

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені Ігоря СІКОРСЬКОГО»
НАВЧАЛЬНО-НАУКОВИЙ ФІЗИКО-ТЕХНІЧНИЙ
ІНСТИТУТ

Кафедра математичного моделювання та аналізу даних

«До захисту допущено»

В.о. завідувача кафедри

_____ Іван ТЕРЕЩЕНКО

«__» _____ 2023 р.

Дипломна робота
на здобуття ступеня бакалавра

зі спеціальності: 113 Прикладна математика
на тему: «Створення торговельних стратегій з урахуванням
інформації про аномалії у просторі DeFi»

Виконала: студентка 4 курсу, групи ФІ-92
Шевцова Марія Сергіївна

Керівник: асистент кафедри ММАД Яворський О. А. _____

Консультант: _____

Рецензент: асистент кафедри ІБ Наконечна Ю. В. _____

Засвідчую, що у цій дипломній
роботі немає запозичень з праць
інших авторів без відповідних
посилань.

Студентка _____

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені Ігоря СІКОРСЬКОГО»
НАВЧАЛЬНО-НАУКОВИЙ ФІЗИКО-ТЕХНІЧНИЙ
ІНСТИТУТ

Кафедра математичного моделювання та аналізу даних

Рівень вищої освіти — перший (бакалаврський)
Спеціальність (освітня програма) — 113 Прикладна математика,
ОПП «Математичні методи моделювання, розпізнавання образів та
комп'ютерного зору»

ЗАТВЕРДЖУЮ

В.о. завідувача кафедри

_____ Іван ТЕРЕЩЕНКО

«__» _____ 2023 р.

ЗАВДАННЯ
на дипломну роботу

Студент: Шевцова Марія Сергіївна

1. Тема роботи: *«Створення торговельних стратегій з урахуванням інформації про аномалії у просторі DeFi»*,

керівник: асистент кафедри ММАД Яворський О. А.,

затверджені наказом по університету №__ від «__» _____ 2023 р.

2. Термін подання студентом роботи: «__» _____ 2023 р.

3. Вихідні дані до роботи: опубліковані джерела за тематикою дослідження.

4. Зміст роботи: Темою дослідження є "Створення торговельних стратегій з урахуванням інформації про аномалії у просторі DeFi" описує процес створення торговельних стратегій, які враховують наявність аномалій у просторі децентралізованих фінансів (DeFi). В роботі вирішуються наступні задачі:

1) дослідження актуальних теоретичних джерел пов'язаних з тематикою роботи;

2) виконання аналіз сучасних моделей для вирішення схожих проблем;

- 3) імплементувати обрані моделі для пошуку аномалій;
- 4) скласти торговельний алгоритм, та перевірити його роботи;
5. Перелік ілюстративного матеріалу: презентація доповіді.
6. Дата видачі завдання: 15 листопада 2022 р.

Календарний план

№ з/п	Назва етапів виконання дипломної роботи	Термін виконання	Примітка
1	Узгодження теми роботи із науковим керівником	15-30 листопада 2022 р.	Виконано
2	Огляд опублікованих джерел за тематикою дослідження	Листопад-грудень 2022 р.	Виконано
3	Ознайомлення з методами побудови торговельних стратегії	Січень-лютий 2023 р.	Виконано
4	Ознайомлення з методами пошуку аномалій в часових рядах	15 січня - 10 лютого 2023 р.	Виконано
5	Проведення дослідження, з метою пошуку кращої моделі для задачі пошуку аномалій	Лютий 2023 р.	Виконано
6	Проведення досліджень, з метою знаходження ефективного торговельного алгоритму	Березень 2023 р.	Виконано
7	Реалізація моделі Prophet для задачі пошуку аномалій	15 - 30 березня 2023 р.	Виконано
8	Імплементация обраної торговельної стратегії	10 - 20 квітня 2023 р.	Виконано
9	Покращення результатів роботи моделі та підвищення прибутковості роботи алгоритму	10 - 20 травня 2023 р.	Виконано
10	Оформлення дипломної роботи	Травень 2023 р.	Виконано

Студентка

_____ Шевцова М.С.

Керівник

_____ Яворський О. А.

РЕФЕРАТ

Кваліфікаційна робота містить: 47 стор., 14 рисунків, 3 таблиці, 18 джерела.

Простір DeFi швидко розвивається і стає все більш популярним серед криптовалютних та фінансових учасників. Забезпеченням торговельних стратегій, які враховують особливості DeFi, ми можемо ефективно використовувати цей новий фінансовий простір для отримання вигоди та забезпечення успішних торгів.

Врахування аномалій у торговельних стратегіях у просторі DeFi допомагає зменшити ризики та зберегти капітал. Завдяки аналізу аномалій можна ідентифікувати потенційно небезпечні ситуації, забезпечити захист від великих змін цін та мінімізувати втрати. Це особливо важливо в сфері DeFi, де ризики можуть бути значними через високу волатильність та недостатню регуляцію. Темою даної роботи є створення торговельних стратегій з урахуванням інформації про аномалій у просторі DeFi. Одним з підходів є використання моделі прогнозування під назвою "Профет"(Prophet), а також трейдінгової моделі на основі інформації про аномалії, волатильність, показник Шарпа та ціни закриття для прийняття рішення щодо купівлі та продажу активів.

Використання такої моделі дозволяє нам підходити до торгівлі в просторі DeFi з більшою точністю за рахунок врахування аномалій.

Основний внесок цієї дипломної роботи полягає у розробці та застосуванні торговельних стратегій, які враховують інформацію про аномалії у просторі DeFi. Це допомагає трейдерам зменшити ризики та збільшити прибутковість своїх угод.

МАШИННЕ НАВЧАННЯ, СТРАТЕГІЇ ІНВЕСТУВАННЯ,
ДЕЦЕНТРАЛІЗОВАНІ ФІНАНСИ, ВИЯВЛЕННЯ АНОМАЛІЙ,
FACEBOOK PROPHET, ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ

ABSTRACT

The explanatory note of the diploma work consists of 47 pages and includes 3 tables and 14 figures. 18 bibliographic references were used for the research.

The DeFi space is rapidly evolving and gaining popularity among cryptocurrency and financial participants. By developing trading strategies that consider the unique characteristics of DeFi, we can effectively leverage this new financial space to gain advantages and ensure successful trades.

Incorporating anomalies into trading strategies in the DeFi space helps to reduce risks and preserve capital. Analyzing anomalies allows for the identification of potentially risky situations, protection against significant price changes, and minimizing losses. This is particularly important in the DeFi sphere, where risks can be substantial due to high volatility and insufficient regulation.

The topic of this work is the creation of trading strategies that take into account anomalies in the DeFi space. One approach is the use of a forecasting model called "Prophet" as well as a trading model based on information about anomalies, volatility, Sharpe ratio, and closing prices to make decisions regarding buying and selling assets.

Using such a model allows us to approach trading in the DeFi space with greater accuracy by considering anomalies.

The main contribution of this diploma work lies in the development and application of trading strategies that consider information about anomalies in the DeFi space. This helps traders reduce risks and increase the profitability of their trades.

MACHINE LEARNING, INVESTMENT STRATEGIES,
DECENTRALIZED FINANCE, ANOMALY DETECTION, FACEBOOK
PROPHET, TIME-SERIES FORECASTING

ЗМІСТ

Перелік умовних позначень, скорочень і термінів	8
Вступ.....	9
1 РОЗДІЛ. Децентралізовані фінанси та часові ряди	11
1.1 Децентралізовані фінанси.....	11
1.2 Часові ряди	17
1.3 Аномалії	21
2 ОГЛЯД АЛГОРИТМІВ ТА ДАНІ.....	25
2.1 Модель Prophet	25
2.2 Датасет.....	31
3 РЕЗУЛЬТАТИ МОДЕЛЮВАННЯ. ОЦІНКА ЕФЕКТИВНОСТІ МОДЕЛІ	34
3.1 Попередня обробка та аналіз даних	34
3.2 Реалізація моделі Prophet	36
3.3 Торговельна стратегія	40
3.4 Порівняння результатів моделі	42
Висновки	45
Перелік посилань	47

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

DeFi — Decentralized Finance

MA — Moving Average

ETH-USD — Ethereum USD

P2P — Peer-to-peer, (англ. — один-одному)

EOA — Externally Owned Accounts (англ. — зовнішні облікові записи)

CA — Contract Accounts (англ. — контрактні облікові записи)

PoW — Proof-of-work (англ. — доказ виконання роботи)

PoS — Proof-of-stake (англ. — Доказ частки володіння)

EVM — Ethereum Virtual Machine

APE — абсолютну відсоткову похибку (англ. — absolute percentage error)

ВСТУП

Актуальність дослідження. Розширення простору DeFi (Децентралізованих фінансів) в останні роки створило нові можливості для фінансових учасників, пропонуючи децентралізовані фінансові інструменти, послуги та платформи. Децентралізованість і прозорість DeFi забезпечують новий рівень фінансової свободи та доступу до послуг для користувачів.

Ринок DeFi характеризується високим рівнем коливань цін та, важливо зазначити, про існування в ньому, аномалій, що можуть впливати на ефективність торговельних стратегій. Трейдери шукають способи оптимізувати свої стратегії, враховуючи ці аномалії, для досягнення кращих результатів у торгівлі DeFi.

Дослідження пропонує використання інформацію про аномалії в процесі трейдингу у просторі DeFi. Це означає аналізувати та використовувати дані про непередбачувані події, зміни цін та ринкову динаміку для створення торговельних стратегій, що мають перевагу перед традиційними підходами.

Інтеграція інформації про аномалії в торговельні стратегії може дозволити мінімізувати ризики та зменшити можливі втрати, зробити більш обґрунтовані торговельні рішення.

Метою дослідження є побудова ефективної торговельної стратегії, що враховуватиме інформацію про аномалії, у фінансових часових рядах. Відповідно до мети дослідження, ми сформуваємо наступні **завдання**: 1. Дослідження різних методів побудови торговельної стратегії; 2. Визначення найкращої стратегії; 3. Дослідження різних методів для знаходження аномалій, визначенням найкращого, для даної стратегії; 4. Порівняння обраних моделей з моделями, які вже існують.

Об'єктом дослідження. Торговельні стратегії у просторі DeFi які враховують інформацію про аномалії.

Предметом дослідження Проблема побудови торговельної стратегії для децентралізованих фінансів, за допомогою використання машинного навчання. У цьому контексті, важливим аспектом дослідження є використання інформації про аномалії для поліпшення ефективності торговельних стратегій.

Методи дослідження: Аналіз даних, моделювання, програмні пакети для аналізу часових рядів.

Науковою новизною є детальний аналіз та оцінка ефективності моделі Prophet для ETH-USD, для пошуку аномалій та побудови торговельної стратегії.

