку – не менше 0,8;
- у разі виявлення будь-яких несправностей роботу не розпочинати, повідомити про це керівника [22].

Вимоги безпеки після закінчення роботи за ПК:

- закінчити та записати у пам'ять комп'ютера файл, що знаходиться в роботі;
вимкнути принтер та інші периферійні пристрої. Штепсельні вилки витягнути з розеток. Накрити клавіатуру кришкою запобігання попаданню в неї пилу;

- прибрати робоче місце;
- ретельно вимити руки теплою водою з милом;
- вимкнути кондиціонер, освітлення і загальне електроживлення
- пройти в спеціально обладнаному приміщенні сеанс психофізіологічного розвантаження і зняття втоми з використанням спеціальних вправ аутогенного тренування [22].

4.3. Безпека у надзвичайній ситуації

Пожежна безпека – це стан об'єкта, при якому з регламентованою ймовірністю відкидається можливість виникнення та розвиток пожежі, і впливу на людей її небезпечних факторів, а також забезпечується захист матеріальних цінностей [23].

Пожежа, також застосовують пожежа – неконтрольоване горіння поза спеціальним вогнищем, що розповсюджується в часі і просторі, знищує матеріальні цінності, створює загрозу для життя людей, тварин, негативно впливає на навколишнє природне середовище [23].

Для успішного проведення дієвих упереджувальних заходів важливо знати основні причини виникнення пожеж.

До чинників, що можуть викликати пожежу саме з цієї причини, належать: короткі замикання, несправності електроустаткування та приладів, струмові перевантаження, що виникають у силових та освітлювальних електромережах, великі значення перехідних опорів.

Короткі замикання виникають внаслідок неправильного монтажу або експлуатації електроустановок, старіння або пошкодження ізоляції.

Струм короткого замикання залежить від потужності джерела струму, відстані від джерела струму до місця замикання та виду замикання [23].

Великі струми замикання зумовлюють іскріння та нагрівання струмопровідних частин до високої температури, що може бути причиною займання ізоляції провідників та горючих будівельних конструкцій, які знаходяться поряд.

Струмові перевантаження виникають при ввімкненні до мережі додаткових споживачів струму або при зниженні напруги в мережі. Тривале 49 перевантаження призводить до нагрівання провідників, що може викликати займання ізоляції [23].

Потенційну небезпеку виникнення пожежі або вибуху в умовах виробництва викликають і хімічні речовини, які в контакті з повітрям чи водою, а також при взаємодії між собою виділяють велику кількість теплової енергії [23].

4.4. Напрями інформаційного забезпечення охорони праці

Пропаганда безпечних та нешкідливих умов праці займає одне із важливих місць в практиці профілактичної роботи збереження життя. Встановлено, що необхідно психологічним чином заохочувати робітників до дотримання вимог безпеки праці, підвищувати їх мотивацію нагадувати про важливість власного життя [24].

Дослідження показують, що найчастіше на підприємствах є куточки охорони праці та пожежної безпеки у вигляді стенду, де розміщені в більшості:

- інформація, що стосується охорони праці (термін охорони праці, Закон України «Про охорону праці», інструктажі з питань охорони праці та графік їх проведення, обов'язки працівників з охорони праці, інформація про СУОП);
- інформація про порядок надання першої медичної допомоги;
- інформація з інструкції;

Досліджено, що найчастіше відсутні плакати, що стосуються: правил користування ЗІЗ та апаратами захисту органів дихання, правил поведінки у різних ситуаціях, що можуть виникнути.

Існують наступні різновиди плакатів: позитивний, підкреслює переваги безпечної праці; страхітливий, що показує шкода від порушення правил безпеки; нейтральний, що містить емоційно не пофарбовані рекомендації без показу і оцінок наслідків їх невиконання; комічний, гумористична різновид позитивного плаката; комбінований, який поєднує в собі перелічені різновиди.

Інформаційні плакати безпеки повинні виконувати три головні мети: привернути до себе увагу, бути зрозумілим правильно і запам'ятатися; викликати бажану емоційну реакцію.