Практичне значення. Побудова більш ефективних трейдингових стратегій для децентралізованих фінансів, зокрема, для такого класу як Ethereum. Що дозволяє компаніям, які працюють у сфері децентралізованих фінансів, використовувати результати цього дослідження для розробки власних торговельних стратегій, що будуть враховувати коливання цін та аномалії на ринку DeFi. Допоможе їм зробити більш обґрунтовані торговельні рішення, знизити втрати та забезпечити стабільний ріст свого бізнесу, що в свою чергу, може значно збільшити прибутковість та конкурентоспроможність таких компаній. Дозволить їм отримувати кращі результати на ринку DeFi, зменшувати ризики та підвищувати ефективність своїх операцій.

1 РОЗДІЛ. ДЕЦЕНТРАЛІЗОВАНІ ФІНАНСИ ТА ЧАСОВІ РЯДИ

ТЕОРЕТИЧНІ ВІДОМОСТІ

Цей розділ надає основи для розуміння DeFi, дозволяє ознайомитись з торговельними стратегіями та аналітичними методами, що використовуються в цьому просторі. Він також наголошує на важливості виявлення аномалій та розумінні особливості та методи аналізу часових рядів.

1.1 Децентралізовані фінанси

Означення 1.1. DeFi - це простір, де фінансові послуги надаються на основі блокчейн-технологій та смартконтрактів. Він характеризується високою волатильністю, швидкими змінами та широким спектром ризиків.

Децентралізована фінансова система (DeFi) є новою технологією у сфері фінансів, яка базується на безпечних розподілених реєстрах, подібних до тих, які використовуються в криптовалютах. DeFi перетворює традиційну централізовану фінансову систему, дозволяючи користувачам здійснювати безпосередні цифрові обміни між собою, уникати плати за користування банківськими послугами та зберігати кошти в безпечному цифровому гаманці. DeFi надає можливість швидкого та доступного фінансового обміну будь-якій особі з підключенням до Інтернету [1].

Децентралізовані фінанси дозволяють учасникам використовувати криптовалюту для надання більшості послуг, які традиційні банки пропонують з використанням державної валюти — позичати, запозичати, заробляти відсотки, торгувати активами, купувати страхування та багато іншого.

Послуги DeFi зазвичай є швидшими, дешевшими та простішими, з новими перевагами та послугами, які пропонуються щодня. DeFi є загальним терміном для послуг фінансових технологій на основі один-одному (англ. — Peer-to-peer) на публічних блокчейнах, переважно на Ethereum. За допомогою DeFi можна виконувати більшість операцій, які зазвичай підтримуються банками, такі як позичання, запозичення, купівля страховок, торгівля похідними інструментами, торгівля активами та інше. Однак, на відміну від традиційних банківських послуг, DeFi дозволяє здійснювати ці операції швидше та без необхідності займатися паперами [1], [2], [3] .

DeFi відрізняється від традиційних, централізованих фінансових установ і банків, де для відправлення або отримання грошей необхідні посередники, такі як банки чи фондові біржі. У DeFi ці посередники замінені програмним забезпеченням, і люди торгують безпосередньо один з одним, а "розумні контракти" на основі блокчейну виконують роль створення ринків, здійснення угод та забезпечення чесного та надійного процесу [1].

Ринок DeFi характеризується своєю специфічністю на сьогодні, але в майбутньому він має потенціал стати більш прозорою, відкритою та стабільною фінансовою інфраструктурою. Діаграма стеку DeFi ілюструє різні рівні та складові цього ринку, підкреслюючи його складність та можливості [4].

Означення 1.2. Ефіріум – це децентралізована, відкрита блокчейн-платформа, яка дозволяє створювати та виконувати смартконтракти. Вона була запропонована Віталієм Бутеріном наприкінці 2013 року і розроблена в 2014 році Ethereum Foundation. Основною криптовалютою Ethereum є Ефір (ETH), який використовується як засіб передачі значення в мережі Ethereum.

Ethereum є програмованим блокчейном, який акцентується на ідеї виконання кодів у децентралізованому середовищі. Це дозволяє розробляти децентралізовані додатки, відомі як DApps, без наявності

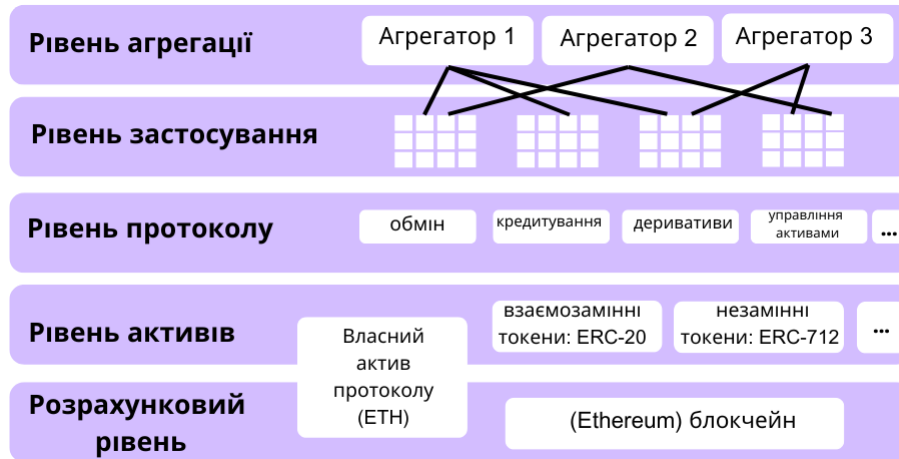


Рисунок 1.1 – DeFi стек

єдиного центрального пункту. Ці додатки можуть використовувати смартконтракти, які є автоматично виконуваними контрактами з умовами угоди, прямо записаними в кодї. Смартконтракти дозволяють автоматизувати, забезпечувати прозорість та безпечне виконання різних видів угод та транзакцій.

Ethereum використовує Ether (ETH) як свою власну криптовалюту, в якій майнери вирішують головоломки, щоб отримувати ETH. Поточна ринкова капіталізація Ethereum становить \$220,97 млрд. Платформа також базується на другому виді токенів, відомих як газ, який потрібен при виконанні кожного смартконтракту для стимулювання майнерів до майнінгу та ланцюгування нових транзакцій.

Мережа Ethereum працює на механізмі доказу роботи (PoW), який знаходиться в процесі переходу до механізму доказу частки володіння (PoS), відомого як Ethereum 2.0. Цей перехід має на меті покращення масштабованості, безпеки та енергоефективності. Ефіріум набув значної популярності і став основою для великої екосистеми проектів децентралізованої фінансової сфери (DeFi), токенів незамінних активів (NFT), децентралізованих бірж та інших інноваційних додатків. Його відкритість привернули розробників, підприємців та інвесторів для дослідження та побудови на його блокчейн-платформі.

Крім того, Ethereum пропонує дві категорії рахунків: зовнішні

рахунки (EOA) та контрактні облікові записи (CA), які ідентифікуються 20-байтовими адресами. Будь-яка зміна стану блокчейну Ethereum ініціюється зовнішнім рахунком, як показано на Рис. 2. Це означає, що єдиним способом взаємодіяти з кодом будь-якого контракту є ініціювання транзакції з зовнішнього рахунку, яка містить входні параметри, необхідні для виконання коду контракту. Крім того, Ethereum пропонує спеціальний тип операцій, які дозволяють запитувати інформацію з блокчейну без будь-яких комісій, відомі як "виклики" (англ. – Calls), які використовуються для запити даних IoT-пристроїв у нашій запропонованій рамці.

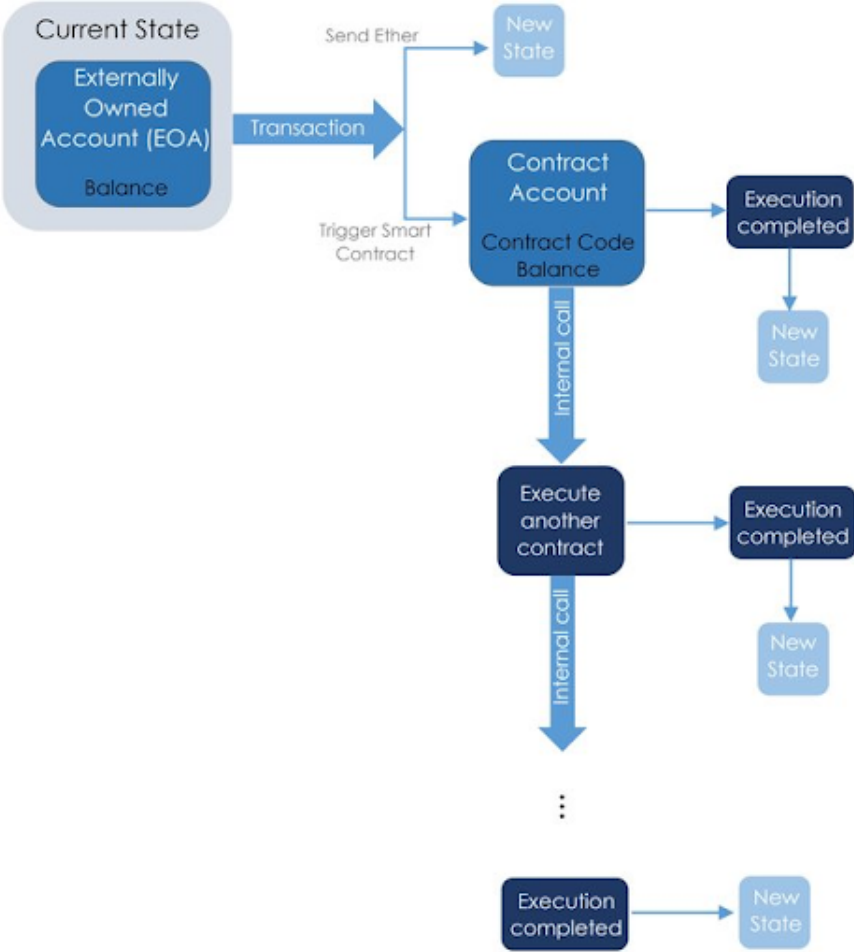


Рисунок 1.2 – Схема робочого процесу транзакцій Ethereum.

Газ належать до одиниці, яка вимірює кількість обчислювальних зусиль, необхідних для виконання певних операцій в мережі Ethereum.

Оскільки кожна транзакція на Ethereum вимагає обчислювальних ресурсів для виконання, для кожної транзакції потрібна комісія. Газ належать до комісії, необхідної для виконання транзакції на Ethereum, незалежно від успіху або невдачі транзакції.

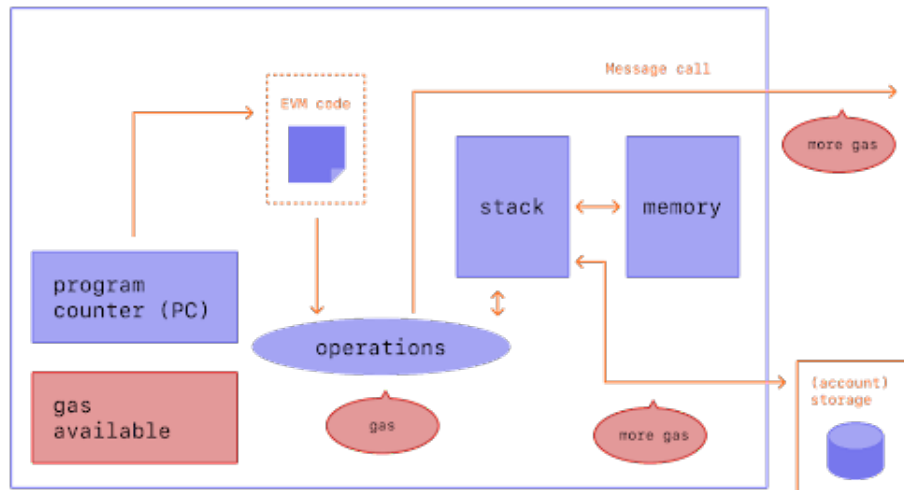


Рисунок 1.3 – Діаграма робочого процесу Ethereum транзакції.

Ця діаграма показує послідовність етапів, які відбуваються під час виконання транзакції на платформі Ethereum, включаючи такі елементи, як доступний газ, лічильник програми (pc), EVM-код, виклик повідомлення та зберігання облікових записів.

Комісії за газ сплачуються в основній валюті Ethereum – ефірі (ETH). Ціни на газ виражаються в гвей, що є під одиницею ETH – кожен гвей дорівнює 0.000000001 ETH (10^{-9} ETH). Наприклад, замість того, щоб сказати, що ваша комісія за газ коштує 0.000000001 ефіра, ви можете сказати, що ваша комісія за газ коштує 1 гвей. Сам термін "гвей" означає "гіга-вей" і дорівнює 1,000,000,000 вей. Вей сам по собі він є найменшою одиницею ETH [5].

Кожен блок має базову комісію, яка виступає як резервна ціна. Щоб бути прийнятим до блоку, запропонована ціна за газ повинна не менше базової комісії. Базова комісія обчислюється незалежно від поточного блоку і визначається попередніми блоками – це робить комісії за

транзакції більш передбачуваними для користувачів. При видобуванні блоку, базова комісія "випаюється" вилучаючи її з обігу.

Базова комісія обчислюється за формулою, яка порівнює розмір попереднього блоку (кількість газу, використаного для всіх транзакцій) з цільовим розміром. Базова комісія збільшується максимум на 12,5% на кожному блоку, якщо цільовий розмір блоку перевищено. Цей експоненційний ріст робить нерентабельним економічно підтримувати високий розмір блоку нескінченно.

Комісії за газ допомагають забезпечити безпеку мережі Ethereum. Вимагаючи комісію за кожну виконану обчислювальну операцію в мережі, запобігається спам-мережі небажаним суб'єктам. Щоб уникнути випадкових або зловмисницьких нескінченних циклів або інших марнослівних обчислень у кодї, кожна транзакція повинна встановити обмеження на кількість обчислювальних кроків виконання коду. Основною одиницею обчислення є "газ"[6]. Всупереч того, що транзакція включає обмеження, будь-який не використаний газ в транзакції повертається користувачу (тобто максимальна комісія – $(\text{max fee} - (\text{base fee} + \text{tip}))$ повертається).

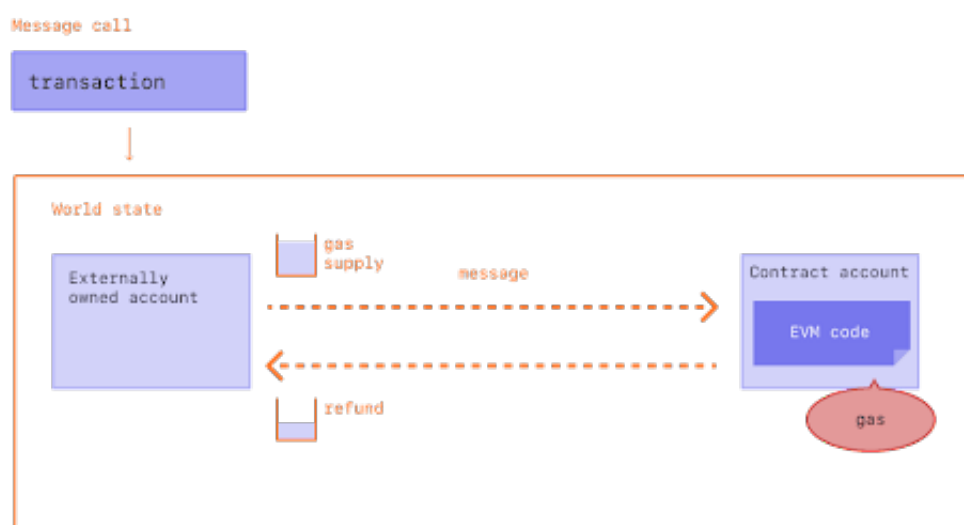


Рисунок 1.4 – Діаграма транзакції.

1.1.1 Торговельні стратегії

Означення 1.3. Торговельні стратегії – це набори об’єктивних правил, які визначають умови для здійснення угоди на вході та виході [7]. Вони використовуються для керування інвестиційними рішеннями та торговельними планами і включають умови для входу в угоду, виходу з угоди, толерантності до ризику та горизонту часу торгівлі [8].

Існує різноманітність типів торговельних стратегій, включаючи торгівлю на основі новин, торгівлю в кінці дня, денну торгівлю, торгівлю за трендом, скайпінг та позиційну торгівлю [9]. Найкраща торговельна стратегія залежить від індивідуальних уподобань, толерантності до ризику та стилю інвестування [10].

Початкові торговельні стратегії використовують найпростіші показники та підходи [7]. Фундаментальні торговельні стратегії спираються на фундаментальні фактори, тоді як технічні торговельні стратегії використовують технічні показники, такі як ковзні середні [11].

1.2 Часові ряди

Означення 1.4. Часовий ряд – це послідовність спостережень за певним параметром у різні моменти часу. Він містить інформацію про те, як змінюється параметр з часом.

Тобто, послідовність точок даних, індексованих у порядку часу. Такі сфери, як фінанси, економіка, прогнозування погоди та управління операціями використовують цей тип даних для відстеження змін у часі.

Властивості часових рядів:

1) Тренд (англ. Trend): Довгострокові зміни у ряду, такі як зростаючий або спадний шаблон.

2) Сезонність (Seasonality): Циклічні шаблони, які повторюються після певного інтервалу часу.

3) Циклічність (англ. Cyclicity): Флуктуації навколо тенденції, які можуть бути спричинені економічними або ринковими циклами.

4) Нерегулярності (англ. Irregularities): Випадкові, не пояснені варіації в даних, які можуть бути спричинені факторами, такими як помилки вимірювання або неочікувані події.

5) Автокореляція (англ. Autocorrelation): Кореляція між поточним значенням змінної та його попередніми значеннями, що може вказувати на наявність шаблонів і взаємозв'язків у даних.

6) Викиди (англ. Outliers): Точки даних, які значно відрізняються від решти даних, що може бути спричинено факторами, такими як помилки вимірювання або екстремальні події.

7) Шум (англ. Noise): Випадкові флуктуації в даних, які можуть ускладнити виявлення основних шаблонів та взаємозв'язків [12].

Означення 1.5. Послідовність випадкових величин

$$\{Y_t : t = 0, \pm 1, \pm 2, \pm 3, \dots\}$$

називається стохастичним процесом і слугує моделлю для спостережуваного часового ряду.

Відомо, що повна імовірнісна структура такого процесу визначається набором розподілів всіх скінченних наборів Y . Багато інформації в цих спільних розподілах можна описати в термінах середніх, дисперсій та коваріацій.

Означення 1.6. Для випадкового процесу $\{Y_t : t = 0, \pm 1, \pm 2, \pm 3, \dots\}$ функція середнього визначається, як $\mu_t = E(Y_t)$ для $t = 0, \pm 1, \pm 2, \dots$

Це означає, що μ_t є очікуваною величиною процесу в часі t . Загалом, μ_t може бути різним на кожній точці часу t .

Означення 1.7. Автоковаріаційна функція $\gamma_{t,s}$ визначається, як

$$\gamma_{t,s} = Cov(Y_t, Y_s) \tag{1.1}$$

для $t, s = 0, \pm 1, \pm 2, \dots$, де $Cov(Y_t, Y_s) = E[(Y_t - \mu_t)(Y_s - \mu_s)] = E(Y_t Y_s) - \mu_t \mu_s$.

Функція автокореляції $\rho_{t,s}$ визначається, як

$$\rho_{t,s} = Corr(Y_t, Y_s) \quad (1.2)$$

для $t, s = 0, \pm 1, \pm 2, \dots$, де $Corr(Y_t, Y_s) = \frac{Cov(Y_t, Y_s)}{\sqrt{Var(Y_t)Var(Y_s)}} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}}$

Згідно з характеристиками аналізу часових рядів виникають такі методи, виявлення та аналізу часових рядів:

(1) зміна довгострокового тренду: під впливом деяких основних факторів дані показують певну тенденцію при зміні у часі, яка збільшується або зменшується стабільно згідно з певними правилами. Використовуються такі методи аналізу: метод ковзного середнього, метод експоненційного згладжування, метод побудови моделі тощо.

(2) Зміна сезонного циклу: під впливом сезонної зміни та інших факторів послідовність регулярно змінюється за фіксованим циклом, що також називається бізнес-циклом. Методи: сезонні індекси.

(3) Циклічні зміни: коливальні зміни з нефіксованими періодами;

(4) Випадкові зміни: послідовність змін, спричинені багатьма факторами. Використовується метод аналізу часового ряду. Як показано на рисунку 5.



Рисунок 1.5 – Метод аналізу часових рядів.

Параметри для часових рядів даних включають: Інтервали часу: Частота, з якою збираються точки даних, такі як щоденні, щотижневі, щомісячні або щорічні. Затримкові ознаки (англ. Lag features): Попередні значення ряду, які можуть бути використані як вхідні ознаки для моделей прогнозування. Ознаки вікон (англ. Window features): Обчислені значення на основі вікна попередніх або майбутніх періодів, такі як ковзні середні або стандартні відхилення [13][18].

Після перед опрацювання транзакцій з них береться набір (англ. Feature extraction) ознак у форматі часових рядів. Для аналізу часових рядів використовується набір числових ознак, які витягуються за фіксованим дискретним інтервалом часу, наприклад, 1 день. Тому до набору перепрацьованих транзакцій застосовується ковзне вікно, і ознаки витягуються послідовно.

Нехай W і s є розміром вікна та кроком зсуву вікна відповідно. Ми припускаємо, що для кожного дискретного часового індексу t_j існує мультіваріантний набір p ознак $f_i(t_j)$, де $i \in [1, p]$ і $j \in [1, n]$. Також позначимо вектор ознак часового ряду як $\mathcal{F} = f_i(t_i)$. Тому $s = t(j + 1) - t_j(1.3)$, де $j \in [1, W - 1]$.

Два ключових поняття в аналізі часових рядів - це автокореляція та стаціонарність.

Автокореляція належить до лінійного зв'язку між часовим рядом та його варіантом з запізненням. Вона важлива, оскільки багато аналітичних інструментів, включаючи регресійний аналіз, передбачають відсутність автокореляції в даних.