Дослідження плакатів, які використовують на підприємствах так і в іншому конторах, показують використання на них жовто-, чорно-, червоної гама кольорів, що наголошує на попередження різного роду небезпек. Практичний досвід показує, що більшість плакатів повинні містити психологічні тези такі як наприклад: «Сьогодні повезло, а завтра може і ні!», «Працівники, пам'ятайте! Дотримання вимог промислової безпеки – запорука вашого життя та здоров'я», «Тебе чекають вдома – працюй безпечно», «Щоб не трапилось біди – про безпеку думай ти», «Використовуйте засоби захисту та справні механізми про виконанні кожної роботи», «Безпеці праці – належну увагу. Щоб недбалість не обернулось бідою».

Встановлено, що оформлення стендів та інформаційних плакатів, які мають психологічний вплив на дотриманні безпеки праці чи наголошували на ролі власного життя в більшому випадку відсутні.

Можна погодитися з думкою ряду авторів, що, як правило, той, хто вивчає плакат, робить висновок про те, як йому необхідно діяти, не з напису на плакаті, а з рисунка. Напис має бути коротким і тільки доповнювати те, що не вдавалось передати рисунком. При цьому сприйняття напису слід полегшувати за допомогою кольору і шрифту. Плакат буде «працювати» в тому випадку, якщо він відповідає таким вимогам: привертає до себе увагу; правильно розуміється і запам'ятовується; викликає бажану емоційну реакцію, бажане відношення. Як відомо, при проведенні пропаганди безпеки праці необхідно наголошувати та пріоритетності життя та здоров'я працюючих.

Таких плакатів, як мають психологічний вплив на дотриманні безпеки праці чи наголошували на ролі власного життя мало. Є різні варіанти соціальних плакатів, серед яких мають місце ті, що несуть в собі зміст який може допомогти усвідомити цінну власного життя людині будь-якої професії чи сфері.

РОЗДІЛ 5.

ВИЗНАЧЕННЯ ЕФЕКТИВНОСТІ

Визначення ефективності розробленої моделі машинного навчання є завершальним етапом магістерського дослідження і має на меті комплексну оцінку отриманих результатів. На цьому етапі узагальнюються результати експериментальних досліджень, перевіряється, наскільки обрані методи відповідають поставленим завданням, а також оцінюється практична доцільність використання моделі для прогнозування фінансових показників у реальних умовах ринку.

Особливістю фінансових часових рядів є їх висока волатильність, наявність шуму, нелінійна динаміка та нестабільність статичних властивостей у часі. Тому для оцінки ефективності моделі було застосовано комплексний підхід, який поєднує використання стандартних регресійних метрик для оцінки точності прогнозування та фінансових показників для оцінки практичної користі прогнозів.

На першому етапі здійснювалась перевірка точності прогнозування. Для цього використовувалися показники середньої абсолютної похибки (MAE), кореня середньоквадратичної похибки (RMSE) та коефіцієнт детермінації (R^2). Середня абсолютна похибка визначається формулою:

$$MAE = \frac{1}{n} \sum_{i=1}^N |y_i - \hat{y}_i|. \quad (5.1)$$

де y_i – фактичне значення;

$\hat{y}_i$ – прогнозне;

n – кількість спостережень.

Припустимо, що модель прогнозує ціну активу на п'ять наступних днів, а середня абсолютна різниця між прогнозом та фактом становить 0,8 одиниці. Це означає, що у середньому прогноз відстає від реальної ціни 0,8 одиниці, що дає відчуття масштабів похибки в грошовому виразі.

Середньоквадратична похибка (RMSE) дозволяє оцінити, наскільки модель реагує на різкі коливання ринку, тому визначається за формулою:

$$RMSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2. \quad (5.2)$$

Наприклад, якщо у один день ринок різко падає на 6 одиниць, RMSE збільшується більше, ніж MAE, підкреслюючи, що модель менш точно відображає екстремальні рухи.

Коефіцієнт детермінації R^2 показує, яка частка дисперсії фактичних даних пояснюється моделлю. Даний показник визначається за такою формулою:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (5.3)$$

Якщо $R^2 = 0,85$, це означає 85% коливань ринку модель змогла відтворити, що є високим показником для фінансових часових рядів.

Однак точні прогнози не гарантують прибутковості у реальній торгівлі, тому що є ризик на певну втрату капіталу. Тому для цього прогнозні сигнали застосовуються у торговій логіці, і на основі капіталу визначаються сумарна прибутковість, максимальне просідання капіталу та коефіцієнт Шарпа.

Сумарна прибутковість обчислюється за формулою:

$$TR = \frac{C_f - C_0}{C_0} * 100\%, \quad (5.4)$$

де C_0 – початковий капітал;

C_f – кінцевий капітал.