Автокореляція допомагає виявляти сезонність та тенденції в даних часових рядів і необхідна для вибору відповідної моделі ARIMA для прогнозування часових рядів.

Стаціонарність є ще одним важливим поняттям в аналізі часових рядів. Багато статистичних моделей та тестів базуються на стаціонарності для здійснення прогнозування. Якщо часовий ряд є нестаціонарним, може бути необхідно перетворити його на стаціонарний процес для

застосування певних функцій прогнозування. Один з поширених методів досягнення стаціонарності - це пошук різниць, який може використовуватися для видалення небажаних тенденцій у часовому ряді.

При аналізі часових рядів важливо враховувати, чи є дані стаціонарними, чи є сезонність і чи є змінні автокорельованими. Ці фактори впливають на вибір моделей та точність прогнозів. Існують різноманітні моделі аналізу часових рядів, включаючи:

- Авторегресія (AR)
- Ковзне середнє (MA)
- Авторегресія-ковзна середня (ARMA)
- Авторегресія з інтегрованою ковзною середньою (ARIMA)
- Авторегресія-ковзна середня з екзогенними входами (ARMAX)

Розуміння автокореляції та стаціонарності є важливими в аналізі часових рядів, оскільки вони впливають на вибір моделей та точність прогнозів. Аналіз цих концепцій допомагає виявляти тенденції, сезонність та інші закономірності у даних, що в кінцевому рахунку допомагає зробити кращі прогнози та інформаційні рішення [14] .

1.3 Аномалії

Ринковою аномалією на фінансовому ринку є передбачуваність, яка, здається, суперечить (зазвичай заснованим на ризику) теоріям цін на активи. Аномалії можуть розумітися, як аутлайери або, як не характерні шматки часового ряду.

Аномалії в просторі DeFi можуть бути спричинені факторами, такими як цінові шоки, протокольні помилки, атаки зловмисників, непередбачувані зміни регуляторної політики та інші небажані події.

Більшість, документів, які документують аномалії, пояснюють їх неправильним ціноутворенням (наприклад, Лаконішок, Шеліфер і Вісний 1994). Пояснення неправильного ціноутворення є природним, оскільки аномалії за визначенням є відхиленнями від базової теорії цін на активи.

Тоді «неправильне ціноутворення» визначається, як відхилення відносно контрольного показника.

Найпоширенішим еталонним показником є CAPM (англ. Capital-Asset-Pricing Model, модель оцінки капіталу-активів). Відхилення від цієї теорії вимірюється ненульовим перехопленням оціненої лінії ринку цінних паперів. Це перехоплення зазвичай позначається грецькою літерою альфа:

$$E(R_t - R(f,t)) = \alpha + \beta[E(R(M,t) - R(f,t))] \quad (1.3)$$

де - R_t віддача від аномалії, $R(f,t)$ – прибутковість безризикової ставки, β – це нахил від регресії віддачі від аномалії до віддачі ринку, і $R(M,t)$ – це прибутковість «ринку», яка часто визначається прибутковістю індексу CRSP (індексу всіх публічних акцій США).

Виявлення аномалій - це процес у машинному навчанні, який ідентифікує точки даних, події та спостереження, що відхиляються від нормальної поведінки набору даних. Існують три основні методи виявлення аномалій:

1) Пошук аномалій, за допомогою навчання з вчителем (supervised anomaly detection): Ця техніка передбачає тренування моделі на попередньо позначених даних, що містять заздалегідь визначені нормальні дані та чітко позначені приклади аномалій [15].

2) Напівавтоматичне навчання з учителем для виявлення аномалій (semi-supervised anomaly detection): Використання методів півнаглядного навчання для виявлення аномалій, на відміну від навчання з учителем, не вимагає наявності позначених аномальних записів у даних. Замість цього методи передбачають, що у наборі даних доступні лише записи з нормального класу [16]

3) Пошук аномалій, за допомогою навчання без вчителя (unsupervised anomaly detection): Ця техніка передбачає, що дані не мають позначень і є найбільш поширеною через її широке та відповідне застосування.

Висновки до розділу 1

Перший розділ даної дипломної роботи присвячений огляду теоретичних відомостей та ознайомленню з базовими поняттями. У рамках цього розділу були ретельно проаналізовані ключові поняття, такі як децентралізовані фінанси, криптовалюта Ethereum, часові ряди, аномалії у часових рядах, торговельні стратегії та методи пошуку аномалій.

В процесі цього аналізу ми з'ясували, які метрики найкраще використовувати при роботі з часовими рядами та їх аналізі. Також ми детально вивчили принципи роботи та особливості криптовалюти Ethereum. Було виявлено такі переваги Ethereum, як масштабованість, безпека та енергоефективність, що робить її ідеальним вибором для використання в децентралізованих фінансах. Крім того, з допомогою огляду літератури, ми прийшли до висновку, що часові ряди є нативним способом представлення інформації про ціни акцій, тому використання часових рядів для побудови торговельних алгоритмів є доцільним та перспективним підходом.

Під час нашого дослідження ми також вказали на існування такого об'єкта, як аномалії та прийшли до висновку, що їх використання допомагає виявляти потенційні ризики та небезпеки. Коли система виявляє незвичайні аномальні показники, це може свідчити про можливі проблеми або небезпеку на ринку. Такі сигнали допомагають приймати обґрунтовані рішення щодо ризикових позицій або виведення з ринку, що робить їх незамінними елементами при розробці торговельної стратегії.

Використання часових рядів дозволяє нам аналізувати та прогнозувати динаміку цін та рухи ринку. Ми можемо виявити патерни та тренди, що допомагають нам приймати обґрунтовані рішення щодо входу в позицію чи виходу з ринку. Засновуючись на історичних даних, ми можемо визначити оптимальні точки для виконання торговельних операцій.

Водночас, використання аномалій у часових рядах надає нам можливість виявляти незвичайні події та непередбачувані зміни, що можуть вплинути на ринок. Це дозволяє нам вчасно реагувати на ризики та небезпеки, а також знаходити потенційні можливості для отримання прибутку. Використання аномалій у торговельній стратегії допомагає нам забезпечити більш надійні та ефективні рішення.

Таким чином, знання про часові ряди, аномалії та їх використання в торговельних стратегіях становлять важливу основу для подальшої розробки нашої стратегії. Вони дозволяють нам більш глибоко аналізувати ринок, розпізнавати непередбачувані ситуації та приймати обґрунтовані торговельні рішення.

2 ОГЛЯД АЛГОРИТМІВ ТА ДАНІ

В даному розділі буде розглянуто важливі аспекти, зокрема визначення вхідних даних - історичних цін на актив, та очікувані результати, які включають стратегію торгівлі, що базується на інформації про аномалії. Також буде розглянуто використання моделі Prophet для виявлення аномалій та їх впливу на прийняття рішень щодо купівлі та продажу активів. Додатково буде проаналізовано прогнозування за допомогою моделі Prophet, враховуючи сезонність, тренди, святкові ефекти та інші фактори, що впливають на дані. Крім того, буде проведений аналіз торговельної стратегії, включаючи оцінку метрик профітності та інших параметрів для оцінки ефективності стратегії торгівлі.

2.1 Модель Prophet

Модель Prophet є моделлю прогнозування часових рядів великого масштабу, запропонованою компанією Facebook у 2017 році [17]. Вона розкладає часовий ряд на складові тренду, сезонності та святкових подій, як показано у формулі:

$$y(t) = g(t) + s(t) + h(t) + \epsilon(t) \quad (2.1)$$

де t - це час; $y(t)$ - часовий ряд; $g(t)$ - трендова складова, який відповідає за аперіодичну частину часового ряду; $s(t)$ - відноситься до сезонності і представляє періодичні зміни даних часового ряду; $h(t)$ - ефект свят, що вказує на зміни, спричинені особливими обставинами, такими як свята; $\epsilon(t)$ - частина шуму часового ряду, що вказує на випадкові флуктуації, які не можна передбачити.

Модель Prophet головним чином складається з трьох моделей:

моделі тренду, моделі сезонності та моделі свят. Модель Prophet має два види функцій зростання тренду: нелінійну функцію зростання і лінійну функцію зростання, які підходять для апроксимації різних видів часових рядів. Крім того, сезонна частина моделі Prophet використовує ряд Фур'є для наближення, щоб гнучко підігнатися під набори даних з різними правилами періодичних змін.

Модель сегментованого лінійного зростання реалізується за допомогою сегментованої лінійної функції з таким рівнянням:

$$g(t) = (k + \alpha(t)\delta) \times \tau + (m + \alpha(t)^T \gamma) \quad (2.2)$$

де k - це темп зростання, m - зсув, δ - кількість зміни k в темпі зростання.

Нелінійна модель зростання реалізується за допомогою логістичної функції з таким рівнянням:

$$g(t) = \frac{C(t)}{1 + e^{-k(t)(t-m(t))}} \quad (2.3)$$

де $C(t)$ позначає межу насичення, $k(t)$ позначає темп зростання, $m(t)$ позначає зсув.

Часові ряди, як правило, містять різноманітні щорічні, місячні, тижневі, денні та інші типи циклічних сезонних тенденцій.

Модель Prophet використовує ряди Фур'є для представлення періодичних змін у часових рядах, конкретний вираз наступний:

$$s(t) = \sum_{n=1}^N (a_n \cos(\frac{2\pi nt}{T}) + b_n \sin(\frac{2\pi nt}{T})) \quad (2.4)$$

Де T позначає період, а $2n$ позначає кількість періодів [18].

Algorithm 2.1 Алгоритм використання моделі Prophet

Вхід: часові ряди, $y_{t\text{rain}}$; дати збору даних, $date_{t\text{rain}}$

Вихід: прогнози значення, $y_{p\text{rediction}}$; верхня та нижня межі довірчого інтервалу, y_{upper} , y_{lower}

Перевести дані до формату $ds_{t\text{rain}}$

Ініціалізувати модель Prophet

Побудувати об'єкт часового ряду з вхідних даних $y_{t\text{rain}}$ та $date_{t\text{rain}}$

Навчити модель на вхідних даних

Згенерувати прогноз для майбутніх значень часового ряду

Отримати прогнози значення $y_{p\text{rediction}}$

Отримати верхні та нижні межі довірчого інтервалу y_{upper} , y_{lower}

Повернути $y_{p\text{rediction}}$, y_{upper} , y_{lower}

Щодо використання моделі Prophet, яка показана в алгоритмі 1, ми вводимо змінну часового ряду (y) разом з форматованою датою (ds), яка відповідає змінній. Потім ми налаштуємо параметри тренду та сезонності моделі Prophet, щоб отримати оптимальні результати прогнозування. Після цього ми робимо прогноз на майбутнє, вимагаючи введення прогнозованого періоду часу та частоти прогнозування. Вихід моделі Prophet включає значення підгонки та верхню та нижню межі довірчого інтервалу.

Модель Prophet для пошуку аномалій є інструментом прогнозування, розробленим компанією Facebook для обробки часових рядів. Вона особливо корисна для роботи з даними, що мають сезонні відмінності та тенденції, які часто спостерігаються в фінансових даних.

Переваги використання моделі Prophet включають:

– Стійкість до відсутніх даних та викидів: Модель може працювати з прогалинами та неправильностями в даних без суттєвого впливу на прогноз.

– Гнучкість у врахуванні свят та особливих подій: Модель може враховувати відомі події, які впливають на дані, що поліпшує точність прогнозу.

– Автоматичне виявлення сезонних відмінностей та тенденцій:

Модель може ідентифікувати та коригувати сезонність у даних, що робить її придатною для фінансових часових рядів.

– Простота використання та інтерпретації: Модель відносно проста у реалізації та розумінні, що робить її доступною для широкого спектру користувачів.

– Для пошуку аномалій за допомогою Prophet використовується набір функцій, які дозволяють моделювати та прогнозувати часовий ряд з урахуванням різних факторів [19].

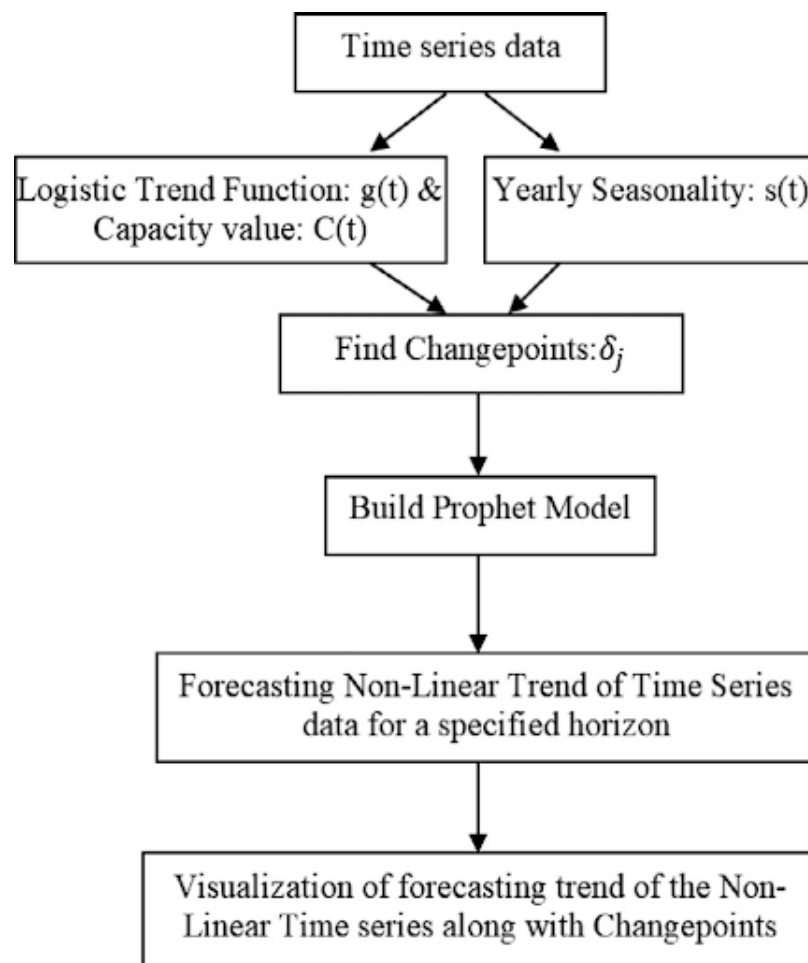


Рисунок 2.1 – Схема роботи моделі Prophet.

У першій ітерації ми налаштовуємо модель Prophet з урахуванням щоденної сезонності. Далі ми додаємо місячну та тижневу сезонності. Після цього ми визначаємо індекс обрізання (cutoff date) і кількість днів, які ми хочемо прогнозувати. Виконуємо підгонку моделі до даних і

генеруємо прогноз на певну кількість днів. Після цього виявляємо аномалії у прогнозі та оновлюємо датафрейм з прогнозованими значеннями. Нарешті, ми візуалізуємо початкові дані, прогноз та аномалії на графіку.

У другій ітерації ми виконуємо аналогічні кроки, але з іншим індексом обрізання та кількістю днів для прогнозу.

Загалом, за допомогою моделі Prophet виявляємо аномалії у часових рядах. Це допомагає отримати більш точні прогнози і зробити обґрунтовані рішення на основі аналізу часових рядів.

Запропонована нижче модель, є адаптацією моделі прогнозування Prophet для задачі виявлення аномалій. Вона базується, на порівнянні значень спостереження з передбаченими та виявленні значні відхилення. Ось загальний алгоритм виявлення аномалій за допомогою Prophet:

1) Збір даних часового ряду, які хочемо проаналізувати на предмет аномалій.

2) Попередня обробка даних. Prophet очікує DataFrame з двома стовпцями: 'ds' для міток часу та 'y' для спостережених значень.

3) Ініціалізувати модель Prophet з потрібною конфігурацією. Це включає визначення потрібних сезонних компонентів та будь-яких додаткових налаштувань.

4) Застосувати модель на обраних даних, за допомогою методу fit.

5) Згенерувати прогнози для всього часового ряду за допомогою методу predict.

6) Порівняти спостереженні значення з передбаченими, щоб виявити аномалії. Існує кілька підходів:

– а Розрахувати залишки, віднімаючи передбачені значення від спостережуваних: $residuals = observed_values - predicted_values$.

– б Розрахувати абсолютну відсоткову похибку (APE), взявши абсолютну різницю між спостереженнями та передбаченими значеннями, поділену на спостережувані значення: $ape = abs((observed_values - predicted_values) / observed_values)$.

– в Використовувати статистичні показники, такі як Z -оцінки або довірчі інтервали, для виявлення аномалій. Наприклад, розрахувати Z -оцінку залишків або APE і вважати значення, що перевищують певний поріг, аномаліями.

7) Визначити поріг для виявлення аномалій на основі обраної міри. Це можна зробити, аналізуючи розподіл залишків або APE та визначаючи точку відсічення, вище якої значення вважаються аномальними.

8) Позначити дані точки, які перевищують поріг, як аномалії.

9) Візуалізувати вихідні дані часового ряду з позначеними аномаліями, щоб зрозуміти виявлені аномалії.

Algorithm 2.2 Використання моделі Prophet для задачі пошуку аномалій

Попередня обробка даних

```
preprocessed-data = preprocess-data(raw-data)
```

Ініціалізація моделі Пророк

```
model = Prophet(configuration)
```

Прогон моделі по даних

```
model.fit(preprocessed-data)
```

Генерація прогнозів

```
predictions = model.predict(preprocessed-data)
```

Розрахунок залишків або APE

```
residuals = preprocessed-data['y'] - predictions['yhat']
```

```
ape = abs((preprocessed-data['y'] - predictions['yhat'])
```

```
/ preprocessed-data['y'])
```

Визначення порогу

```
threshold = calculate-threshold(residuals, ape)
```

Виявлення аномалій

```
anomalies = identify-anomalies(residuals, ape, threshold)
```

Візуалізація результатів

```
visualize-data(preprocessed-data, anomalies)
```

2.2 Датасет

Для використання нашої моделі було розглянуто декілька різних варіантів даних, які були агреговані з Google та Kaggle. Кінцевою даними були обрані YFINANCE, зважаючи на його якості властивості та актуальність. Датасет, містить історичні дані про ціни криптовалюти Ethereum (ETH) у доларах США.

Обраний проміжок часу, в зазначеному датасеті починається від 09 листопада 2017 року по 22 травня 2023 року. Частота даних – щоденно. Тип даних – часові ряди.

Цей датасет містить інформацію про ціни Ethereum у доларах США на щоденній основі, що дозволяє вивчати його динаміку та тренди протягом зазначеного періоду. Він може бути використаний для аналізу руху цін Ethereum, виявлення аномалій, розробки торговельних стратегій та дослідження впливу подій у просторі DeFi на ціни криптовалюти.

За допомогою цього датасету можна вивчити характеристики цін Ethereum, такі як зміна вартості з часом, коливання, та знайти відношення між ціною Ethereum та іншими фінансовими даними для подальшого аналізу та прийняття рішень.

Date	Open	High	Low	Close	Adj Close
2017-11-09	308.64	329.45	307.06	320.88	320.88
2017-11-10	320.67	324.71	294.54	299.25	299.25
2017-11-11	298.59	319.45	298.19	314.68	314.68
2017-11-12	314.69	319.15	298.51	307.90	307.90
2017-11-13	307.02	328.41	307.02	316.71	316.71

Таблиця 2.1 – Приклад даних з обраного датасету

В таблиці представлений приклад даних, з якими ми працюємо. Де Open (укр. відкриття): це ціна активу або індексу на початку торговельного періоду; High (укр. максимальна ціна): це найвища ціна,

яку досягнув актив або індекс протягом торговельного періоду; Low (укр. мінімальна ціна): це найнижча ціна, яку досягнув актив або індекс протягом торговельного періоду; Adjusted Close (укр. скориговане закриття): це ціна закриття, враховуючи корекції, такі як ділити акції, спліт акцій або виплати дивідендів, дозволяє порівнювати ціни на різних етапах ігрового процесу, враховуючи вплив корекцій; Volume (укр. обсяг): це кількість акцій або контрактів, які були торговані протягом даного періоду. Обсяг може вказувати на зацікавленість торговців та ліквідність ринку; Close (укр. ціна закриття) – це дані про закриття, вимірює кінцеві значення певних величин або параметрів протоколу, наприклад, обсягу торгів чи вартості активів.



Рисунок 2.2 – Графік трендів Ethereum за обраний період.

Висновки до розділу 2

У цьому розділі була розглянута модель побудови торговельної стратегії, що буде спиратись на інформацію про аномалії, в процесі прийняття рішень.

Був проведений огляд моделі Prophet, яка використовується для пошуку аномалій. Зазначено переваги моделі Prophet, такі, як: простоту

використання, висока точність та гнучкість, що дозволяє ефективно виявляти незвичайні зміни та потенційні аномалії в даних. Крім того, наявність ряду вбудованих функцій та можливість роботи з екзогенними факторами дозволяють враховувати додаткові впливи на часовий ряд та забезпечують більш точний аналіз.

Зважаючи на гнучкість моделі Prophet вона дозволяє виявляти аномалії в даних. Також, оскільки модель Prophet є доволі універсальною та простою в використанні, її результати легко інтерпретувати, а тому можна використовувати велику кількість додаткових суміжних алгоритмів, для передбачення аномалій.

3 РЕЗУЛЬТАТИ МОДЕЛЮВАННЯ. ОЦІНКА ЕФЕКТИВНОСТІ МОДЕЛІ

3.1 Попередня обробка та аналіз даних

Беремо вище зазначений датасет Ethereum USD (ETH-USD) в проміжку від 09 листопада 2017 року по 22 травня 2023 року.



Рисунок 3.1 – Денний тренд ціни ETH-USD.

Для роботи з моделлю Prophet потрібно мати дві колонки вхідних даних, які називаються "ds" і "y". Колонка "ds" (formatted date): колонка містить форматовану дату або часову мітку для кожного точкового часу в часовому ряді. Вона використовується для моделювання та прогнозування залежності між часом і значеннями в часовому ряді.