Наприклад, якщо початковий капітал 1000 грн, в результаті після серії прогнозів і торгів становить 1150 грн, сумарна прибутковість складе 15%. Тому такі прогнози стають корисними на практиці.

Максимальне просідання капіталу визначається як:

$$MDD = \max_{t \in [1, T]} \frac{C_{p,t} - \max(V_t)}{\max(V_{p,t})} * 100\%. \quad (5.5)$$

де $C_{p,t}$ – локальний максимум капіталу на момент t ;

C_t – капітал у будь-який момент.

Наприклад, якщо на піку капіталу було 1200 грн, а найнижче значення опустилося до 1020 грн, то MDD складе 15%. Цей показник допомагає оцінити ризики та підготуватися до можливих втрат.

Коефіцієнт Шарпа розраховується за формулою:

$$SR = \frac{\bar{R}_p - R_f}{\sigma_p}, \quad (5.6)$$

де $\bar{R}_p$ – середня дохідність портфеля,

R_f – безризикова ставка;

σ_p – стандартне відхилення доходності.

Значення коефіцієнта Шарпа дозволяє оцінити, як ефективно стратегія компенсує прийнятий ризик. Значення $SR > 1$, свідчить про прийнятий баланс між доходністю та ризиком.

Для узагальненої оцінки ефективності моделей було використано інтегральний показник, який поєднує точність прогнозування та фінансову результативність

$$E = \alpha \times (1 - RMSE_{norm}) + \beta \times SR - \gamma \times MDD_{norm}, \quad (5.7)$$

де $RMSE_{norm}$, MDD_{norm} – нормовані значення відносних показників;

α, β, γ – вагові коефіцієнти, що відображають пріоритетність точності, доходності та ризику.

Такий підхід дозволяє порівнювати різні моделі за єдиним критерієм та обґрунтовано обирати оптимальну модель для практичного використання.

Результати експериментів досліджень показали, що модель Bagging Regression забезпечує найкраще поєднання точності прогнозування та фінансової стабільності. Вона характеризується зниженими MAE та RMSE, меншою чутливістю до шуму у вхідних даних та більш рівномірною динамікою прогнозів і підвищити надійність результатів, що є критично важливим для фінансових часових рядів. Для підтвердження практичної ефективності моделі результати прогнозування було інтегровано у торгову логіку, на основі побудовано криву капіталу та оцінено максимальну просадку.