Колонка "y" (time series variable): колонка містить значення часового

ряду, які будуть проаналізовані та прогнозовані. Вона може представляти числові дані, такі як ціни акцій, продажі, кількість відвідувачів тощо.

Ці колонки визначають залежність між часом і значеннями часового ряду. Їх колонок дозволяє моделі Prophet автоматично розпізнавати та моделювати сезонність, тренди та інші характеристики часового ряду, що полегшує прогнозування майбутніх значень.

В нашому випадку колонка "ds" – містить щоденні дані, за заданий період, а колонка "y" – містить числові дані про кінцеві значення ETH-USD величин.

На основі аналізу графіка – Рис. 3.1. денного тренду ціни ETH-USD, можна спостерігати такі основні зміни:

1) З 2017 року до початку 2018 року: Ціна ETH-USD демонструвала суттєвий ріст. Це був період швидкого зростання ціни Ethereum, коли багато людей зацікавилися цією криптовалютою.

2) Від 2018 до 2021 року: Протягом цього періоду спостерігалось загальне падіння ціни Ethereum. Ціна підлягала значним коливанням і втратила значну частину свого вартості. Це може бути пов'язано зі змінами на ринку криптовалют, регуляторними втручаннями і іншими факторами, такими як пандемія Коронавірусу.

3) Від 2022 року: Починаючи з 2022 року, ціна Ethereum почала стрімко зростати. Цей період характеризується значним підйомом ціни ETH-USD. Ймовірно, це може бути пов'язано зі зростанням інтересу до криптовалют та блокчейн-технологій, а також зі збільшенням використання Ethereum у децентралізованих фінансових (DeFi) та інших проектах.

4) Зараз: Останнім часом спостерігається знову зниження ціни Ethereum. Це може бути пов'язано зі змінами на ринку криптовалют, включаючи волатильність біткоіна та загальний стан ринку криптовалют.

Варто зазначити, що ці зміни в ціні Ethereum є результатом багатьох факторів, включаючи попит та пропозицію, інтерес інвесторів, новини та події в галузі криптовалют. Тому важливо розуміти, що ринок

криптовалют може бути дуже непередбачуваним, і прогнозування цін на основі попередніх трендів може бути складним.

3.2 Реалізація моделі Prophet

Алгоритм виявлення аномалій у часових рядах базується на виявленні аномалій на основі інтервалу прогнозування часового ряду.

Виконання включає такі кроки:

- 1) Побудувати модель прогнозування часового ряду.
- 2) Здійснити прогнози на історичних даних за допомогою моделі прогнозування часового ряду.
- 3) Порівняти фактичні значення з інтервалами прогнозування. Аномалії визначаються як точки даних з фактичними значеннями поза інтервалами прогнозування.
- 4) Налаштування.

Цей алгоритм базується на порівнянні фактичних значень з очікуваним діапазоном значень, що отримані з моделі прогнозування. Аномалії вважаються точками даних, які виходять за межі цього інтервалу.

На рисунку 10 показано загальну структуру процесу моделювання часових рядів у моделі Prophet.

Перш ніж використовувати модель Prophet для прогнозування, ми розділяємо набір даних. У цій роботі ми розділяємо набір даних на тренувальну та тестову вибірку відповідно до співвідношення 9 до 1. Розділена тренувальна вибірка використовується для навчання моделі та знаходження оптимальних ваг, а потім навчена модель використовується для прогнозування майбутнього періоду.

При ініціалізації моделі Prophet, параметри `yearly-seasonality` та `weekly-seasonality` встановлюємо на значення `True`, а потім модель навчається на тренувальних даних. Параметр `interval-width` встановлюється на значення `0.99`, що означає, що інтервал невизначеності становить 99%.

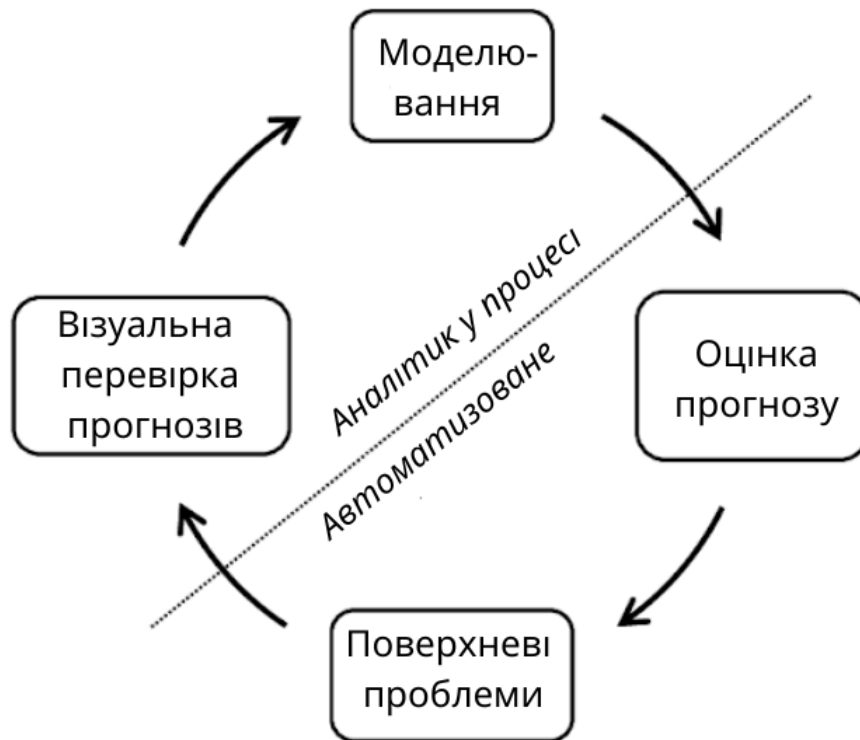


Рисунок 3.2 – Модель Prophet.

Проводимо нормалізацію цільової змінної.

Далі робимо передбачення з використанням моделі Prophet.

Графік прогнозу показує, що прогнози загалом узгоджуються з фактичними значеннями. Ми також можемо перевірити графік компонентів щодо тенденції, тижневої сезонності та річної сезонності.

На графіку представлені чотири графіку аналізу компонентів моделі Prophet. З їх допомогою можна зрозуміти внутрішні характеристики та тренди. Наприклад, можна спостерігати зростання цін закриття акцій у середу.

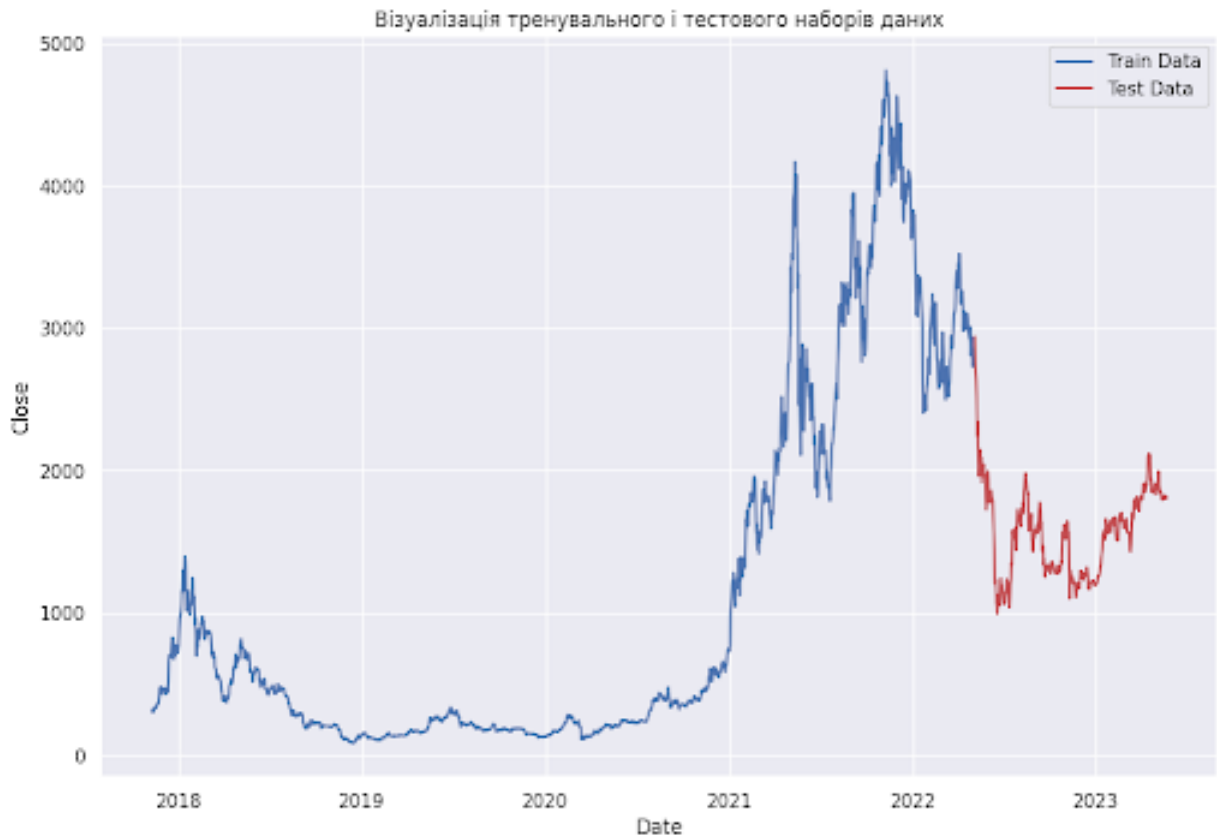


Рисунок 3.3 – Візуалізація тренувального та тестового наборів даних.

3.2.1 Пошук аномалій

Далі знаходимо аномалії. Створюємо індикатор аномалії. На основі порівняння дійсних значень ціни (стовпець y) з нижньою межею прогнозного діапазону (стовпець $yhat_{lower}$) та верхньою межею прогнозованого діапазону (стовпець $yhat_{upper}$) індикатор дозволяє виділити дні, коли спостерігаються аномальні значення ціни, що виходять за межі прогнозного діапазону.

В тестувальному наборі даних знайдено 65 аномалій.

Записуємо дані про знайдені аномалії, у датасет, з яким будемо працювати над торговельною стратегією.