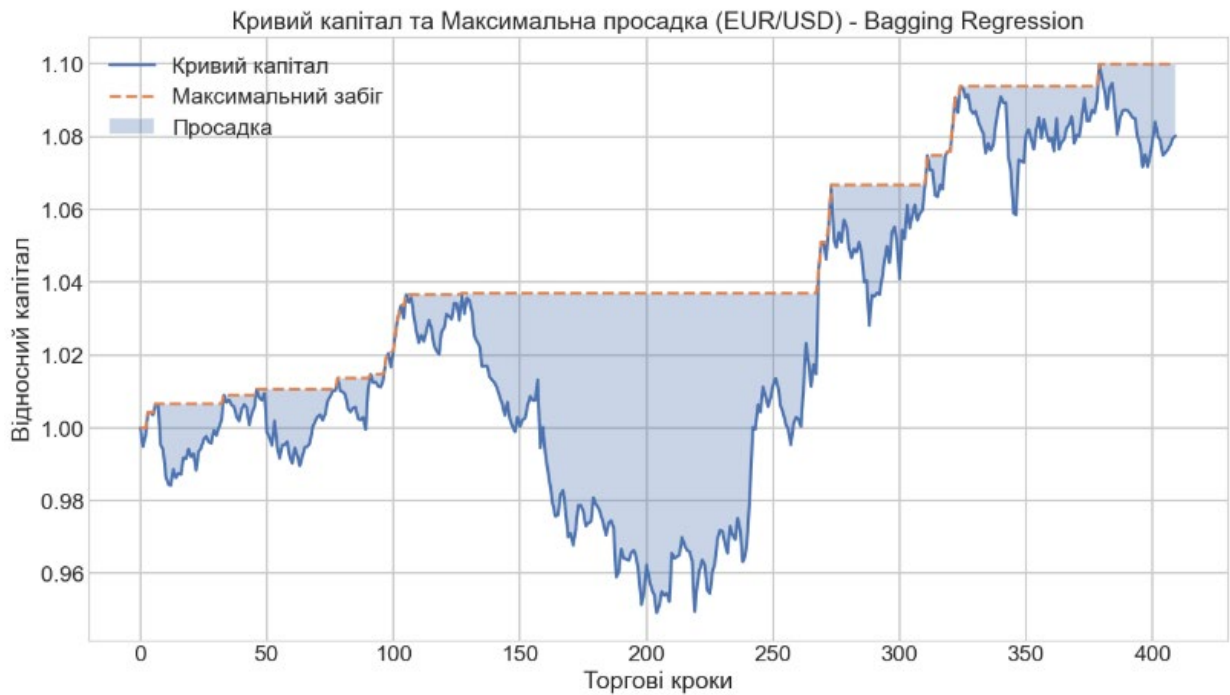


Рисунок 5.1 – Крива капіталу та максимальна просадка для моделі Bagging Regression

Як видно з рисунка 5.1, крива капіталу має загальну висхідну динаміку з наявністю тимчасових просадок, що є характерним для реальних фінансових ринків. Водночас модель демонструє здатність до відновлення після періодів зниження та оновлення локальних максимумів, що свідчать про її фінансову стійкість. Обмежений рівень максимальної просадки підтверджує збалансоване співвідношення між прибутковістю та ризиком при використанні прогнозів моделі Bagging Regression.

У підсумку, проведене визначення ефективності підтвердило, що розроблена модель машинного навчання є як статично обґрунтованою, так і практично придатною для використання. Вона забезпечує прийнятний рівень точності прогнозування, стабільні фінансові результати та збалансоване співвідношення між прибутковістю і ризиком. Отримані результати мають здатність використати при розробці аналітичних досліджень у сфері застосування методів машинного навчання у фінансовому аналізі.

ВИСНОВКИ

У магістерській роботі було проведено комплексне дослідження ефективності стратегій трейдингу на бінарних опціонах із застосуванням сучасних методів машинного навчання. Бінарні опціони, попри значну популярність, характеризує високою волатильністю, мінімальною прозорістю та коротким інвестиційним горизонтом, що ускладнює прогнозування ринкових рухів.

У ході дослідження виконано аналіз сучасних наукових та практичних розробок, пов'язаних із застосуванням ML-моделей у фінансовому прогнозуванні. Розглянуто особливості бінарних опціонів та фактори, що впливають на якість прогнозів такі як, часовий горизонт, рівень волатильності, вплив новинного фону, специфіку торгової платформи та обмеженість доступних даних.

На основі зібраного історичного набору даних було здійснено попередню обробку та розширення ознакового простору шляхом створення додаткових технічних індикаторів: ковзних середніх різної довжини, показників волатильності, зміщених цін, дохідностей, індикаторів імпульсу та статистичних характеристик.

Було реалізовано та порівняно декілька алгоритмів такі як лінійна регресія, градієнтний бустинг, дерево рішень Random Forest, а також проведено оптимізацію гіперпараметрів і такі інші. Оцінювання виконувалося за ключовими метриками, а саме: RMSE, MAE, R^2 , точність напрямку руху та аналізом стратегічної ефективності – потенційного фінансового результату застосування моделі у вигляді торгових угод.

За результатами порівняння встановлено, що моделі ансамблевого типу, демонструють вищу точність прогнозування короткострокових змін ціни порівняно з класичними алгоритмами. Проте у найточніших моделей зберігається обмеження у вигляді високого рівня шуму на ринку бінарних опціонів, який не дозволяє створити повністю стабільно прибуткову автоматизовану стратегію без урахування ризик-менеджменту.

Проведено визначення ефективності показав, що впровадження системи

машинного аналізу є економічно доцільним за наявності достатньої кількості історичних даних, правильно налаштованої стратегії ризику та регламенту торгівлі.

У подальших дослідженнях є зосередження уваги на вдосконаленні алгоритмів прогнозування, підвищенні стійкості моделей до різких ринкових змін та дослідженні можливостей інтеграції методів глибокого навчання для аналізу фінансових часових рядів.

Щодо рекомендацій, можна запропонувати:

- 1) Використання глибинних нейронних мереж (LSTM, GRU) для кращого прогнозування часових рядів;
- 2) Використання текстового аналізу новин для врахування впливу новинного фону;
- 3) Введення показників настроїв ринку;
- 4) Розробка автономного робота, який використовує ML-моделі для прийняття рішень;
- 5) Впровадження моніторингу в реальному часі для оцінки точності прогнозів;
- 6) Порівняння результатів ML-моделей з класичними економетричними методами.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Бінарні опціони в Україні: Спосіб заробітку чи шахрайство? URL: <http://securities.usmdi.org/?p=22&n=97&s=983> (дата звернення: 12.10.2025).
2. Géron A. *Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow*. - 3rd Edition. - O'Reilly Media, 2023. pp. 45-50.
3. López de Prado M. *Advances in Financial Machine Learning*. - Wiley, 2018. pp. 25-26.
4. Hastie T., Tibshirani R., Friedman J. *The Elements of Statistical Learning*. - Springer, 2017. pp. 112-114.
5. Cortes C., Vapnik V. *Support-Vector Networks*. - *Machine Learning*, 1995. pp. 273-279.
6. Breiman L. *Random Forests*. - *Machine Learning Journal*, 2001. pp 10-15.
7. Murphy K. P. *Machine Learning: A Probabilistic Perspective*. - MIT Press, 2012. pp. 203-210.
8. MacQueen J. *Some Methods for Classification and Analysis of Multivariate Observations*. - *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 1967. pp. 285-290.
9. Sutton R., Barto A. *Reinforcement Learning: An Introduction*. - MIT Press, 2018. pp. 45-50.
10. Bishop C. M. *Pattern Recognition and Machine Learning*. - Springer, 2006. pp. 320-225.
11. Індикатори технічного аналізу в торгівлі та для чого вони потрібні. URL: <https://rates.fm/ua-uk/cryptocurrency/indikatori-tehnichnogo-analizu-v-torgivli/> (дата звернення: 15.11.2025).
12. Tsay R. S. *Analysis of Financial Time Series*. Wiley, 2010.
13. Hastie T., Tibshirani R., Friedman J. *The Elements of Statistical Learning*. Springer, 2009.
14. de Prado M. *Advances in Financial Machine Learning*. Wiley, 2018 - 432 p.
15. Fischer T., Krauss C. Deep learning for financial market prediction. *European Journal of Operational Research*, 2018. pp. 654-669.

16. Geron, A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. - 2nd Edition. - O'Reilly Media, 2019. 232 p.
17. Hull, J. C. *Options, Futures, and Other Derivatives*. 10th ed. Pearson, 2017. pp. 200-205.
18. Fawcett, T. "An Introduction to ROC Analysis." *Pattern Recognition Letters*, 27(8), 2006. pp. 861-884.
19. Pedregosa F. et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2011. 285 p.
20. VanderPlas J. *Python Data Science Handbook*. - O'Reilly, 2016. 548 p.
21. Kluyver T. et al. Jupyter Notebooks - a publishing format for reproducible computational workflows, 2016. pp. 87-90.
22. Розвиток системи управління закладом освіти: матеріали Всеукраїнської науково-практичної онлайн-конференції; [за заг. ред. проф. Р. Черновол-Ткаченко, проф. О. Мармази, доц. О. Гречаник]. Харків : ХНПУ імені Г. С. Сковороди, 2021. 49 с.
23. Пожарова О.В. Охорона праці. Навчальний посібник. Одеса: НУ «Одеська юридична академія», 2022. 48 с.
24. Охорона праці (гігієна праці та виробнича санітарія): навчальний посібник / Пістун І.П., Березовецький А. П., Тимочко В.О., Городецький І. М.; За ред. І.П. Пістуна. - Львів: "Тріада плюс", 2017. 620 с.

ДОДАТКИ

ДОДАТОК А.