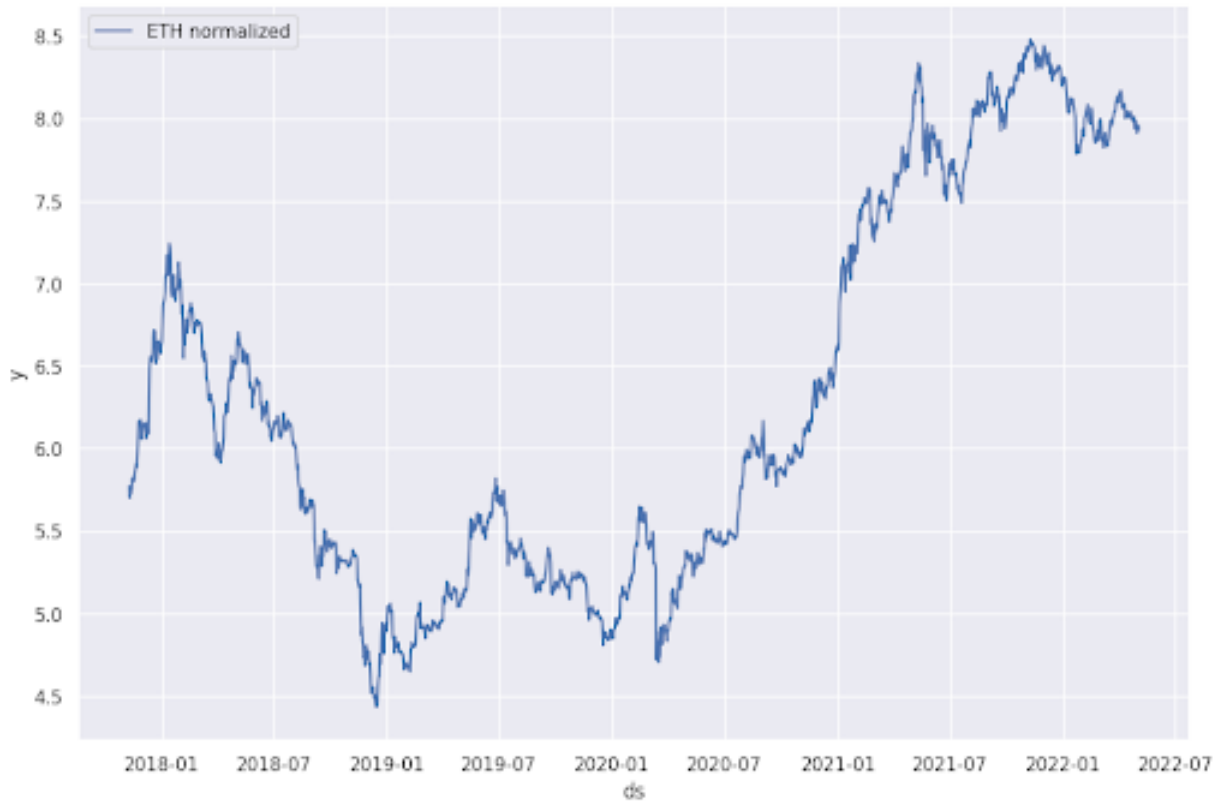


Рисунок 3.4 – Нормалізовані дані.

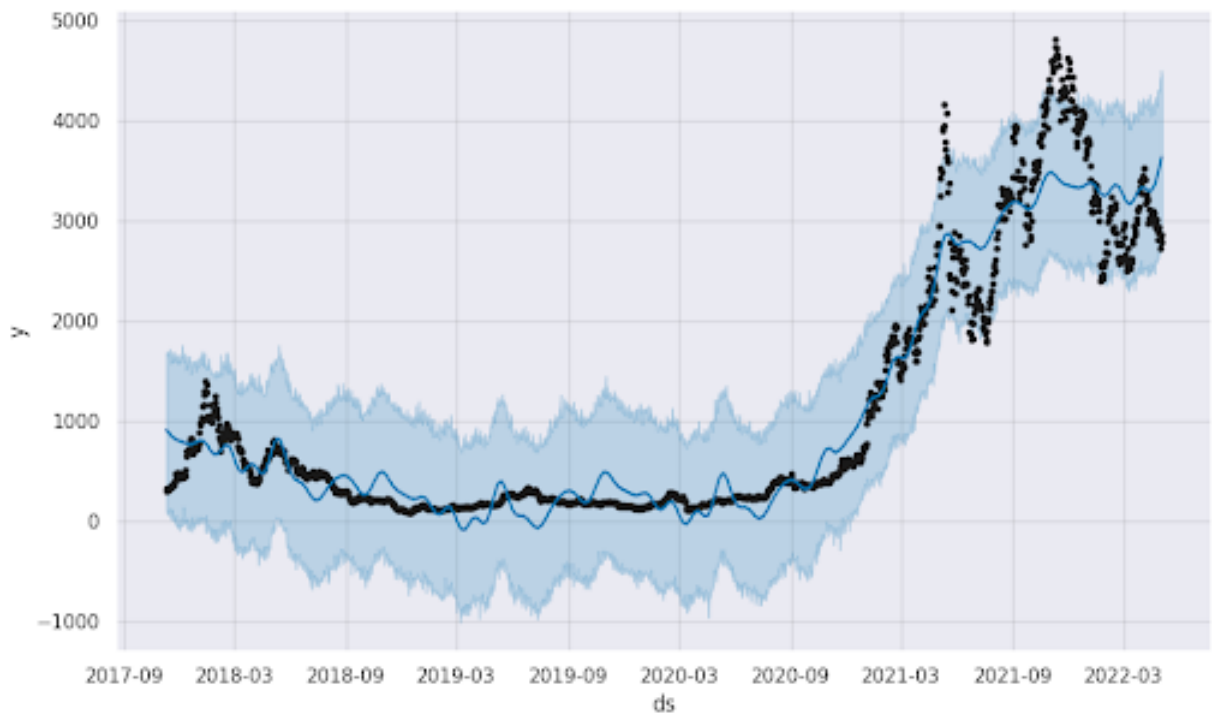


Рисунок 3.5 – Графік прогнозування з використанням моделі Prophet.

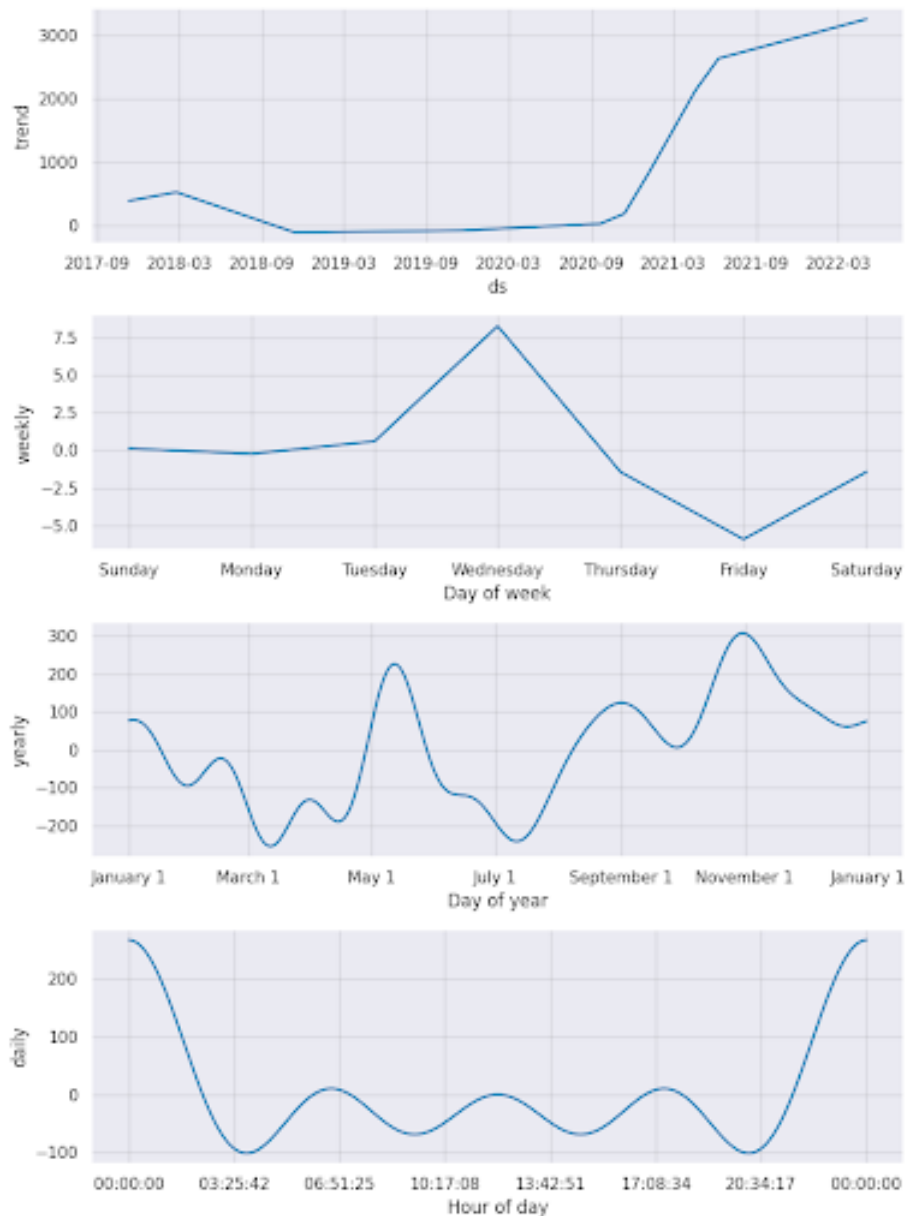


Рисунок 3.6 – Графік сезонності.

3.3 Торговельна стратегія

Запропонований алгоритм використовує інформацію про аномалії, волатильність, показник Шарпа та ціни закриття для прийняття рішення щодо купівлі та продажу активів. За умови виявлення аномалії та виконання відповідних умов для купівлі або продажу, алгоритм здійснює відповідну операцію. Цей підхід дозволяє використовувати інформацію про аномалії та інші показники для прийняття рішень з покупки та

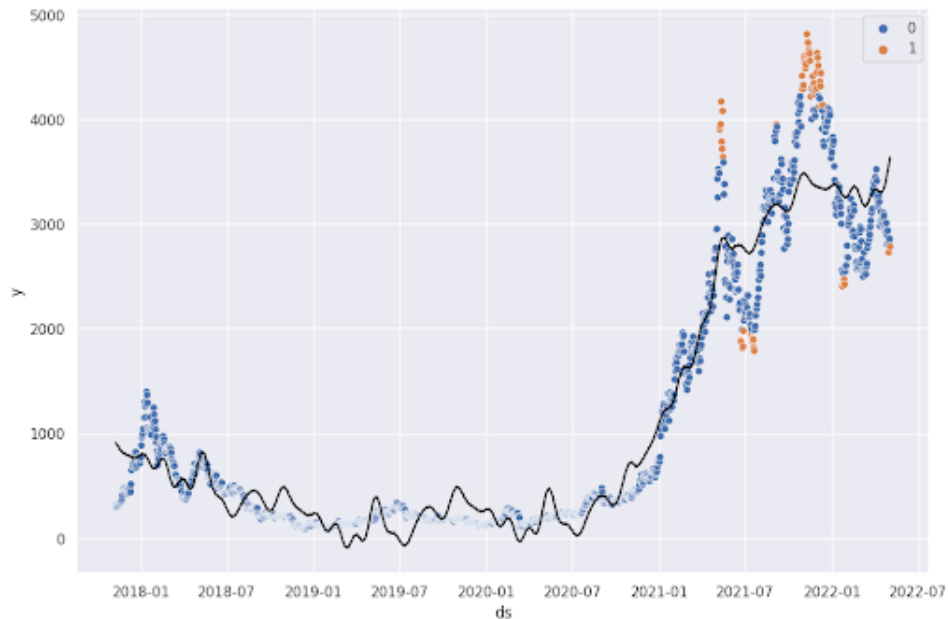


Рисунок 3.7 – Графік знаходження аномалій.