Функція `generate_features` для головних ознак середні ковзні

```
def generate_features(df):
    """ Генерація ознак для акції/індексу/валюти/товару на основі історичної ціни
    та ефективності

    Аргументи:
        df (dataframe with columns "open", "close", "high", "low", "volume")

    Повертає:
        фрейм даних, набір даних з новими ознаками
    """
    df_new = pd.DataFrame()

    # 6 оригінальні ознаки
    df_new['open'] = df['open']
    df_new['open_1'] = df['open'].shift(1)
    df_new['close_1'] = df['close'].shift(1)
    df_new['high_1'] = df['high'].shift(1)
    df_new['low_1'] = df['low'].shift(1)
    df_new['volume_1'] = df['volume'].shift(1)

    # 50 оригінальні ознаки
    # середня ціна
    df_new['avg_price_5'] = df['close'].rolling(window=5).mean().shift(1)
    df_new['avg_price_30'] = df['close'].rolling(window=21).mean().shift(1)
    df_new['avg_price_90'] = df['close'].rolling(window=63).mean().shift(1)
    df_new['avg_price_365'] = df['close'].rolling(window=252).mean().shift(1)

    # співвідношення середніх цін
    df_new['ratio_avg_price_5_30'] = df_new['avg_price_5'] / df_new['avg_price_30']
    df_new['ratio_avg_price_905_'] = df_new['avg_price_5'] / df_new['avg_price_90']
    df_new['ratio_avg_price_5_365'] = df_new['avg_price_5'] / df_new['avg_price_365']
    df_new['ratio_avg_price_30_90'] = df_new['avg_price_30'] / df_new['avg_price_90']
    df_new['ratio_avg_price_30_365'] = df_new['avg_price_30'] / df_new['avg_price_365']
    df_new['ratio_avg_price_90_365'] = df_new['avg_price_90'] / df_new['avg_price_365']

    # середній обсяг
    df_new['avg_volume_5'] = df['volume'].rolling(window=5).mean().shift(1)
    df_new['avg_volume_30'] = df['volume'].rolling(window=21).mean().shift(1)
    df_new['avg_volume_90'] = df['volume'].rolling(window=63).mean().shift(1)
```

```

df_new['avg_volume_365'] = df['volume'].rolling(window=252).mean().shift(1)

# середнє відхилення обсягу

df_new['ratio_avg_volume_5_30']=df_new['avg_volume_5']/df_new['avg_volume_30']
df_new['ratio_avg_volumee_5_90']=df_new['avg_volume_5']/df_new['avg_volume_90']
df_new['ratio_avg_volume_5_365']=df_new['avg_volume_5']/df_new['avg_volume_365']
df_new['ratio_avg_volume_30_90']=df_new['avg_volume_30']/df_new['avg_volume_90']
df_new['ratio_avg_volume_30_365']=df_new['avg_volume_30']/df_new['avg_volume_365']
df_new['ratio_avg_volume_90_365']=df_new['avg_volume_90']/df_new['avg_volume_365']

# стандартне відхилення цін

df_new['std_price_5'] = df['close'].rolling(window=5).std().shift(1)
df_new['std_price_30'] = df['close'].rolling(window=21).std().shift(1)
df_new['std_price_90'] = df['close'].rolling(window=63).std().shift(1)
df_new['std_price_365'] = df['close'].rolling(window=252).std().shift(1)

# стандартне відхилення відхилення цін

df_new['ratio_std_price_5_30']=df_new['std_price_5']/df_new['std_price_30']
df_new['ratio_std_price_5_90']= df_new['std_price_5']/df_new['std_price_90']
df_new['ratio_std_price_5_365']=df_new['std_price_5']/df_new['std_price_365']
df_new['ratio_std_price_30_90']=df_new['std_price_30']/df_new['std_price_90']
df_new['ratio_std_price_30_365']=df_new['std_price_30']/df_new['std_price_365']
df_new['ratio_std_price_90_365']=df_new['std_price_90']/df_new['std_price_365']

# стандартне відхилення обсягів

df_new['std_volume_5'] = df['volume'].rolling(window=5).std().shift(1)
df_new['std_volume_30'] = df['volume'].rolling(window=21).std().shift(1)
df_new['std_volume_90'] = df['volume'].rolling(window=63).std().shift(1)
df_new['std_volume_365'] = df['volume'].rolling(window=252).std().shift(1)

# коефіцієнт стандартного відхилення об'ємів

df_new['ratio_std_volume_5_30']=df_new['std_volume_5']/df_new['std_volume_30']
df_new['ratio_std_volume_5_90']=df_new['std_volume_5']/df_new['std_volume_90']
df_new['ratio_std_volume_5_365']=df_new['std_volume_5']/df_new['std_volume_365']
df_new['ratio_std_volume_30_90']=df_new['std_volume_30']/df_new['std_volume_90']
df_new['ratio_std_volume_30_365']=df_new['std_volume_30']/df_new['std_volume_365']
df_new['ratio_std_volume_90_365']=df_new['std_volume_90']/df_new['std_volume_365']

```

```

# дохідність
df_new['return_1']=(df['close']-
df['close'].shift(1))/df['close'].shift(1)).shift(1)

df_new['return_5']=(df['close']-
df['close'].shift(5))/df['close'].shift(5)).shift(1)

df_new['return_30']=(df['close']-
df['close'].shift(21))/df['close'].shift(21)).shift(1)

df_new['return_90']=(df['close']-
df['close'].shift(63))/df['close'].shift(63)).shift(1)

df_new['return_365']=(df['close']-
df['close'].shift(252))/df['close'].shift(252)).shift(1)


# середнє значення дохідності
df_new['moving_avg_5'] = df_new['return_1'].rolling(window=5).mean()
df_new['moving_avg_30'] = df_new['return_1'].rolling(window=21).mean()
df_new['moving_avg_30'] = df_new['return_1'].rolling(window=63).mean()
df_new['moving_avg_365'] = df_new['return_1'].rolling(window=252).mean()


# ціль
df_new['close'] = df['close']
df_new = df_new.dropna(axis=0)
return df_new


data = generate_features(eu)

```

ДОДАТОК Б.

Функція `evaluate_financial` для обчислення фінансових метрик для прогнозів

```
# Фінансові метрики
def evaluate_financial(y_pred, y_true, close_prices, risk_free_rate=0.0):
    """
    Обчислення фінансових метрик для прогнозів.

    Аргументи:
    y_true      - справжні ціни
    y_pred      - прогнозовані ціни
    close_prices - ціни закриття для розрахунку equity curve
    risk_free_rate - безризикова ставка для Sharpe ratio

    Повертає:
    metrics_dict - словник з метриками
    equity_curve - кумулятивна дохідність у вигляді pd.Series
    """
    # Дохідність від прогнозів
    returns = (y_pred[1:] - y_pred[:-1]) / y_pred[:-1]
    returns = np.insert(returns, 0, 0)

    equity_curve = np.cumprod(1 + returns)

    # Точність напрямку
    dir_true = np.sign(y_true[1:] - y_true[:-1])
    dir_pred = np.sign(y_pred[1:] - y_pred[:-1])
    direction_accuracy = np.mean(dir_true == dir_pred)

    # Win rate (частка виграшних прогнозів)
    profit_per_trade = y_true - np.roll(y_true, 1)
    profit_per_trade[0] = 0
    win_rate = np.mean((y_pred - np.roll(y_pred, 1)) * profit_per_trade > 0)

    # Середній прибуток на угоду
    avg_profit = np.mean(np.abs((y_pred - np.roll(y_pred, 1))[1:]))

    # Коефіцієнт Шарпа
    excess_returns = returns - risk_free_rate / 252
    sharpe_ratio = np.mean(excess_returns) / (np.std(excess_returns) + 1e-9) *
```

```
np.sqrt(252)
```

```
# Максимальна просадка
cum_returns = equity_curve
running_max = np.maximum.accumulate(cum_returns)
drawdown = (cum_returns - running_max) / running_max
max_drawdown = drawdown.min()

metrics_dict = {
    'Cumulative Return (last)': equity_curve[-1],
    'Sharpe Ratio': sharpe_ratio,
    'Max Drawdown': max_drawdown,
    'Win Rate': win_rate,
    'Average Profit per Trade': avg_profit,
    'Direction Accuracy': direction_accuracy
}

return metrics_dict, equity_curve
```

ДОДАТОК В.

Код реалізації всіх моделей машинного навчання

```

# Linear Regression
from sklearn.linear_model import LinearRegression
lin = LinearRegression()

lin.fit(X_scaled_train, y_train)
predictions_lin = lin.predict(X_scaled_test)
metrics_lin, eq_curve_lin = evaluate_financial(y_test.values, predictions_lin,
data_test['close'].values)

print('RMSE: {0:.3f}'.format(mean_squared_error(y_test, predictions_lin)**0.5))
print('MAE: {0:.3f}'.format(mean_absolute_error(y_test, predictions_lin)))
print('R^2: {0:.3f}'.format(r2_score(y_test, predictions_lin)))
print("LR Фінансові метрики:")
for k, v in metrics_lin.items():
    print(f"{k}: {v:.4f}")

# SGD Regressor
from sklearn.linear_model import SGDRegressor

param_grid = {
    'penalty': ['l1', 'l2', 'elasticnet'],
    "alpha": [1e-5, 3e-5, 1e-4],
    "eta0": [0.01, 0.03, 0.1],
}

sgd = SGDRegressor()
grid_search = GridSearchCV(sgd, param_grid, cv=5,
scoring='neg_mean_absolute_error', n_jobs=-1)
grid_search.fit(X_scaled_train, y_train)

print(grid_search.best_params_)

sgd_best = grid_search.best_estimator_
#print(grid_search.best_score_)

predictions_sgd = sgd_best.predict(X_scaled_test)
metrics_sgd, eq_curve_sgd = evaluate_financial(y_test.values, predictions_sgd,
data_test['close'].values)