продажу активів.

Алгоритм, що наведений, виконує наступні дії:

- 1) Фільтрація даних за певний період. А саме: беремо початкову дату 2021-11-0, та кінцеву дату – 2023-05-31.
- 2) Обчислення волатильності на основі котирувань закриття цін.
- 3) Обчислення середньої денної доходності (приросту) на основі котирувань закриття цін.
- 4) Обчислення середньої річної доходності на основі середньої денної доходності.
- 5) Обчислення стандартного відхилення річної доходності на основі середньої денної доходності.
- 6) Обчислення показника Шарпа на основі середньої річної доходності та стандартного відхилення.
- 7) Додавання показників Шарпа і волатильності до датасету.
- 8) Визначення умов для купівлі і продажу, які включають умову про наявність аномалії, порівняння ціни закриття наступного дня після аномалії з ціною закриття в день аномалії та перевищення волатильності певного порогу.

- 9) Ініціалізація змінних для балансу, обсягу утримання активів, ціни купівлі та ціни продажу.
- 10) Проходження по кожному рядку датасету.
- 11) Розрахунок волатильності та показника Шарпа на основі вікна попередніх даних.
- 12) У разі виконання умови для купівлі, здійснюється покупка активів.
- 13) У разі виконання умови для продажу, здійснюється продаж активів.
- 14) Обчислення прибутку на основі зміни балансу.
- 15) Виведення початкового балансу, кінцевого балансу та прибутку.

3.4 Порівняння результатів моделі

Початкові результати моделі були покращені за допомогою: нормалізації даних, зменшення параметру *interval-width*, який визначає ширину довірчого інтервалу для прогнозу.

Значення цього параметру лежать в діапазоні від 0 до 1 і відображає відсоток довірчого інтервалу, в цьому випадку найкращий результат було визначено при $interval-width = 0.95$, тобто 95%.

Розраховуємо відсоток заробітку кожної моделі. Відсоток заробітку = $(\text{Прибуток} - \text{Початкова ціна}) / \text{Початкова ціна} * 100$

$interval - width = 0.95$: Відсоток заробітку = $(9314.26 - 100) / 100 * 100 = 9214.26\%$

При $interval - width = 0.97$: Відсоток заробітку = $(660.73 - 100) / 100 * 100 = 560.73\%$

При $interval - width = 0.99$: Відсоток заробітку = $(196.22 - 100) / 100 * 100 = 96.22\%$

Також було проведено порівняння запропонованої моделі, з моделлю, яка просто закупає на початку вказаного періоду на всі гроші ЕТН, задля того, щоб спростувати твердження, що прибутковість моделі

interval_width	Початковий баланс:	Кінцевий баланс:	Прибуток:
0.95	100	9314.26	9214.26
0.97	100	760.73	660.73
0.99	100	296.22	196.22

Таблиця 3.1 – Результати торговельної стратегії з використанням моделі Prophet.

присутня тільки через те, що ETH завідома прибуття криптовалюта.

Початковий баланс:	Кінцевий баланс:	Прибуток:
100	102.94	2.94

Таблиця 3.2 – Пасивний прибуток.

У такому випадку прибуток стратегії складає 3%, що свідчить про те, що запропонована модель є ефективнішою за припущення.

Висновки до розділу 3

За результатами проведених досліджень, можна зробити наступні висновки.

Модель Prophet показала ефективність у виявленні аномалій в часових рядах. Вона здатна ідентифікувати високі та низькі значення, що відхиляються від очікуваного шаблону, та маркувати їх як потенційні аномалії.

Модель Prophet може бути корисною у трейдингу, оскільки її

виявлені аномалії можуть використовуватися для прийняття торговельних рішень. Наприклад, умова про купівлю або продаж активу на основі ціни закриття наступного дня після аномалії може допомогти уникнути негативних фінансових наслідків.

Отримані результати підтверджують ефективність запропонованого алгоритму для виявлення аномалій в цінах активів. Шляхом порівняння фактичних значень ціни з нижньою та верхньою межами прогнозованого діапазону, було створено індикатор аномалії, що дозволив виділити дні, коли спостерігалися значення ціни, що виходили за межі прогнозного діапазону.

Даний алгоритм було поєднано з іншими показниками, такими як волатильність, показник Шарпа та ціни закриття, для прийняття рішень щодо купівлі та продажу активів. Такий підхід дозволив використовувати інформацію про аномалії та інші показники для прийняття обґрунтованих рішень з покупки та продажу активів.

Отже, застосування запропонованого алгоритму разом з оптимізацією параметрів моделі та використанням додаткових показників дозволило покращити результати моделювання, що дозволяє використовувати цю модель.

ВИСНОВКИ

У даній дипломній роботі було досліджено простір децентралізованих фінансів (DeFi) і розглянуто його особливості. Особлива увага була приділена Ethereum як провідній платформі для розвитку децентралізованих додатків, а також його криптовалюти, Ethereum. Також було вивчено торговельні стратегії і важливість розробки ефективних стратегій у просторі DeFi. В рамках дослідження було детально розглянуто поняття часових рядів, їх особливості, методи аналізу та виявлення аномалій.

Використання моделі Prophet для пошуку аномалій виявилось перспективним підходом. Модель Prophet дозволяє прогнозувати часові ряди з урахуванням сезонності та знаходити аномалії в цих рядах. В даній роботі модель Prophet була успішно використана для виявлення аномалій у фінансових даних DeFi.

На основі інформації про аномалії було побудовано торговельний алгоритм, який використовує ці дані для прийняття торговельних рішень. Цей підхід показав потенціал у створенні торговельних стратегій, що збільшують прибуток шляхом використання інформації про аномалії.

Отже, модель, розроблена у даній дипломній роботі, має потенціал стати основою для створення трейдингових стратегій, що збільшують прибуток в просторі DeFi за рахунок використання інформації про аномалії. Дана модель може бути використана трейдерами та інвесторами для прийняття обґрунтованих торговельних рішень і покращення їх фінансових результатів.

Однак, важливо врахувати, що модель є лише інструментом, і успішність торговельної стратегії залежить від багатьох факторів, таких як волатильність ринку, ризику, ліквідність активів та ефективне управління ризиками. Тому рекомендується поєднувати дану модель з додатковими аналітичними інструментами та професійним досвідом для

досягнення кращих результатів.

У майбутньому можливо покращення моделі шляхом оптимізації архітектури та додавання додаткових умов трейдингу. Також, можливі розширення дослідження можуть включати в себе розгляд інших методів пошуку аномалій, використання більш складних алгоритмів прогнозування та інтеграцію з іншими факторами, що впливають на ринок DeFi.

У цілому, дана дипломна робота внесла важливий внесок у розвиток області торговельних стратегій в просторі DeFi. Використання моделі Prophet для пошуку аномалій та побудови трейдингового алгоритму відкриває нові можливості для трейдерів та інвесторів у цьому швидко зростаючому секторі.

ПЕРЕЛІК ПОСИЛАНЬ

1. Sharma, Rakesh. What Is Decentralized Finance (DeFi) and How Does It Work? / Rakesh Sharma // Investopedia. — 2022. Режим доступу: <https://www.investopedia.com/decentralized-finance-defi-5113835>.
2. Popper, Nathaniel. What Is DeFi? A Guide to Decentralized Finance / Nathaniel Popper, Erin Griffith // The New York Times. — 2022. Режим доступу: <https://www.nytimes.com/interactive/2022/03/18/technology/what-is-defi-cryptocurrency.html>.
3. DLT, Radix. Why DeFi is the Future of Exchanges and Trading. Режим доступу: <https://www.radixdlt.com/blog/why-defi-is-the-future-of-exchanges-and-trading>.
4. Metelski, Dominik. Decentralized Finance (DeFi) Projects: A Study of Key Performance Indicators in Terms of DeFi Protocols' Valuations / Dominik Metelski, Janusz Sobieraj // International Journal of Financial Studies. — 2022.
5. Ethereum. Ethereum Gas Documentation. — <https://ethereum.org/uk/developers/docs/gas/>. — N/A.
6. Ethereum. Gas. — N/A. Режим доступу: <https://ethereum.org/uk/developers/docs/gas/>.
7. Hazes, Adam. Trading Strategy / Adam Hazes, Somer Anderson, Kristen Schmitt // Investopedia. — 2023. Режим доступу: <https://www.investopedia.com/terms/t/trading-strategy.asp>.
8. Konchar, Phillip. Why Do Traders Need a Strategy? / Phillip Konchar. — 2023. Режим доступу: <https://mytradingskills.com/why-do-traders-need-a-strategy>.
9. CMC Markets. Trading Strategies. — 2023. Режим доступу: <https://www.cmcmarkets.com/en/trading-guides/trading-strategies>.

10. MOVING AVERAGE STRATEGY GUIDE – 5 MOVING AVERAGE STRATEGIES. — 2019. Режим доступа: <https://www.tradingstrategyguides.com/moving-average-strategy-guide/>.

11. CFI Team. Trading Strategy. — 2021. Режим доступа: <https://corporatefinanceinstitute.com/resources/capital-markets/trading-strategy/>.

12. Components of Time Series Data. Режим доступа: <https://www.geeksforgeeks.org/components-of-time-series-data/>.

13. Lazzeri, Francesca. Introduction to feature engineering for time series forecasting / Francesca Lazzeri // Data Science at Microsoft. — 2021. Режим доступа: <https://towardsdatascience.com/introduction-to-feature-engineering-for-time-series-forecasting-afd4c5f10270>.

14. Panwar, Abhash Anil. Introduction to Time Series Analysis and key concepts / Abhash Anil Panwar // Analytics Vidhya. — 2021. Режим доступа: <https://www.analyticsvidhya.com/blog/2021/01/introduction-to-time-series-analysis-and-key-concepts/>.

15. Larsson, Frans. Algorithmic trading surveillance / Frans Larsson. — 2019. Режим доступа: <https://www.diva-portal.org/smash/get/diva2:1339906/FULLTEXT01.pdf>.

16. Chandola, V. Anomaly detection: A survey / V. Chandola, A. Banerjee, V. Kumar // ACM Computing Surveys (CSUR). — 2009. — Vol. 41, no. 3. — Pp. 15:1–15:58.

17. Facebook. Facebook Prophet Documentation. — <https://facebook.github.io/prophet/>. — N/A.

18. Zhou, You. Study on Natural Gas Price Forecasting Based on Prophet-GRU Nonlinear Combination / You Zhou, Shuai Chen, Di Xiao // 7th International Conference on Computational Intelligence and Applications / IEEE. — 2022. — P. N/A.

19. Taylor, S. J. Forecasting at scale / S. J. Taylor, B. Letham // The American Statistician. — 2018. — Vol. 72, no. 1. — Pp. 37–45.