```

```

#evaluating the predictions
print('RMSE: {0:.3f}'.format(mean_squared_error(y_test, predictions_sgd)**0.5))
print('MAE: {0:.3f}'.format(mean_absolute_error(y_test, predictions_sgd)))
print('R^2: {0:.3f}'.format(r2_score(y_test, predictions_sgd)))
print("XGB Фінансові метрики:")
for k, v in metrics_sgd.items():
    print(f"{k}: {v:.4f}")

# Extreme Gradient Boosting Regressor
from xgboost import XGBRegressor, DMatrix
xgb = XGBRegressor()

data_dmatrix = DMatrix(data=X_train,
                        label=y_train)

xgb_param_grid = {'learning_rate': [0.001, 0.01, 0.1, 1],
                  'n_estimators': [50, 100, 200, 300],
                  'subsample': [0.3, 0.5, 0.7, 1]}

grid_search = GridSearchCV(estimator=xgb,
                           param_grid=xgb_param_grid,
                           scoring='neg_mean_squared_error',
                           cv=4,
                           verbose=1,
                           n_jobs=-1)

grid_search.fit(X_train, y_train)

print("Best parameters found: ",grid_search.best_params_)

xgb_best = grid_search.best_estimator_

xgb_best.fit(X_train,y_train)
predictions_xgb = xgb_best.predict(X_test)
metrics_xgb, eq_curve_xgb = evaluate_financial(y_test.values, predictions_xgb,
data_test['close'].values)

print('RMSE: {0:.3f}'.format(mean_squared_error(y_test, predictions_xgb)**0.5))
print('MAE: {0:.3f}'.format(mean_absolute_error(y_test, predictions_xgb)))

```

```

print('R^2: {0:.3f}'.format(r2_score(y_test, predictions_xgb)))
print("XGB Фінансові метрики:")
for k, v in metrics_xgb.items():
    print(f"{k}: {v:.4f}")

# Bagging Regressor
bgr = BaggingRegressor(estimator=lin, n_estimators=100, oob_score=True, n_jobs=-1)

bgr.fit(X_scaled_train, y_train)
predictions_bgr = bgr.predict(X_scaled_test)
metrics_bgr, eq_curve_bgr = evaluate_financial(y_test.values, predictions_bgr,
data_test['close'].values)

print('OOB: {0:.3f}'.format(bgr.oob_score))
print('RMSE: {0:.3f}'.format(mean_squared_error(y_test, predictions_bgr)**0.5))
print('MAE: {0:.3f}'.format(mean_absolute_error(y_test, predictions_bgr)))
print('R^2: {0:.3f}'.format(r2_score(y_test, predictions_bgr)))
print("BGR Фінансові метрики:")
for k, v in metrics_bgr.items():
    print(f"{k}: {v:.4f}")

# Random Forest
param_grid = {
    "max_depth": [30, 50],
    "min_samples_split": [5, 10, 20],

}

rf = RandomForestRegressor(n_estimators=100, random_state=42)
grid_search = GridSearchCV(rf, param_grid, cv=5,
scoring='neg_mean_absolute_error', n_jobs=-1)
grid_search.fit(X_train, y_train)

print(grid_search.best_params_)
# print(grid_search.best_score_)

rf_best = grid_search.best_estimator_
predictions_rf = rf_best.predict(X_test)
pred_rf = pd.Series(predictions_rf, index=y_test.index)
truth = y_test

```

```

metrics_rf, eq_curve_rf = evaluate_financial(y_test.values, predictions_rf,
data_test['close'].values)

print('RMSE: {0:.3f}'.format(mean_squared_error(y_test, predictions_rf)**0.5))
print('MAE: {0:.3f}'.format(mean_absolute_error(y_test, predictions_rf)))
print('R^2: {0:.3f}'.format(r2_score(y_test, predictions_rf)))

print("Random Forest Фінансові метрики:")
for k, v in metrics_rf.items():
    print(f"{k}: {v:.4f}")

```